

Book of abstracts

**EuroSLA33**

The 33<sup>rd</sup> Conference of  
the European Second Language Association

3–6 Jul 2024, Montpellier



[eurosla33.sciencesconf.org](https://eurosla33.sciencesconf.org)



UNIVERSITÉ DE  
MONTPELLIER



Conception et impression : Imprimerie DMG - Université Paul-Valéry – Montpellier 3

# Exploring the Interplay of L2 Learning Experience and Executive Functions: A Longitudinal Study

Mehmet Akinci \* <sup>1</sup>, Gülcan Ercetin \*

2

<sup>1</sup> İstanbul 29 Mayıs University – Turkey

<sup>2</sup> Boğaziçi University – Turkey

Executive functions (EFs) referring to the processes necessary for regulating thoughts and behaviors (Miyake & Friedman, 2012) play a crucial role in a myriad of real word tasks (Diamond, 2013; Engle, 2018) including second language (L2) success. Particularly, inhibition and working memory (WM) represent two major and extensively studied components of EFs in relation to L2 with neurological and behavioral studies highlighting their impact on L2 success (e.g., Ghaffarvand Mokari & Werner, 2019; Linck, Osthus, Koeth, & Bunting, 2014). While individual differences in EFs can predict L2 success, it is noteworthy that L2 experience could also have facilitative effects on EFs (Grundy, 2020; Grundy & Timmer, 2017).

The joint activation of two languages in a bilingual mind demands intensive and frequent use of cognitive control processes for effective language management by maintaining attention to the target language while avoiding interference from the competing one (Bialystok, Craik, Green, & Gollan, 2009). However, scant research has explored the relationship between EFs and L2 learning experience particularly as of/after puberty, and the extant behavioral studies present inconclusive findings alongside methodological challenges. To address this gap, the current study aims to investigate the longitudinal effects of L2 learning experience on EFs and the predictive validity of EFs for L2 outcomes. Employing a longitudinal pre-test/post-test design involving both EFs tasks and an L2 proficiency test, this approach would enable to unveil the underlying mechanisms through establishing causal effects and reducing confounds (Gallo & Abutalebi, 2023).

Participants in the experimental group comprised 165 Turkish high-school graduates undergoing intensive L2 instruction in a university setting for six months (600 class contact hours), and the control group included 103 freshman students taking undergraduate courses in their first language. Two visual complex span tasks (symmetry and rotation) to measure working memory, antisaccade and flanker tasks to measure inhibition, and one standardized English proficiency task were administered before and after the instruction. Multivariate and univariate repeated measures ANOVA results indicated that both instruction types led to the enhancement of EFs except for flanker performances, yet no group differences were observed. Exploratory factor analysis revealed that complex span task and antisaccade performances had loadings on the same factor, named as executive attention, whereas flanker performance was independent from this dimension. Regression findings demonstrated that, albeit little, flanker but not executive attention could explain L2 success. Prior L2 experience was found to be the best predictor. The study concluded that L2 experience can contribute to the enhancement of EFs, yet executive

---

\*Speaker

attention might not play a significant role in L2 success at the end of a six-month intensive L2 instruction.

Various paradigms encompassing domain-general and domain-specific perspectives and considerations related to instruction could contribute to a comprehensive understanding of the results. The study also offers valuable conceptual and methodological implications along with critical insights for further research.

**Keywords:** Executive functions, Bilingual advantage, L2 experience, Working memory, Inhibition, L2 instruction

# Studying Language Control in High Frequency Code-Switchers: Combining Experimental and Ecological Approaches

Layana Awada <sup>\*† 1</sup>, Barbara Köpke <sup>1</sup>

<sup>1</sup> Université Toulouse - Jean Jaurès – Université de Toulouse, Laboratoire de neuropsycholinguistique – France

For bilingual speakers, producing a word in a specific language requires to select one language and inhibit the other, therefore implying a certain level of control. This has led to modeling language control (LC) taking into account code-switching (CS) (*e.g.*, *Inhibitory Control Model*, Green, 1986). Nevertheless, the interaction between switching and control remains poorly explored, especially the influence of CS on LC. In their *Control Process Model of CS* (CPM), Green and Wei (2014) suggest a LC scale ranging from open to competitive control, allowing for more or less CS to take place in the bilingual’s discourse. Furthermore, the sociolinguistic context has an impact on the use of CS, as claimed in Green’s *Adaptive Control Hypothesis* (2013). Green suggests three different interactional contexts, among which dense CS, which calls for an open control mode. Dense CS is common in bilingual communities, where speakers frequently mix languages in the course of a single utterance. This is the case of the Lebanese community, known for frequently using Lebanese Arabic (LA), French (FR) and/or English within the same discourse, sentence, and even word (Bassam, 2022).

Recent research on the role of CS in LC takes place in laboratory settings, involving experimental tasks measuring the cost of switching languages (switch and mixing costs) (Declerck & Koch, 2023), while fewer studies explore CS in an ecological environment. In the current research, we aim to conciliate experimental and ecological domains within an integrative protocol that helps to better understand LC abilities of frequent code-switchers, hence French speaking Lebanese bilingual adults.

The *LEAP-Q* questionnaire (Marian et al., 2007), *BLP* (Birdsong et al., 2012) and *BCSP* (Olson, 2022) are used to establish type of bilingualism, dominance and CS habit respectively, L2 (FR) proficiency is assessed using a language test (CEL, 2018).

The experimental protocol, undergone by 50 LA/FR bilingual adults, involves assessing three main domains: discursive, lexical and cognitive.

Discursive: An ecological discussion with multiple speakers, where language mode (Grosjean, 2008) is manipulated, in order to test the participants’ adaptation to their respective interlocutors’ language knowledge.

Lexical: A bilingual picture-naming task, following Costa and Santesteban’s (2004) cued language-switching paradigm for LA and FR, allowing us to measure switch costs and

---

\*Speaker

†Corresponding author: layana.awada@univ-tlse2.fr

mixing costs.

Executive: Assessment of inhibitory processes with a computerized version of the *Stroop test*, the *Flanker task*, the *Wisconsin Card Sorting Test* and digit-span.

We will present the protocol and some first results. Among the hypotheses, we expect to find symmetrical switch costs in highly proficient French speakers, as L2 proficiency was found to reduce switch cost asymmetry (Costa & Santesteban, 2004). In addition, a mixing benefit is expected for voluntary switch tasks, as suggested by De Bruin et al. (2018), who state that voluntarily using two languages can be less costly than having to stay in one language. *In fine*, executive functioning, in terms of inhibition, shifting and updating (Miyake et al., 2000), will be correlated to language switching, allowing us to disentangle domain-general and domain-specificity hypotheses (Bobb, Wodniecka & Kroll, 2013).

**Keywords:** bilingualism, code, switching, language control, French, Lebanon

# Chinese and Japanese L2 learners’ perceptions of the use of generative AI in language learning

Nozomi Abe <sup>\*† 1</sup>, Xiaoli Wu <sup>\* ‡ 1</sup>

<sup>1</sup> Cardiff University – United Kingdom

The integration of artificial intelligence (AI) in language learning has emerged as a transformative force in the realm of second language acquisition (SLA), presenting both unprecedented opportunities and potential challenges. This research navigates the dynamic intersection of generative AI and language education, focusing on the specific context of Chinese and Japanese second language (L2) learners.

In response to the escalating influence of AI in academia, the Russell Group, a consortium of leading UK universities, has endorsed a set of principles designed to cultivate AI literacy among staff and students. These principles underscore the importance of AI-literacy support, ethical considerations, equal access, academic rigor, and collaborative best practices (Russell Group, 2023). Our study aligns with these principles, seeking to delve into the nuanced landscape of AI-supported language learning and its implications for Chinese and Japanese L2 learners. While existing studies have explored similar themes in the context of European languages (Dakakni and Safa, 2023; Chan and Hu, 2023), the unique writing systems of Chinese and Japanese and different learner’ motivations might pose distinctive challenges and opportunities in the integration of AI tools. The current study contends that a more language-specific approach is imperative.

This study aims to bridge existing gaps in literature by conducting an investigation into the perceptions of generative AI-supported language learning among Chinese and Japanese second language (L2) learners. Our research employs a mixed-methods approach, combining an online survey with additional in-depth interviews, providing a holistic view of participants’ perspectives. The online survey, adapted from Chan and Hu’s comprehensive questionnaire (2023), includes demographic information such as gender and language proficiency levels, an assessment of participants’ knowledge of Generative AI technologies like ChatGPT, an exploration of perceptions regarding generative AI tools, and an examination of the ethical considerations associated with their use. The survey encompasses a sample of 50 Chinese and Japanese L2 learners across different academic years within a UK university. Following the online survey, a select group of participants expressing voluntary interest is invited to engage in more in-depth interviews. This intentional sampling strategy allows for a more nuanced exploration of L2 learners’ experiences, expectations and concerns.

This proposed poster presentation aims to share our research findings on the current perceptions of Chinese and Japanese second language (L2) learners concerning generative AI-supported language learning. This type of investigation is timely in an era characterized by technological fluidity and uncertainty, where advancements in AI are rapid and profound. In many ways, we are standing at a crossroad of changes and for this reason, understanding current perceptions of students will be a great starting point for language instructors to update our teaching methods,

---

\*Speaker

†Corresponding author: AbeN@cardiff.ac.uk

‡Corresponding author: wuxla@cardiff.ac.uk

tailor more engaging and relevant materials and enhance students' learning experience. It will also help us identify what kinds of ethical considerations are required in this AI-driven world we live in. By sharing our research outcomes, we aim to contribute to the ongoing discourse surrounding the ethical use of AI, inform instructional practices, and guide future investigations in the ever-evolving landscape of second language acquisition.

**Keywords:** Artificial Intelligence (AI), L2 learning at HE, ChatGPT, ethical considerations, Chinese and Japanese L2 learners

# Acquisition of phraseological competences and the spacing effect

Fredina Addo \* <sup>1</sup>

<sup>1</sup> Métropole Nice Côte d'Azur – University Côte d'Azur – France

The spacing effect is a phenomenon that has been extensively researched and described as robust in cognitive psychology (Gerbier & Toppino, 2015). It refers to the benefits that accrues to retention from engaging in distributed as opposed to massed practice (Nakata & Elgort, 2021). Its robustness has also been attested to in the fields of cognitive neuroscience, music, genetics, among others. The spacing effect is also being studied in second language acquisition (SLA) with authors focusing on various aspects of language learning. The studies so far have tested the effects of different spaced conditions on the learning of L2 grammar (Bird, 2009; Kasprowicz, 2019; Serfaty & Serrano, 2020, 2022; Suzuki, 2017); vocabulary knowledge (Nakata & Elgort, 2020; Rogers & Cheung, 2018, 2020; Serfaty & Serrano, 2023); as well as on more global measures of L2 knowledge (e.g., Collins & White, 2011; Serrano & Muñoz, 2007; Serrano, 2012). These studies have thus far presented mixed results, there is therefore the need for more research on this phenomenon.

Until recently, most studies investigating the effect of spacing on L2 vocabulary acquisition have concentrated on the acquisition of single words and at present, the results of these studies appear robust. However, the acquisition of the phrasal lexicon under spaced conditions may yield different results owing to the complex nature of their form and meaning. Currently, there are only a few studies that have been concerned with the effect of spacing on the phrasal lexicon (e.g., Farvadin, 2019; Ferguson et al, 2021; Macis et al., 2021; Snoder, 2017; Yamagata et al., 2022). Although all these studies have focused on collocations, results from the studies on spacing and collocational knowledge have yielded conflicting results. More research into this effect is thus needed. Hence, our thesis aims to contribute to this aspect of vocabulary knowledge by investigating the effect of spacing on the learning of two phraseological units (collocations and conventional expressions) in French as a foreign language (FLE). Our motivation for also including conventional expressions stems from claims of how acquisition of these expressions is invaluable to social communication (Bardovi-Halig & Vellenga, 2012). Besides, to the best of our knowledge, no study on the spacing effect has dwelled on conventional expressions.

Following a pretest-posttest design with a pedagogical intervention, we are adopting a phraseo-didactic approach focused on the acquisition of phraseological skills for collocations and conventional expressions. 5 learning sessions will be organized according to two spacing schedules: 1 day and 8 days. Retention will be tested via an immediate and two delayed posttests. The planned participant pool will be third-year Ghanaian trainee teachers in the FLE stream. Our work seeks to probe how effective this teaching intervention will be on developing participants' lexical knowledge of the taught phraseological units, and in what way spacing influences the benefits gained from the teaching intervention. We will present the methodological aspect of this ongoing thesis.

---

\*Speaker



**Keywords:** Phraseology, Teaching FLE, Spacing Effect, Trainee teachers

# Writing development in adult L2 learners of varying educational backgrounds

Anders Agebjörn \*<sup>1</sup>, Robert Walldén<sup>1</sup>

<sup>1</sup> Malmö Höskola = Malmö University – Sweden

Second-language (L2) research, including L2 writing research, primarily recruits participants from academic contexts (Pettitt et al., 2021). Existing studies involving adult L2 learners with limited schooling typically focuses on teaching practices rather than on language or writing development (Nordanger et al., 2023). There is an increasing awareness that such sampling biases and limited scopes is detrimental both to the generalizability of emerging L2 theories and to the applicability of research findings in heterogeneous L2 classrooms (Godfroid & Andringa, 2023). To our knowledge, previous research has not systematically compared L2 writing development among adult learners with varying educational backgrounds. The present study addresses this knowledge gap by examining the relationship between educational background and L2 writing development in diverse groups of adult L2 learners.

The participants were 38 adults enrolled in course C or D in Swedish for Immigrants (SFI). These two courses target proficiency levels corresponding approximately to the A2 and the B1 levels of the Common European Framework of Reference. Depending on their educational background, the participants followed three different study paths. On two occasions separated by three months, they completed a communicative writing task. The texts were assessed holistically by five experienced L2 teachers using the method of Comparative Judgement (Pollitt, 2012). The assessment reached an excellent level of reliability ( $SSR = .95$ ).

A mixed-effects regression model, including course, study path, and data point as fixed effects, revealed that educational background was positively associated both with the participants' general writing ability and with their rate of L2 writing development. This result confirms that educational background is a factor that must not be neglected in L2 research. However, a large intra-individual variation among the participants with the most limited schooling indicates that the writing task used may not be a valid measure of L2 writing development in this population. Such methodological challenges are discussed (cf. Gujord Helland et al., 2023). Furthermore, we highlight implications for curriculum design and further research in diverse L2 classrooms.

## References

Godfroid, A., & Andringa, S. (2023). Uncovering sampling biases, advancing inclusivity, and rethinking theoretical accounts in Second Language Acquisition: Introduction to the special issue *SLA for All?*. *Language Learning*. <https://doi.org/10.1111/lang.12620>

Gujord, A. K. H., Søfteland, Å., & Emilsen, L. E. (2023). Researching the language development of learners with limited schooling – methodological choices and challenges. *Nordand*, 18(2), 125–139. <https://doi.org/10.18261/nordand.18.2.4>

---

\*Speaker

Nordanger, M., Carlsen, C. H., & Bugge, E. (2023). Forskning på samtidig andrespraks- og literasitetstillegg. En sonderende oversikt over LESLLA-nettverkets konferansepublikasjoner. In: Bugge, E. & Carlsen, C. H. (Eds.), *Norsk som andresprak – voksne innvandrere som utvikler skriftkyndighet på et andresprak*. Cappelen Damm Akademisk, pp. 21–48.

Pollitt, A. (2012). The method of adaptive Comparative Judgement. *Assessment in Education: Principles, Policy & Practice*, 19(3), 281–300. <https://doi.org/10.1080/0969594X.2012.665354>  
Pettitt, N., Gonzalves, L., Tarone, E., & Wall, T. (2021). Adult L2 writers with emergent literacy: Writing development and pedagogical considerations. *Journal of Second Language Writing*, 51, 1–5. <https://doi.org/10.1016/j.jslw.2021.100800>

**Keywords:** L2 writing, Comparative Judgement, Educational Background, basic L2 education, Swedish for Immigrants

# Investigating the contribution of accuracy and complexity in characterizing language proficiency: a machine learning experiment

Soroosh Akef <sup>\*† 1,2</sup>, Detmar Meurers<sup>‡ 2,3,4</sup>, Amália Mendes<sup>§ 1</sup>, Patrick Rebuschat<sup>¶ 2,5</sup>

<sup>1</sup> Center of Linguistics of the University of Lisbon – Portugal

<sup>2</sup> LEAD Graduate School and Research Network – Germany

<sup>3</sup> Leibniz-Institut für Wissensmedien (IWM) – Germany

<sup>4</sup> University of Tübingen – Germany

<sup>5</sup> Lancaster University – United Kingdom

The complexity, accuracy, and fluency (CAF) triad introduced by Skehan (1989) has become a prominent conceptualization of language proficiency. The CAF constructs are multidimensional and interrelated (Norris & Ortega, 2009), so much so that even expert teachers find it difficult to disentangle them (Riemenschneider et al, 2019). But while the conceptualization and operationalization of complexity have received substantial attention in recent years (cf. Kuiken, 2022), how to combine and balance those aspects with an equally rich characterization of accuracy is underexplored. The current study uses machine learning as an experimental test bed to investigate the following research question: What kind of evidence on linguistic complexity and accuracy and their interaction is informative in characterizing proficiency?

To explore this question, we distinguish 49 fine-grained types of grammatical, orthographical, lexical, and discourse errors in the 1634 texts of the Portuguese learner corpus COPLE2 (Mendes et al., 2016) based on its token-level error annotations. We calculated both absolute counts and length-independent ratios of errors per linguistic unit (token, phrase, clause). For complexity, we computed 489 measures, encompassing syntactic, lexical, morphological, discourse, and psycholinguistic features using CTAP (Chen & Meurers, 2016). Based on overall 589 features, we investigated predicting the proficiency level of the learner (A1–C1) as specified in the corpus meta-information.

We trained random forest classifier models for a three-class task (A, B, C) and models distinguishing all five levels. The three-class model reached an accuracy of 74.66% (10-fold cross-validation), outperforming the previous best proficiency classification result of 72% (del Río Gayo, 2019). Accuracy features by themselves are predictive, but much less so than complexity features: our three-class model trained only on the accuracy features reached 62.18%.

Investigating which features are the most informative for the five-class classifier using the feature importance method in scikit-learn (Pedregosa et al., 2011), we found that among the top-20 features, lexical sophistication, lexical richness, and morphological inflection features dominate. Among the accuracy measures, the overall token accuracy rate ( $\#errors/\#tokens$ ) is among the

---

\*Speaker

†Corresponding author: sorooshakef@edu.ulisboa.pt

‡Corresponding author: dm@sfs.uni-tuebingen.de

§Corresponding author: amaliamendes@letras.ulisboa.pt

¶Corresponding author: p.rebuschat@lancaster.ac.uk

ten most informative features, but strikingly none of the more specific error features is even among the top-200. More fine-grained analyses for binary classifiers distinguishing each proficiency level show that the token accuracy rate drops from A1 to A2, followed by an increase from B1 onwards. One way to interpret this is that increasingly elaborate and varied structures become available at A2 compared to A1, which then are more accurately mastered later – the qualitative discussion will explore such issues in more detail. Relatedly, the observation that the more fine-grained accuracy measures were not informative deserves more attention. Complementing the above approach with statistical modeling using generalized additive models will make it possible to zoom into interaction effects between complexity and accuracy features, even where they are non-linear as in the development of the token accuracy measure. While the results obtained so far constitute only a first step towards understanding the interplay at work here, they clearly illustrate the need for more research empirically analyzing and interpreting multi-layered models of complexity and accuracy together.

**Keywords:** CAF, linguistic complexity, accuracy, learner corpora, intelligent computer assisted language learning, automatic proficiency assessment

# The impact of podcast creation on EFL students' language learning emotions

Mari Alger <sup>\*†</sup> <sup>1</sup>, June Eyckmans <sup>1</sup>

<sup>1</sup> Universiteit Gent = Ghent University – Belgium

This study focuses on an out-of-class podcast project undertaken by second-year Bachelors students enrolled on an English as a Foreign Language (EFL) vocabulary course in Belgium. The aim of the project was to provide students with the opportunity to use English more frequently and authentically while creating something tangible together through interaction. Research is beginning to emerge demonstrating the advantages of engaging students in collaborative podcast production, such as an improvement in spoken performance and a reduction in foreign language anxiety. This study contributes to the current wave of applied linguistics research which aims to explore learners' foreign language learning experiences through the coexistence of multiple emotions. We present data from 38 students who were tasked with creating a podcast in small groups. The first set of data comprised students' perceptions of making a podcast at two time intervals to examine emotions 1) in anticipation of the project and 2) upon completion. The second set – an online survey – contained Likert-scale items from validated foreign language anxiety and enjoyment instruments as well as open-answer questions to gain insights on the wider-ranging emotions students experienced during 1) their formal language learning education and 2) the podcast project. The thematic analysis guided by the control-value theory of achievement emotion identifies which particular aspects of the podcast project invoked emotions in students to better understand what worked, in what situations, and why. Furthermore, the approach of comparing two time periods and two conditions helped us to demonstrate how emotions evolve across time and context, thus informing pedagogical implications especially with regards to teaching a language-focused podcast project. In response to the need for increased researcher reflexivity and interdisciplinary orientation in qualitative research, the methodological challenges we encountered while investigating something as interpretative and multidimensional as emotions are also deliberated.

**Keywords:** emotions, qualitative research, mixed methods, language learning and technology, podcast

---

\*Speaker

†Corresponding author: mari.alger@ugent.be

# The acquisition of L2 French phonology by L1 Portuguese adult learners

Letícia Almeida <sup>\*† 1</sup>, Charlotte Alazard-Guiu <sup>\* ‡ 2</sup>, Margarida Possidónio

1

<sup>1</sup> University of Lisbon, Faculty of Arts, Center of Linguistics – Portugal

<sup>2</sup> Octogone-Lordat, Université Toulouse Jean Jaurès – OCTOGONE-Lordat – France

Despite the growing interest for research on second language pronunciation (Suzukida & Saito, 2022; Santiago et al, 2022) and pronunciation instructions (Darcy et al, 2021; Darcy, 2018), from our experience, phonetics is still not seen as a priority in second language classes, especially in French as a Foreign Language (FFL) course. Nevertheless, knowing the specific areas of difficulties of students according to their native language offers a valuable opportunity to focus on these specific areas in FFL classes. In this study, we focus on European Portuguese learners of L2 French, an under-studied language combination. To our knowledge, only one study has previously looked at this particular language combination and reported difficulties producing specific vowels, even at an advance level (Almeida, 2006).

In this pilot study, our primary goal is thus to make an assessment that measures the pronunciation skills of university students attending a FFL course in Portugal. We decided to compare two levels of acquisition: beginner and advanced in order to answer two main questions: What are the difficulties of L1 Portuguese learners of French at the pronunciation level? Which skills have improved and which skills remain unchanged with a higher level of French training?

We recorded 40 Portuguese L2 French learners who study at the University of Lisbon attending either an A2 (n = 20) or a B2 (n = 20) FFL class. We used two tasks: a reading task and a narrative task. We did both perceptive and acoustic analyses of the data. All the data were phonetically transcribed and acoustics analyses were performed on speech rhythm measures, disfluencies and vocalic measures. Additionally, the productions of the students were evaluated in terms of accentedness, comprehensibility and fluency by native raters, following (Bails et al, 2022) study.

The preliminary results suggest that the main difficulties attested in our corpus affect the pronunciation of anterior rounded vowels, absent in L1 inventory, as well as nasal vowels, which are attested with different phonetic properties in the students' L1. The pronunciation of nasal vowels improves with training: advance learners perform better than beginners. However, the difficulty with anterior rounded vowels is still present at the advanced level. The main difference between beginners and advanced students is the type of compensatory strategies rather than the numbers of errors: while beginners systematically substitute the anterior rounded vowels by a vowel from the L1 inventory, advanced learners are also capable of producing vowels that do not exist either in the L1 nor the L2 inventory, showing a high level of knowledge of the L2 and distancing from their L1. The different strategies will be further discussed in our presentation from both a research and a pedagogical point of view.

---

\*Speaker

†Corresponding author: lalmeida@letras.ulisboa.pt

‡Corresponding author: alazard@univ-tlse2.fr

**Keywords:** phonology, vowels, L1 Portuguese, L2 French, adult learners



# How Vulnerable are Verb-noun Collocations in Attrition? The Impact of Extralinguistic Factors on L2 Lexical Attrition among Saudi Arabic-English Returnees

Hadil Alraddadi \*<sup>1</sup>, Jeanine Treffers-Daller<sup>1</sup>

<sup>1</sup> University of Reading – United Kingdom

One of the key aims in studies of heritage speakers (HSs) is understanding which aspects of their linguistic repertoire are vulnerable to attrition. While most studies focus on heritage speakers' knowledge of their heritage language, much less attention has been paid to attrition in their second language (L2). In particular, knowledge and use of multiword units (MWUs) have hardly been investigated among HSs and returnees. It is well-known that MWUs are difficult for L2 learners because extensive naturalistic input is needed to master those. Much less is known about the vulnerability of MWUs under conditions of reduced input. Studying returnees who experience a drop in contact with their L2 after return to their home country can throw new light on the vulnerability of these constructions to attrition and the conditions under which attrition takes place. Clearly, the extent and rate of attrition can vary depending on extralinguistic factors, such as age at immigration or remigration, length of residence, and exposure to both languages. The few available studies of L2 attrition in returnees have focused on Japanese, Portuguese or Turkish HSs that returned to their countries of origin, but little is known about Saudi returnees, who grew up in the US but returned to their home country in early or late childhood. We study 44 returnees who had lived in the US for an extended period of time and returned to Saudi Arabia, either in early childhood or in adolescence. We compare the returnees to 54 US-based child and adolescent Saudi HSs and a group of 20 adult HSs to establish to what extent the target structures are normally acquired by HSs in the US. In total 118 bilinguals took part in the study. Productive knowledge of the collocations was measured with a novel online gap-filling task consisting of English verb-noun collocations which were either congruent, such as *'have experience'* or incongruent between Arabic and English, for example, *'do homework'*. In addition, an online lexical decision task focusing on verb-noun collocations was used to measure receptive collocational knowledge. A range of baseline tests were administered to test vocabulary and grammar knowledge, and the Bilingual Language Profile was used to collect background information. Following the recommendation in Matos and Flores (2022), mixed effect modelling was used to analyse the data. As predicted, returnees' performance in recognising non-existing English collocations was highly correlated with their degree of current use of English as well as their age of return and length of residence in the US and Saudi Arabia. The results revealed that returnees obtain significantly lower scores than HSs on gap-filling task. However, in the receptive task, the adolescent returnees perform at similar levels as adult and adolescent HSs. This could mean attrition was more pronounced among returnees in the productive task as opposed to the receptive task. The findings also indicate a strong influence of Arabic among returnees, possibly L2 attrition after return. We finish by formulating implications for the role of L2 attrition in HSs and returnees.

---

\*Speaker

**Keywords:** language attrition, vocabulary, collocations, heritage speakers, returnees

# Continuity and contrast in L2 Italian dialogues

Cecilia Andorno \* <sup>1</sup>, Christine Dimroth \* <sup>† 2</sup>, Ludovica Monte \*

<sup>1</sup>, Sandra Benazzo \*

3

<sup>1</sup> università di Torino – Italy

<sup>2</sup> Universität Münster – Germany

<sup>3</sup> Université de Vincennes – UMR 7023 Structures formelles du langage – France

Native speakers' preferences for discourse organisation are language-specific and reflect entrenched and economic solutions for context integration. Matching these preferences is a late achievement in L2 acquisition (Flecken et al. 2013). We study context integration in Italian dialogues and ask whether L2 learners with different learning profiles acquire the relevant preferences and whether their choices for adding information in continuity with the previous discourse differ from those adopted when contrasting information in reaction to their interlocutors.

In an elicited oral production study, speakers compared two slightly different pictures and identified a maximal number of differences. Participants were two groups of adult German learners of Italian (Group1: intermediate classroom learners in Germany, N=22; Group2: near-native speakers with  $\approx 20.3$  years of residence in Italy, N=20), as well as two groups of German and Italian native speakers (N=20 each).

For all speakers, utterances consisted of at least 2 information units: an NP referring to a depicted entity in utterance-final (focal) position; an anaphoric locative expression in pre-verbal position ('next to the X is an orange'). Reference to the speaker ('below that I have a teapot') is optional but nearly indispensable in contrastive contexts, when speakers indicate that the location specified by their interlocutor hosts a different entity in their own picture. Italian native speakers use three types of constructions with an increasing degree of contrastivity:

- Existentials without reference to the speaker; only in continuity contexts: *sotto c'è un barattolo blu*.

- Possessives with a slot for a subject pronoun; mostly left implicit in continuity contexts: *alla sinistra della candela ho un tubetto di dentifricio*.

- Possessives with overt pronoun and local adverbial in pre-verbal position; strongly preferred in contrastive contexts: *io di fianco al bicchiere ho una pallina da tennis*.

German speakers use copula constructions in all contexts; a post-verbal adverbial (V2-constraint) can express the contrastive information: *darunter ist bei mir ein Teelicht*.

---

\*Speaker

<sup>†</sup>Corresponding author: christine.dimroth@uni-muenster.de

Both learner groups show preferences conforming to their L1, as well as development toward the TL. Group 1 adopts target-like existentials for continuity but slides back to ungrammatical copula constructions for contrast: *accanto all'arancia per me è la pasta per i denti*. Group 2 resembles Italian native speakers for continuity, but shows some backsliding in contrastive contexts, using existentials without a slot for a contrastive pronoun in pre-verbal position: *In alto lì a sinistra accanto all'arancia c'è una pera*.

Overall, our findings confirm that the adoption of target-like preferences for information structure is a difficult task. Italian existentials are more salient candidates for reference introduction (and potentially also taught): albeit structurally different from German, they were more frequently produced than possessives that are available in L1. Learners with long-term immersion outperformed intermediate classroom learners, suggesting that natural interaction supports the acquisition of target-like variants. Both groups showed some backwards development in contexts requiring the expression of a contrast to the interlocutors' turn, suggesting that information structure management in reaction to a dialogue partner requires an extra effort.

Flecken, M. et al. (2013). Principles of information organization in L2 use. *IRAL* 51, 229-242.

**Keywords:** L2 Italian, discourse, information structure, syntax

# Higher level of biliteracy is associated with better executive function in Greek-English bilingual children

Froso (effrosyni) Argyri <sup>\*† 1</sup>

<sup>1</sup> UCL, Institute of Education – United Kingdom

**Background:** Recent research demonstrates that higher engagement in bilingual language environments boosts some features of cognition, including executive function, among bilinguals. Most studies however have explored cognitive effects of bilingualism by examining mainly the role of language use and proficiency without taking into account bilinguals’ literacy experience in both languages. Biliteracy is a crucial aspect of bilingual development that requires acquiring but segregating distinct orthographic, phonological and grammatical rules. Previous research suggests a strong relationship between level of dual language proficiency/use and cognition (e.g., Smith & Briggs Baffoe-Djan, 2019). Extending this premise to the context of biliteracy, the current study examines whether bilingual children with a higher level of biliteracy have better executive function (EF) abilities than bilinguals with lower level of biliteracy. In contrast to previous studies (e.g., Andreou, 2015), we examine biliteracy effects using a range of EF tasks and measure children’s reading abilities directly.

**Method:** Twenty-eight English-Greek bilingual children, exposed to both languages within the first 5 years of life (age  $M \pm SD = 10.1 \pm 1.2$  years 19 females) were tested on Greek and English measures of literacy and online executive function measures. The literacy measures comprised decoding, reading comprehension and reading rate (DADA reading test for Greek and YARC reading test for English). Expressive English and Greek vocabulary (Crichton vocabulary scale) were also assessed. Based on those measures, we constructed three biliteracy indices (mean, cumulative percentage, and difference) for the three literacy measures (decoding, reading rate and reading comprehension) of each language, resulting in nine indices. The executive function measures include the Hearts & Flowers (H&F), Flanker, Corsi span and the CELF-4 Digit Span. Children additionally completed the Raven’s Progressive Matrices (nonverbal reasoning test). Bayesian model comparisons were performed with uninformative priors to test the associations between the biliteracy indices and cognitive performance.

**Results:** Higher biliteracy decoding mean was associated with faster performance in the incongruent condition of the H&F ( $BF=1.83$ ). Better biliteracy reading comprehension mean was associated with higher Corsi forward span ( $BF=1.55$ ). The model with the strongest evidential strength comprised the biliteracy decoding in predicting reaction time (Flowers – Hearts) in the H&F task ( $BF=16.4$ , strong). This model remained the most supported when tested against others which included age, gender, reasoning ability, Greek and English vocabulary.

**Conclusions:** Our findings suggest that, in school-aged bilingual children, better biliteracy decoding ability is associated with decreased slowing related to increased inhibitory and switching task demands, indicating enhanced efficiency. Secondly, better biliteracy reading comprehension mean is associated with higher nonverbal memory. These results support earlier research stud-

---

\*Speaker

†Corresponding author: f.argyri@ucl.ac.uk

ies with monolingual young readers in showing an association between executive function and bilingual reading abilities (e.g., Spencer & Cutting, 2020; Ober, Brooks, Homer & Rindskopf, 2020). Overall, these preliminary findings indicate a mutually beneficial effect of biliteracy and EF development in childhood.

**Keywords:** bilingualism, biliteracy, executive function

# Vocabulary learning from captioned video viewing in primary school learners: The influence of word- and context-related factors.

Daniela Avello \* <sup>1</sup>, Carmen Muñoz <sup>2</sup>

<sup>1</sup> University of O'Higgins – Chile

<sup>2</sup> University of Barcelona – Spain

The existing body of research examining the potential of audio-visual input has demonstrated that foreign language learners' exposure to captioned videos results in significant vocabulary gains (Montero Perez, 2022; Muñoz, 2022; Reynolds et al., 2022). However, these findings have primarily emerged from studies with adolescents and adults, and much less is known about the extent to which these positive outcomes may be translated to primary school learners (e.g. Casulleras, 2023). To this aim, the present investigation assessed the effects of an extensive treatment (11 captioned episodes of an animated cartoon) by means of pretest and immediate posttest administration in a group of 120 primary school learners from Chile (4th-5th grade; 9-11 years old; L1 Spanish-L2 English). More specifically, students' vocabulary gains were assessed at the level of written-word form recall, and receptive written form-meaning mapping. In addition, considering the variability in gains among target items as a function of the learning burden involved in each word (Barclay, 2021), this investigation explored the role of word-related factors (i.e. orthographic regularity, length, concreteness) and context-related factors (frequency of occurrence in the input) in the outcomes. The results revealed that the participants obtained significant vocabulary gains, especially at the level of receptive form-meaning mapping. In addition, fifth graders were found to obtain higher gains in comparison with fourth graders, suggesting that the older group was better equipped (e.g. vocabulary knowledge and cognitive development) to pick up words from the input. As regards word- and context-related factors, the results indicated that the odds of learning the target items increased with orthographic regularity, concreteness, and frequency. As for word-length, the analyses suggested that shorter word forms were easier to recall, whereas longer word forms appeared to be more salient in the input and enhanced viewers' need to figure out their meaning.

## References

Barclay, S. (2021). *Examining the Learning Burden and Decay of Second Language Vocabulary Knowledge* (Unpublished doctoral dissertation). University College London, Institute of Education.

Casulleras, M. (2023) *Language learning from subtitled TV series in the primary EFL classroom*. Unpublished PhD thesis, University of Barcelona.

Montero Perez, M. (2020). Incidental vocabulary learning through viewing video: The role of vocabulary knowledge and working memory. *Studies in Second Language Acquisition* 42(4), 749–773. <https://doi.org/10.1017/S0272263119000706>

---

\*Speaker

Muñoz, C. (2022). Audiovisual input in L2 learning. *Language, Interaction and Acquisition*, 13(1), 125-143. <https://doi.org/10.1075/lia.22001.mun>

Reynolds, B. L., Cui, Y., Kao, C., & Thomas, N. (2022). Acquisition through viewing captioned and subtitled video: A scoping review and meta-analysis. *Systems*, 10 (133). <https://doi.org/10.3390/systems10>

**Keywords:** vocabulary learning, foreign language learning, young learners, audiovisual input.



# Modal overspecification in L2 argumentative writing: A study on L1 Turkish learners of French

Sülün Aykurt Buchwalter \* <sup>1</sup>

<sup>1</sup> LInguistique et Didactique des Langues Étrangères et Maternelles – Université Grenoble Alpes, INALCO – France

By expressing some features redundantly, learners may reduce ambiguity and cognitive load, resulting in overspecification (Lenart & Leclercq, 2021). Existing studies have focused on overspecification in adult learners' oral narratives, in particular in nominal references, tense and aspect, and to a lesser extent, modality (Leclercq & Lenart, 2014). There has been little research on overspecification in modal references and in non-narrative L2 discourse. This study investigates modal overspecification in argumentative writing at intermediate/advanced levels. Modality is a crucial component of argumentation, as it expresses uncertainty (Perelman & Olbrechts-Tyteca, 2008). Is modal overspecification a common strategy in L2 argumentative discourse? To what extent do L2 learners make more explicit references to modality than native writers? To what extent advanced learners differ from intermediate learners when it comes to modal overspecification?

In this study, we are focusing on the following French markers of epistemic modality (Le Querler, 1996):

- modal verb "pouvoir"
- structures with "si" ("if")
- conditional mood

In L2 French, modal verbs are acquired early, while conditional mood is acquired at later stages (Véronique, 2009).

## Methodology

60 learners (undergraduate students with L1 Turkish, enrolled in an intensive French program at a Turkish university) participated in the study. Their mean age is 20 and they have previously studied English as well.

- 30 are at B1 level

---

\*Speaker

- 30 are at B2 level (proficiency level estimated by university)

A control group of 30 students (undergraduate/graduate students in humanities in France) took part in the study. Their L1 is French, and they have studied other languages. Students were asked to write an argumentative text of 250 words in one hour, with two identical prompts (a formal letter to persuade the mayor to not cancel a cultural event; an email to other students asking for contributions to a vegetable garden). Texts have been coded for the three categories of modal markers mentioned above.

## Results & discussion

Learners strongly overuse "pouvoir", with a frequency three times higher than natives. The frequency of "si" structures is about twice higher among learners in comparison with native writers. However, the frequency of the conditional mood is lower among learners, especially at B1 level, who use it three times less frequently than natives. Learners use the verb "pouvoir" within "si" structures, especially at B1 level, even though this combination is rarely found in the control group. Learners do not use the same tense combinations as the control group in "si" structures: they often use future tenses in either or both propositions, while native writers do not use future tenses in "si" structures.

Learners' argumentative writing is characterized by modal overspecification, especially at B1 level. It is possible that learners who do not yet master the conditional mood tend to use "pouvoir" and "si" structures instead. The combination of "pouvoir" and the future tense may be another strategy to compensate for the lesser use of the conditional mood. By overspecifying references to modality, learners may try to reduce ambiguity on the degree of uncertainty and reduce the level of cognitive difficulty.

**Keywords:** overspecification, modality, L2 writing

# The effect of additional CLIL instruction on young learners' oral grammatical complexity

Raúl Azpilicueta-Martínez \* <sup>1</sup>

<sup>1</sup> Universidad Pública de Navarra [Espane] = Public University of Navarra – Spain

Over the past few decades, Europe has witnessed a widespread surge in the use of English as a medium of instruction (Graddol, 2006; Heras & Lasagabaster, 2015). Despite this trend, the extent to which Content and Language Integrated Learning (CLIL) enhances proficiency in young learners (YLS), compared to traditional English as a Foreign Language (EFL) programs, remains uncertain. Muñoz (2015) suggested a 300-CLIL-hour yardstick as the exposure threshold beyond which sizeable language gains would be expected, although her claim was underpinned by findings from studies carried out with older populations, complicating its generalisability with YLS (Lázaro-Ibarrola, 2012; Villarreal & García Mayo, 2009). Adding complexity to the matter, potential mediating factors such as socioeconomic status and extramural exposure have been rarely accounted for in the existing literature in the field (Muñoz, 2015). In addition, there is evidence that CLIL influences different skills and subskills in different ways (Admiraal et al. 2006; Dalton-Puffer, 2008; Ruiz de Zarobe, 2011, 2015), among which oral production has received comparatively little attention (Pérez Cañado & Lancaster, 2017). Specifically, there is a paucity of knowledge regarding the impact that varying degrees of CLIL exposure may have on YLS' grammatical complexity.

This study analysed the oral grammatical complexity in four groups of L1-Spanish YLS ( $n=108$ ) performing an oral task. The sample comprised i) an EFL-only group ( $n=23$ ; mean age= 10.91), who had received 1766 EFL hours including grammatical instruction; ii) a low-CLIL group ( $n=21$ ; mean age= 10.24) who had received the same amount of EFL instruction (1766 hours) plus 707 CLIL hours; iii) a high-CLIL group ( $n=32$ ; mean age= 10.91; EFL: 1766 hours, plus 2473 CLIL hours) and iv) a younger high-CLIL group ( $n=32$ ; mean age= 9.84; EFL: 1545 hours, plus 2164 CLIL hours). All four cohorts were matched for socioeconomic status and extramural exposure in order to isolate the CLIL variable. Data analysis was twofold and comprised a) computational measures of overall sentence complexity (complex T-unit ratio), amount of subordination (dependent clause per T-unit ratio) and amount of coordination (phrase per T-unit ratio), and b) the assessment of two independent raters following a holistic scale previously used in similar studies (Lasagabaster & Merino, 2018). Kruskal-Wallis and Dunn-Bonferroni post-hoc tests were conducted.

The results from the computational analyses revealed significantly higher values for the older high-CLIL group over the EFL-only group at the overall sentence complexity, and amount of subordination levels. No other differences were spotted between any of the groups. The results of the independent ratings showed significantly higher values for the high-CLIL cluster (both age groups) over the EFL-only group. These findings underscore the potential of high-exposure CLIL to produce significant oral grammatical complexity gains at the same time they question the suitability of low-CLIL programs to provide an advantage over EFL-only instruction in said subskill. The results also suggest that the comparatively higher exposure of the younger high-CLIL group might have overridden the one-year cognitive advantage of the older, EFL-only learners.

---

\*Speaker

**Keywords:** computational measures, exposure, grammatical complexity, young learners

# L'acquisition du syntagme nominal en grec L3 : l'influence de la L2 et de l'activité métalinguistique

Eftychia Belia \* 1,2,3

<sup>1</sup> Paris 8 - Vincennes Saint-Denis – Université Paris VIII Vincennes-Saint Denis : EDCLI 224 – France

<sup>2</sup> CNRS Structures Formelles du langage UMR 7023 – Centre national de la recherche scientifique -  
CNRS (France) – France

<sup>3</sup> ED 224 CLI – ED 224 CLI – France

Nous étudions l'acquisition du syntagme nominal en grec L3 par des francophones ayant appris tous préalablement l'anglais, l'allemand ou l'espagnol. L'allemand tout comme le grec possède une morphologie nominale casuelle importante tandis que l'espagnol se différencie du grec par l'absence de marquage casuel. Le projet VILLA (J. Hinz et alii 2013, M. Watorek 2020) a étudié des apprenants de L1 différentes et a montré que les débutants en contexte d'enseignement guidé développent très tôt une sensibilité au marquage casuel. Notre étude interroge l'influence de la L2 des enquêtés dans l'acquisition du marquage casuel dans une L3 en contexte guidé. Nous partons de l'idée (R. Rast 2008 ; P. Trévisiol 2006) que les langues apprises avant une nouvelle acquisition peuvent avoir un impact sur celle-ci tout comme la L1. Face à la différence/similitude entre le grec d'une part, et l'allemand et l'espagnol d'autre part, notre question est de savoir si le système casuel appris en allemand pourrait faciliter l'apprentissage de la morphologie nominale casuelle en grec.

Nous avons recueilli nos données auprès d'apprenants en troisième année de licence FLE en France où nous avons enseigné un cours d'initiation en grec, dans un dispositif universitaire d'enseignement à distance. Chaque unité commence par un texte qui illustre une situation de communication, nous fournissons du lexique et des points grammaticaux du grec expliqués en français (S. E. Carroll 1999). Le cours donne des explications théoriques et illustre par des exemples concrets les substantifs masculins singuliers en *-s*, qui se terminent au nominatif en *-o s*, *-a s* et à l'accusatif en *-o*, *-a* (G. Babiniotis, C. Clairis 1999). Pour réaliser les tâches de l'enquête, les apprenants peuvent consulter le cours et un dictionnaire. La première tâche demande aux apprenants de faire la description écrite d'une photo qui illustre deux entités, un homme sur un âne, qui avancent sur un chemin dans une île des Cyclades. Les substantifs " un homme " et " un âne " se terminent au masculin singulier en *-s*, sont explicités et utilisés fréquemment dans le cours (N. E. Ellis 2002). La deuxième tâche demande aux étudiants d'écrire les difficultés qu'ils ont rencontrés lors de la production de la description en grec et comment ils ont fait pour les résoudre. Nous avons recueilli quinze descriptions et quinze verbalisations métalinguistiques pour le groupe L2 allemand et quinze pour le groupe L2 espagnol. Nos résultats, au terme des six premières unités au premier semestre, montrent que le groupe L2 allemand distingue le nominatif de l'accusatif dans la description mieux que le groupe L2 espagnol. L'analyse des difficultés rencontrées, et des solutions apportées, dans la tâche écrite montre deux modes différents de travail. Le groupe L2 allemand développe son activité d'apprentissage alors que le groupe L2 espagnol avance moins rapidement, développe des contraintes à l'apprentissage. Question vers l'audience : Y a-t-il un rapport entre l'acquisition des marques casuelles du nom en grec L3 et l'activité métalinguistique de l'apprenant, sollicitant sans doute L1 et L2 ?

---

\*Speaker

**Keywords:** Acquisition L3, input, apprenants débutant ou stade initial, tâche de description, activité métalinguistique, morphologie nominale casuelle

# Improving pragmatic skills in a foreign language through embodied prosodic training

Florence Baills <sup>\*† 1</sup>, Núria Esteve Gibert <sup>2</sup>, Lorraine Baqué <sup>3</sup>

<sup>1</sup> Universitat de Lleida – Spain

<sup>2</sup> Universitat Oberta de Catalunya [Barcelona] – Spain

<sup>3</sup> Universitat Autònoma de Barcelona – Spain

Prosody plays a crucial role in shaping meaning and effective communication in spoken language (e.g., Romero-Trillo & Sadeghi, 2023). Non-verbal cues such as facial expressions and gestures equally enrich and modulate the literal content of verbal messages (e.g., Wharton, 2009). Gestures and prosody timely align to fulfil a variety of functions, among them signalling discourse structure (e.g., Krivokapić, 2014), expressing pragmatic intentions (e.g., González-Fuente, 2015), and marking focused elements in the discourse (e.g., Esteve-Gibert et al., 2017). It is particularly challenging for language learners to acquire these multimodal pragmatic skills (e.g., Safavi & Kermad, 2022). Hence, it seems essential to address these multimodal cues in both teaching and research practice aiming at developing learner’s pragmatic competence in the target language (e.g., Kang & Kermad, 2019).

In this study, we propose a classroom intervention to test a range of multimodal activities aiming at teaching French pronunciation with a focus on pragmatic meaning. Fifty Catalan learners of French will engage in five 30-minutes training sessions in one of two conditions: in the *embodied prosody* condition, the instructor will use techniques involving head and manual gesture to teach the prosody of simple and complex speech acts, while marking discourse and information structure; in the *control* condition, pronunciation will be taught following the pronunciation exercises proposed in learners’ regular textbook. Participants will be tested before and after the intervention in oral presentation and debating tasks requiring pragmatic abilities and centred around topics relevant to their curriculum. Evaluation will encompass global assessment of comprehensibility and fluency, along with an analysis of their pragmatic skills in terms of speech acts’ prosodic accuracy and gestural marking of information and discourse structure, which will be annotated in Praat and Elan respectively. Generalized linear mixed models will be applied to each of these variables to assess significant differences between the two groups.

Esteve-Gibert, N., Borràs-Comes, J., Asor, E., Swerts, M., & Prieto, P. (2017). The timing of head movements: The role of prosodic heads and edges. *The Journal of the Acoustical Society of America*, 141, 4727–4739.

González-Fuente, S., Escandell-Vidal, V., & Prieto, P. (2015). Gestural codas pave the way to the understanding of verbal irony. *Journal of Pragmatics*, 90, 26-47.

Kang, O., & Kermad, A. (2019). Prosody in L2 pragmatics research. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 78-92). Routledge.

---

\*Speaker

†Corresponding author: florencebaills@gmail.com

Krivokapić, J. (2014). Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning processes. *Philosophical Transactions of the Royal Society B*, 369, Article 20130397.

Romero-Trillo, J., & Sadeghi, Y. (2023). Prosodic pragmatics in context. In J. Romero-Trillo (Ed.), *The Cambridge Handbook of Language in Context* (pp. 314-328). Cambridge University Press.

Safavi, A. & Kermad, A. (2022). Prosodic patterns of refusals: differences between English native, non-native, and bilingual speakers. In J. Levis & A. Guskaroska (eds.), *Proceedings of the 12th Pronunciation in Second Language Learning and Teaching Conference*, held June 2021 virtually at Brock University, St. Catharines, ON.

Wharton, T. (2009). *Pragmatics and non-verbal communication*. Cambridge University Press.

**Keywords:** prosody, gesture, pragmatics, oral skills, classroom intervention



# Development of syntactic and phonological processing during a residence abroad: eye-tracking evidence from L1 Mandarin speakers in the UK.

Lewis Baker <sup>\*†</sup> <sup>1</sup>, James Turner <sup>\*</sup>

2

<sup>1</sup> Technische Universität Braunschweig – Germany

<sup>2</sup> University of Southampton – United Kingdom

Research into Residence Abroads (RAs) has typically focused on students with European L1s (see Borràs and Llanes, 2021 for a state-of-the-art). Furthermore, many employ only off-line measures of linguistic development. The current study addresses this with a longitudinal investigation into the development of L1 Mandarin speakers’ online processing of both syntax and phonology during a RA at a UK university. This study is – to the best of our knowledge – the first to use eye-tracking data to investigate the development of multiple linguistic levels in a study abroad context.

At the beginning of their RA (timepoint 1), 30 L1 Mandarin speakers (and 30 L1 English controls) completed two Visual World eye-tracking experiments testing syntactic and phonological processing in English. The syntax experiment tested predictive processing of number with the determiners *this/these/that/those*, each followed by a balanced number of singular and plural nouns, as well as *the+noun* to test sensitivity to nominal number marking independent of determiner number marking (English and Mandarin differ regarding number morphology). The phonology experiment tested processing of 4 minimal pairs contrasting /i/ vs. // (e.g., *sheep* vs. *ship*), a tense-lax contrast which is known to be difficult for L1 Mandarin learners of English (Yang et al., 2016), as well as 4 items contrasting word-final /s/ vs. /z/ (e.g., *rice* vs. *rise*). Given that the /i/--// contrast helps to distinguish the English determiners *these* and *this* (along with /s/--/z/), we hypothesised that participants who struggle with the processing of this vowel contrast would also struggle to use these two determiners predictively in online processing.

Results of the syntax experiment reveal that for the congruent conditions (e.g., *that + singular noun*, *these + plural noun*), L1 Mandarin speakers do correctly predict the upcoming noun. The exception appears to be for the determiner *this* (Figure 1). Mixed effects binomial logistic regression modelling confirms that the difference between the proportion of looks to the target and competitor is significantly lower for the *this* condition compared to the other determiners. Further analyses suggest this may be linked to a phonological processing difficulty: in the phonology experiment, the difference between the proportion of looks to target and competitor increases significantly more quickly for /i/ over time than for // (i.e., the L1 Mandarin speakers find the processing of the vowel in *this* more challenging than the vowel in *these*). This is likely due to the phonetics of the Mandarin /i/ vowel mapping on to English /i/ more

---

\*Speaker

†Corresponding author: lewis.baker@tu-braunschweig.de

straightforwardly than // . Furthermore, results of the syntax experiment suggest that even once participants have heard the final consonant cluster of the nominal in the *this + singular noun* condition i.e., the absence of plural marking, the L1 Mandarin speakers never fully discount the plural noun (Figure 2). This is again confirmed by statistical modelling and we explore possible explanations for this. Results from timepoint 2 (end of the RA) will be available to compare with timepoint 1 by the time of the conference.

**Keywords:** L2 processing, syntax, phonology, eye, tracking, residence abroad

# Fostering intercultural communicative competence through interaction with a task-based conversational agent aligned with the school curriculum

Elizabeth Bear \* <sup>1</sup>

<sup>1</sup> University of Tübingen – Germany

Despite the recognized importance of intercultural communicative competence (ICC; Byram, 1997) in many second/foreign language (L2) programs, the teaching of culture, and its integration into the curriculum, remains a challenge (Wang, 2023). One pedagogical approach that has yielded positive outcomes in ICC development is interaction with a target language dominant speaker (TLDS) through computer-mediated communication (CMC; e.g., Tecedor & Vasseur, 2020; Hetrovicz, 2022). Notwithstanding the affordances, interaction with a TLDS through CMC may entail extensive coordination with another class (Belz, 2002; O’Dowd, 2021), or for paid services such as TalkAbroad, monetary expenses (Hetrovicz, 2022). Drawing upon recent advances in speech and language technologies, we explore whether, in place of a human TLDS, learners can increase their ICC through interaction with a task-based conversational agent. While prior work has found that conversational agents can foster ICC (Johnson & Valente, 2009), few studies have targeted the school context and integration with the curriculum.

The tasks in our conversational agent are based on task-based language teaching (TBLT; Ellis 2009) and aim to align with the linguistic and ICC goals of the L2 English secondary school curriculum of a federal state in Germany. Unlike other applications which present language in isolation and are not perceived to teach the target culture (Kessler, Loewen, & Gönülal, 2023), we investigate whether learners can develop ICC through interaction with a conversational agent in contextualized, meaningful tasks, e.g., paying at a restaurant (RQ1). Given findings that cultural elements in an interaction may be impacted by task characteristics (Hetrovicz, 2022) and awareness raising (Liddicoat & Crozet, 2001), we further examine whether tasks designed to draw attention to cultural elements, e.g., cultural differences in paying at a restaurant, foster larger ICC gains (RQ2).

Our study will take place as a randomized controlled field trial in 7th grade English classes in Germany. We plan to measure gains in ICC through a questionnaire assessing five dimensions of ICC, adapted from Chao (2014), and through reflective writing, which will be evaluated via a coding scheme. To assess RQ1, we will compare ICC gains in the conversational agent condition to a waitlist control group. To address RQ2, within the conversational agent condition, half of the classes will receive tasks with awareness raising of cultural elements, and half of the classes will receive tasks without such awareness raising. While we will also evaluate language learning gains and learner confidence, which has been shown to increase through interaction with a TLDS (Hetrovicz, 2022; Kessler, Loewen, & Trego, 2023), we focus this talk on ICC.

As we will conduct our piloting in the spring and the main study in the fall of 2024, in this

---

\*Speaker

doctoral workshop, we hope to benefit from a discussion on our assessment procedures, e.g., our questionnaire and coding scheme, and our theoretical framework. The results would shed light on whether curricular ICC targets can be fostered through a fully automated TLDS and on effective task design features for the integration of cultural elements within instructed second language acquisition.

**Keywords:** intercultural competence, computer assisted language learning, task based language teaching

# Insights from multimodal analysis on longitudinal interview data: the case of I don't know.

Pauline Beaupoil-Hourdel <sup>\*† 1,2</sup>, Pascale Leclercq <sup>\*</sup>

<sup>3</sup>, Elisa Sneed German <sup>\*</sup>

<sup>4</sup>, Amanda Edmonds <sup>\*</sup>

5

<sup>1</sup> Centre de Linguistique en Sorbonne – Sorbonne Université – France

<sup>2</sup> Sorbonne Université, INSPE de Paris, CeLiSo – Université Paris IV - Paris Sorbonne – France

<sup>3</sup> Université Paul Valéry Montpellier 3 – Université Paul Valéry Montpellier 3, Université Paul-Valéry - Montpellier 3 – France

<sup>4</sup> Université Paul-Valéry Montpellier 3 – Université Paul-Valéry - Montpellier 3 – France

<sup>5</sup> Université Côte d'Azur – Université Côte d'Azur, Université Côte d'Azur, Université Côte d'Azur, Université Côte d'Azur, Université Côte d'Azur – France

Discourse markers (DM) are defined as expressions with little referential meaning but multiple pragmatic functions (Pichler, 2009, p. 561). One such DM is the expression *I don't know* (*IDK*), which is widely used by native speakers and learners alike. While recent research has focused on the different discursive functions that *IDK* and its French equivalent '*je sais pas*' can take, another line of studies takes a multimodal approach to this linguistic phenomenon, investigating gaze and gestures that may accompany DMs. In the current study, we investigate the association between *IDK*, gaze, and shrugs (frequently associated with the use of *IDK*). Previous research on shrugs indicates that there are associations between the uses of specific formational components of the shrug (the hands, the arms, the shoulders, or the head) and specific functions across languages (Boutet & Morgenstern, 2019) that also differ depending on the language the speakers use (Morris, 1994). Crosslinguistic effects have indeed been observed in the acquisition of gestural patterns (Gullberg, 2022).

As such, do French learners of English appropriate the gestural patterns of their target language over a year abroad? We hypothesize that gestural patterns evolve along with linguistic development and investigate whether there is a transfer effect of L1 gestural patterns in L2 interactions.

We used a case-study approach to investigate the evolution of the use of multimodal *IDK* by one L2 learner of English over a year abroad, a learning context demonstrated to increase DM use. Five filmed semi-guided interviews were conducted over the course of 12 months. The resulting 138mn of recordings were transcribed, and 92 instances of *IDK* were identified. We coded them in ELAN and in a grid for pragmatic function, drawing on both linguistic and multimodal cues,

---

\*Speaker

†Corresponding author: pauline.beaupoil-hourdel@sorbonne-universite.fr

taking into consideration the speaker's body posture, gestures (shrugs and other gestures), gaze and the prosodic contours of their vocal productions.

We used exploratory Multi-Correspondence Analysis (MCA) to identify clusters which illuminate the way DM and lexical *IDK* are used by the L2 participant: Results show a tendency to phonological reduction, combined with creaky voice, when *IDK* is used as a discourse marker. Regarding crosslinguistic and developmental effects, our qualitative analysis of shrug production suggests that the participant is influenced by L1 gestural patterns while developing more fluid co-verbal gestures over time abroad.

Boutet, D., & Morgenstern, A. (2019). *A cross-linguistic approach to the formal and functional features of recurrent gestures: Shrugging in 5 languages*. ICLC 15 Crosslinguistic Perspectives on Cognitive Linguistics, Nishinomiya, Japan.

Gullberg, M. (2022). Why Second Language Acquisition of sign languages matters to general SLA research. *Language, Interaction and Acquisition*, 13(2), 231-253. <https://doi.org/10.1075/lia.22022.gul>

Morris, D. (1994). *Bodytalk: A World Guide to Gestures*. Random House.

Pichler, H. (2009). The functional and social reality of discourse variants in a northern English dialect: I DON'T KNOW and I DON'T THINK compared. *Intercultural Pragmatics*, 6(4), 561-596.

**Keywords:** Discours Marker, I don't know, gesture, shrugs

# Unraveling Grammatical Gender: Insights from Structural Priming in Bilingual and Monolingual Speakers

Giuseppe Dario Benigno \*<sup>1</sup>, Pedro Guijarro-Fuentes<sup>1</sup>, Estela García Alcaraz<sup>1</sup>, Marta Rivera Zurita<sup>1</sup>

<sup>1</sup> Universidad de las Islas Baleares – Spain

Questions persist regarding the psycholinguistic aspects of grammatical gender representation, especially in the processing of it. Previous evidence suggesting shared syntactic representations in bilinguals prompts the hypothesis of an integrated lexicon, where representations of one language interact with the other (Hartsuiker et al., 2015; Paolieri et al., 2010; 2020).

The present study aims to uncover how co-activation of gender information impacts the processing of grammatical gender. Especially relevant for this study is the analysis of gender agreement violations in ninety monolingual Spanish speakers and in ninety Spanish-Catalan bilinguals with different language profiles (accounted as the Age of Onset of Acquisition and the linguistic dominance). Unlike prior work, this research, using a structural priming paradigm, targets grammatical gender (dis)agreement both within-language (between the determiner and the noun) and between languages. Employing a structural priming paradigm in this context contributes novel insights into psycholinguistics, offering a cognitive perspective on gender processing mechanisms.

The experiment involves presenting prime sentences including a NP that will always be grammatical (i.e., the determiner and the noun will always show gender agreement) and gender-congruent between languages (Spanish and Catalan nouns will have the same grammatical gender) (e.g., "LA-fem IGLESIA-fem (the church)" - "EL-mas – TELÉFONO-mas (the phone)"). With respect to the target sentences exclusively ungrammatical NPs have been included (i.e., mismatch between determiner and noun), so that a masculine noun is always preceded by a feminine determiner "la", and a feminine noun is always preceded by a masculine determiner "el". The NP in the target can be either congruent between languages (MESA-esp-fem – TAULA-cat-fem (table) or ESCRITORIO-esp-mas – ESCRITORIO-cat-mas (desk)) or incongruent (CABEZA-esp-fem – CAP-cat-mas (head) or TENEDOR-esp-mas – FORQUILLA-cat-fem (fork)). Spanish monolinguals will perform the task only in Spanish, while Spanish-Catalan bilinguals will engage in tasks across both languages, examining cross-linguistic processing.

The expected results are the following: differential effects among monolinguals, Spanish-dominant, and Catalan-dominant bilinguals are expected, influenced by cross-linguistic congruence in terms of grammatical gender representation and language dominance. Specifically, we posit that the co-activation at the lemma node of both languages' gender information (given by grammatical and gender congruent words) would either enhance or reduce the transfer effect during the processing of ungrammatical and gender (in)congruent structures and that these effects will vary as a function of proficiency as suggested by Hartsuiker et al. (2015) and language dominance. We also propose that the co-activation of both languages' gender systems during the lexical retrieval of gender agreement violations in the target language would produce differential effects (either diminished or enhanced interference effects) for Spanish monolingual, Spanish dominant bilingual and Catalan dominant bilingual speakers. We finally argue that these effects would depend

---

\*Speaker

both on the lexical facilitation/interference generated by the gender (dis)agreement between the determiner and the noun and the cross-linguistic (in)congruence in the grammatical gender of the noun, as suggested by Paolieri et al. (2010; 2020) and Casado et al. (2023).

**Keywords:** Bilingualism, grammatical gender, agreement violations, syntactic processing, structural priming, language dominance



# Unravelling the processing of missing – object syntactic constructions. A comparative study between Spanish monolinguals and Spanish – Catalan bilinguals.

Giuseppe Dario Benigno \* <sup>1</sup>, Pedro Guijarro-Fuentes <sup>1</sup>, Estela García  
Alcaraz <sup>1</sup>, Marta Rivera Zurita <sup>1</sup>

<sup>1</sup> Universidad de las Islas Baleares – Spain

This study investigates whether monolingual and bilingual speakers of object-drop languages represent the missing object abstractly in the syntax, and whether in the case of the bilingual that representation is shared among both syntactic systems. Previous research with monolinguals has evidenced a syntactical-represented account (Cai et al., 2015) suggesting that an indefinite object is syntactically represented even in constructions where the object is not phonologically expressed. Nevertheless, little is known regarding bilingual speakers, specifically regarding the effects of a potential cross-linguistic transfer between languages that differ in the conditions that allow such omission. Spanish-Catalan bilinguals were considered a good fit to explore this phenomenon, due to the differential linguistic constructions of the two languages when it comes to the indefinite object-drop. Specifically, whereas Spanish allows for object omissions when the object can be semantically retrieved by the context and certain semantic features as definiteness are allowed, Catalan employs the partitive clitic "en" whenever the object is not overtly expressed and, thus, does not permit such syntactic constructions (Campos, 1986; Clements, 2006).

This study aims to compare bilingual speakers to Spanish monolinguals to address the following research questions: (i) Do monolingual and bilingual speakers construct a syntactic representation for null objects? If so, how is it mitigated by the effects of semantics? (ii) Is the representation of null objects represented in both syntax systems of the bilingual regardless of the linguistic allowance of Catalan? (iii) To what extent do monolingual and bilingual speakers' syntax resemble one another if the representation of null objects is (not) shared in the two languages? (iv) Is the priming effect modulated by linguistic proficiency/dominance?

Ninety Spanish monolingual speakers and ninety Spanish–Catalan bilingual speakers with different Ages of Onset of Acquisition (AoA) and different language dominance profiles will be recruited for the study.

All participants will take a structural priming task designed to test whether the syntactic structure of a sentence can prime the structure of a subsequent sentence with varying combinations of definiteness and specificity in missing-argument vs. full-form clitic constructions ("los + definite + specific (the)" - "unos + definite - specific (some)"). Additionally, in the Spanish version of the task, an ungrammatical syntactic condition with the "en" partitive has been included, whereas in the Catalan version an ungrammatical condition with the Spanish mirror situation

---

\*Speaker

was included: the omission of the clitic "en". Spanish – Catalan bilinguals will perform the task in both languages.

Monolingual speakers are expected to show a syntactically-represented account as posited by Cai et al. (2015), while no significant differences are expected with changing patterns of specificity. We also predict that depending on the linguistic dominance and the AoA, participants would be subjected to differential priming effects. Specifically, Catalan dominants would experience a higher priming effect for the ungrammatical condition in the Spanish task, whereas a lower effect for the ungrammatical condition in the Catalan task is expected. Probably, Catalan dominant speakers would also show a lower priming effect for object-dropped grammatical conditions in the Spanish task. A mirror-like situation is expected for Spanish dominants.

**Keywords:** Bilingualism, object omission, syntactic processing, structural priming, language dominance

# The inter-relationship between word learning, native phonology and production practice through cross-situational statistics

Sophie Bennett <sup>\*† 1</sup>, João Dinis Fernandes <sup>1,2</sup>, Susana Correia <sup>2</sup>, Padraic Monaghan <sup>1</sup>, Patrick Rebuschat <sup>1,3</sup>

<sup>1</sup> Lancaster University – United Kingdom

<sup>2</sup> Linguistics Research Centre of NOVA University Lisbon (CLUNL) – Portugal

<sup>3</sup> University of Tübingen – Germany

Second language (L2) acquisition requires learners to associate the correct meaning with novel words and acquire the non-native phonology. According to the Revised Speech Learning Model (SLM-r), learners may not produce L2 words in a target-like manner due to interference from their native phonological system (Flege & Bohn, 2021). To develop this L2 knowledge, the skills specificity hypothesis suggests that comprehension and production proficiency is guided by practice of the skill itself (Li & DeKeyser, 2017). This original study tests the inter-relationship between L2 word learning, native phonology and practice modality from a single implicit learning task, which trained participants on novel word-referent mappings from cross-situational statistics. We determined whether production practice improved word production, and whether this was beneficial to acquiring the words' meanings.

Eighty native speakers of British English (aged 18–30 years) with no prior exposure to Portuguese learned to associate eight disyllabic pseudowords with eight novel objects via a cross-situational word learning task (CSWL, e.g., Ge et al., in press). Four pseudowords contained sounds that exist in both the participants' native language (English) and in the target language (Portuguese); the other four contained a nasal vowel that only exists in Portuguese. Participants were randomly assigned to either the listen-only (n=40) or listen-and-repeat (n=40) condition. In every trial of the CSWL task, all participants observed two objects, listened to a pseudoword and selected the object to which it referred, without feedback. Whilst participants in the listen-and-repeat condition repeated the word aloud after selection, those in the listen-only condition remained silent. After this task, participants completed a comprehension test (word recognition) and a production test (picture naming).

For the CSWL task, both groups rapidly acquired the novel word-object associations. This was evidenced by their above-chance performance in selecting the correct object, as predicted. However, the listen-and-repeat group's production accuracy declined as participants completed the CSWL task. Contrary to our predictions based on the SLM-r, participants produced non-native sounds more accurately than native sounds in the production test. We also predicted that the listen-only group would outperform the listen-and-repeat group in the comprehension test and vice versa in the production test, following the skills specificity hypothesis. However, this only occurred in the production test where there was a significant effect of production practice on accuracy. Our findings challenge preconceived expectations of the learnability of non-native sounds across different modalities, whilst opening the door to further production-based CSWL research.

---

\*Speaker

†Corresponding author: s.c.bennett@lancaster.ac.uk

## References

Flege, J., & Bohn, O. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3-83). Cambridge: Cambridge University Press

Ge, C., Monaghan, P., & Rebuschat, P. (in press). The role of phonology in non-native word learning: Evidence from cross-situational statistical learning. *Bilingualism: Language and Cognition*.

Li, M., & DeKeyser, R. (2017). Perception practice, production practice, and musical ability in L2 Mandarin tone-word learning. *Studies in Second Language Acquisition*, 39(4), 593–620.

**Keywords:** statistical learning, implicit word learning, speech production, production practice, European Portuguese

# Teachers' use of semiotic resources for meaning-making in the bimodal-multilingual classroom

Maike Beyer \* 1

<sup>1</sup> Universität zu Köln = University of Cologne – Germany

Teaching newly immigrated Deaf and Hard-of-Hearing (DHH) multilingual learners (IDML, Cannon/Marx, 2023) in bimodal-multilingual classroom contexts requires the flexible use of varied semiotic resources, including different languages, gestures, pointing, (digital) media, and further resources (Wolbers et al., 2023). This study descriptively analyzes which semiotic resources teachers draw on for joint meaning-making with IDML. The analysis was guided by the following research questions:

Which semiotic resources do teachers use for meaning-making with IDML, and how are they jointly used during interactive meaning-making in interaction?

Using a linguistic ethnographic approach, data was collected by means of classroom observations (field notes, video recordings) and semi-structured teacher interviews. Data collection occurred between March and May 2023 at a school for DHH students in Germany. This study included five hours of two different 'German as a Second Language' classroom settings taught in various constellations by three teachers in German Sign Language (DGS) or spoken German. Multimodal data, including sign, spoken and written language, and further semiotic resources, were annotated and subsequently transcribed in ELAN. Building on categories developed by Holmström/Schönström in a higher education setting, categories of semiotic resources from video recordings were adapted in the classroom settings. In the following steps, these categories were continuously refined using inductive methods. Annotated data from the observed teaching situations were described and analyzed to identify which semiotic resources were used and how they were combined and connected. This presentation examines teachers' observed language teaching practices in the two classroom settings. The presentation will show the individual and varied use of semiotic resources to gain shared understanding and to connect understood concepts using (accessible) named languages.

## References:

- Holmström, I., & Schönström, K. (2018). Deaf lecturers' translanguaging in a higher education setting: A multimodal multilingual perspective. *Applied Linguistics Review*, 9(1), 89-111. <https://doi.org/10.1515/applirev-2017-0078>
- Wolbers, K., Holcomb, L., & Hamman-Ortiz, L. (2023). Translanguaging framework for Deaf education. *Languages*, 8, 59-83. <https://doi.org/10.3390/languages8010059>

**Keywords:** semiotic resources, sign language, teacher instruction, SLA, linguistic ethnography

---

\*Speaker

# Observing very-short-term speaking fluency development in computer-delivered interviews

Serge Bibauw \* <sup>1</sup>

<sup>1</sup> Université catholique de Louvain – Belgium

**BACKGROUND.** Utterance fluency, as a dimension of L2 speaking, has been demonstrated to be a key component of L2 proficiency and being able to measure it precisely is essential for testing and research. While various utterance fluency metrics have been proposed, compared and validated against various standards in the literature (e.g., Segalowitz, French & Guay, 2017; Tavakoli, Nakatsuhara & Hunter, 2020), their sensitivity to short-term learning gains has been more rarely studied and, when it has, it was usually during interventions lasting for 4 weeks or more (e.g., Campbell, Colin and McCormack, 2016; Segalowitz et al, 2017). The present study presents a methodology for measuring and analysing speech fluency in computer-delivered interviews that allowed us to observe significant L2 fluency development over less than 3 weeks and only 2 hours of intervention.

**METHODS.** We used a custom-made computer-delivered oral interview consisting of 26 short questions to record individually and simultaneously 215 teenage low-intermediate learners of French in a pre- and a posttest, separated by 1 to 3 weeks. In the experimental group, learners practised the L2 through written interactions with automated agents (chatbots) in a dialogue-based game. The resulting 12'000 audio files were transcribed by automatic speech recognition, manually corrected and annotated for a series of disfluencies. We computed both signal-based (via de Jong & Wempe 2009) and transcription-based fluency metrics, with many variations in terms of pruning (e.g., including or excluding filled pauses, false starts, repetitions, self-talk, L1 words...). We compared learning gains from pre- to posttest and across conditions for most utterance fluency metrics (speech rate, articulation rate, syllable duration, length of runs, duration of runs, duration of silent pauses, silent pausing rate, speech-time ratio, pause duration, etc.). To go beyond a single metric, we computed a composite fluency index based on the first component of a principal component analysis of these metrics.

**RESULTS.** Results show that the most sensitive metric, speech rate, is well able to detect a significant gain in L2 speaking fluency between the pre- and posttest ( $d = 0.51$ ,  $p < .001$ ), i.e. after 1 to 3 weeks, but fails to observe a significant effect between the experimental and the control group, which shows that part of the pre-post difference is due to a task repetition effect ( $d = 0.39$ ). However, the analysis made on the first principal component shows promising sensitivity and sufficient power to distinguish the task-retain effect from the intervention effect.

**DISCUSSION.** Our semi-automated method for data collection and analysis of speaking fluency offers promising precision and sensitivity for measuring very-short-term L2 proficiency gains. A stronger fluency-oriented intervention would certainly produce even more effects in such a timeframe.

---

\*Speaker

**Keywords:** speech fluency, dialogue systems, conversational AI, text chat, transfer

# Assessing the Benefits of Meaning Inferencing: The Case of Idioms

Frank Boers <sup>\*† 1</sup>, Xi Yu <sup>\* ‡ 1</sup>, Xiaofei Wang <sup>\* § 1</sup>

<sup>1</sup> The University of Western Ontario – Canada

While it is often suggested that language learners should be encouraged to try and guess the meaning of new lexical items, concerns have been raised that inaccurate interpretations may linger in memory (e.g., Elgort, 2017). Yu and Boers (2023) examined the effect of an inferencing task on ESL learners' recall of 21 idiomatic expressions (e.g., *jump the gun*) in comparison to directly giving them the meanings, and they found that the inferencing task did not lead to better post-test results when the inferences had needed to be rectified. It was only when the inferences had been accurate that meaning recall was better than in the condition where the meanings were explained directly—a less time-consuming procedure. However, Yu and Boers gauged learning outcomes from a learning episode that was not followed by any additional steps to consolidate the new knowledge. Well-designed language courses do provide such opportunities to consolidate newly acquired knowledge. This presentation therefore reports on an approximate replication study (in progress) of Yu and Boers, using the same counterbalanced design, same-profile participants, and mostly the same 21 target idioms. ESL learners ( $N \geq 30$ ) were presented on an online platform (Qualtrics) with the idioms in brief contexts and they were either prompted to infer the meaning of the idioms, after which they received feedback, or they were given the meaning of the idioms directly. The time spent on the tasks was recorded. As in Yu and Boers, the inferencing condition was designed so that the likelihood of accurate interpretations varied thanks to the provision of a note about the original, literal use of the expression (for example, *jump the gun* comes from a racing event where an athlete starts running before the starting pistol is fired). These notes were included with all the idioms, but they either preceded or followed the inferencing attempts. Different from the original study, this initial learning stage was followed in the same session by a meaning-recall exercise, which was again followed by feedback. One week later, the learners took a test to see how well they remembered the idioms. Logistic regression modeling was used to determine what factors influenced the likelihood of accurate post-test responses. The fixed factors included in the modeling included treatment (inferencing vs. meaning given), inferencing accuracy, time on task, and accuracy in the exercise. Individual learners' inferences and their exercise and post-test responses were compared as well, to examine whether inaccurate interpretations proffered at the learning stage re-emerged later. The findings of the study help to re-assess the effects of inaccurate inferencing on L2 vocabulary learning and retention.

References

Elgort, I. (2017). Incorrect inferences and contextual word learning in English as a second language. *Journal of the European Second Language Association*, 1(1), 1–11. <https://doi.org/10.22599/jesla.3>

Yu, X., & Boers, F. (2023). Inferring the meaning of idioms: Does accuracy matter for retention

---

\*Speaker

†Corresponding author: [fboers@uwo.ca](mailto:fboers@uwo.ca)

‡Corresponding author: [xyu369@uwo.ca](mailto:xyu369@uwo.ca)

§Corresponding author: [xwan2853@uwo.ca](mailto:xwan2853@uwo.ca)



in memory? *RELC Journal*. Epub ahead of print. <https://doi.org/10.1177/00336882231181771>

**Keywords:** guessing, vocabulary, idioms, memory, interference

# Factors influencing L2 learners' single-word productive vocabulary in German: a longitudinal study

Griet Boone \* <sup>1</sup>, Vanessa De Wilde <sup>2</sup>

<sup>1</sup> Universiteit Antwerpen = University of Antwerpen [Antwerpen] – Belgium

<sup>2</sup> Universiteit Gent = Ghent University (UGENT) – St. Pietersnieuwstraat 33, 9000 Gent., Belgium

In vocabulary research, factors affecting second/foreign language (L2) receptive vocabulary have been investigated extensively (e.g., De Wilde et al., 2020, 2022; Peters, 2020; Puimège & Peters, 2019; Reynolds & Ding, 2022). However, few studies have examined factors affecting productive knowledge (i.e., form recall), especially concerning Languages Other Than English (LOTEs). Moreover, not much is known about the influence of time on L2 learners' productive single word knowledge. Nevertheless, it is crucial to understand which variables affect productive vocabulary learning in order to optimize the learning process. This 3-year longitudinal study aims to investigate to what extent time, frequency, cognateness, concreteness, and age of acquisition (AoA) play a role in the development of productive single word knowledge in L2 German. Our participants were 64 Dutch-speaking university learners of German, who participated three times in a German productive vocabulary test, once a year. This test, modelled after Nation's Productive Vocabulary Levels Test for English (Nation, 1990) and developed by ITT Leipzig, consists of 90 items, which measure students' productive vocabulary on the levels 1000, 2000, 3000, 4000, and 5000. Each year, a different version of the test was administered, which means that our dataset contained 270 different words in total. The results of a GLMM showed that time, frequency, cognateness, and AoA together explained 28% of the variance. L2 frequency was found to be the main predictor for L2 form recall. Cognateness also played a role, although the impact was lower compared to the findings of studies on receptive vocabulary knowledge. Concreteness did not affect L2 learners' form recall scores. In our presentation, we will discuss our findings and talk about how these insights can be applied in practice.

**Keywords:** productive vocabulary, form recall, L2 German, frequency, cognateness, concreteness, age of acquisition

---

\*Speaker

# Individual-Level Data and the Property-by-Property/Wholesale Transfer Debate: The Case of L1-English/L2-Spanish and French Learners of L3 Mandarin

Hunter Brakovec \*† <sup>1</sup>

<sup>1</sup> Indiana University – United States

The recent L3 transfer literature has focused on the property-by-property vs. wholesale transfer debate (e.g., Rothman et al., 2019). The Linguistic Proximity Model (LPM) states that crosslinguistic influence occurs from both previously acquired grammars (PAGs) simultaneously for each property being acquired (Westergaard, 2021). The Typological Primacy Model (TPM) states that an individual will wholly transfer one PAG on the basis of structural similarity, as determined by a cue hierarchy (Rothman et al., 2019). The Abbreviated Grappling Period Model (AGPM) states that an individual will wholly transfer one PAG on the basis of lexical or phonological similarity; however, if neither are sufficiently similar, the learner will transfer their dominant PAG (Sprouse & Schwartz, 2023).

To address the three models above, L1-English/L2-Spanish and L1-English/L2-French learners of L3 Mandarin completed an acceptability judgement task with three target contrasts: (1) pre-/post-verbal negation, (2) pre-/post-verbal adverbs (between subject and verb/between verb and direct object), and (3) preposition stranding/pied-piping. Data collection is ongoing and will be completed by the conference date. Thus far, all participants are dominant in English, and all have beginner proficiency in the L3. Participants were also tested on their knowledge of the relevant structures in their respective L1s and L2s.

The TPM, on the basis of morphological similarity, and the AGPM, on the basis of dominance, predict transfer of the English PAG by both groups, indicated by acceptance of pre-verbal negation, pre-verbal adverbs, and both preposition stranding and pied-piping, but rejection of the other conditions. The LPM predicts differing response patterns by group. The L2 Spanish group is predicted to accept pre-verbal negation, pre-verbal adverbs, and preposition pied-piping; reject post-verbal negation; and display intermediate responses for post-verbal adverbs and preposition stranding. The L2 French group is predicted to accept preposition pied-piping; reject nothing; and display intermediate responses for all other conditions.

Preliminary data indicate that no individuals in either group display the response patterns predicted by the LPM. Instead, participants in both groups display a response pattern consistent with transfer of the English PAG. Additionally, while some individuals do show intermediary performance on some conditions, the intermediary performance is only found in those conditions where transfer from English is non-facilitative, indicating subsequent restructuring toward the target Mandarin representation.

The individual-level analysis adopted here helps distinguish between intermediary acceptance

---

\*Speaker

†Corresponding author: hzbrakov@iu.edu

resulting from transfer and intermediary acceptance due to restructuring. The data thus far suggest support for the TPM and AGPM, i.e., wholesale transfer, over the LPM, i.e., property-by-property transfer.

## References

Rothman, J., González Alonso, J., & Puig-Mayenco, E. (2019). *Third language acquisition and linguistic transfer*. Cambridge University Press.

Sprouse, R. A., & Schwartz, B. D. (2023). L3 cognitive states and the Abbreviated Grappling Period Model. In V. Sheu, A. Zhou, & J. D. Weirick (Eds.), *Variation in linguistics: Second language acquisition, discourse studies, sociolinguistics, syntax* (pp. 2–25). Cambridge Scholars Publishing.

Westergaard, M. (2021). Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research*, 37(3), 379–407.

**Keywords:** L3, Transfer, CLI

# Between order and disorder: an information-theoretic approach to linguistic complexity

Paolo Brasolin <sup>\*†</sup> <sup>1</sup>, Arianna Bienati <sup>\*</sup>

1,2

<sup>1</sup> European Academy Bozen/Bolzano – Italy

<sup>2</sup> Università degli Studi di Modena e Reggio Emilia = University of Modena and Reggio Emilia – Italy

The *surprisal* associated with an event of probability  $p$  is defined as  $\log(1/p)$ . This information-theoretic measure has been proposed and used in both psycholinguistics and corpus linguistics (Gries and Ellis, 2015) to capture the intuitive notion of unexpected events (e.g., words, n-grams, constructions) in language. On the processing side, the tension between the expected and unexpected seems to have important effects on both the processing costs and the activation of mechanisms of explicit learning (Jaeger and Snider, 2008). Regarding production, surprisal-based measures could shed light on the complexity of a user’s linguistic system. More proficient language users may not only apply the *normal, conventional usage* of words and constructions, but also exploit them, creating novel, original structures.

In linguistic complexity studies, there have been calls to capture the balance between the expected and the unexpected, conventionality and surprise, order and disorder, structure and lack thereof (Pallotti, 2015). This necessity can be satisfied by the information-theoretical notion of Gell-Mann (effective) complexity, defined as the amount of non-random information in a system: it peaks in the region between order and disorder, where complex structures can manifest. It is often opposed to Kolmogorov (algorithmic) complexity, the total amount of information in a system, which monotonically increases with disorder instead.

In this contribution we focus on lexical complexity to answer the following research question: *which information-theoretical notion of complexity (Kolmogorov vs. Gell-Mann) is measured by the most-used lexical complexity indices?*

To address this question, we conduct a simulation that evaluates lexical complexity on four Italian corpora, which represent the spectrum from expert- to learner-authored texts. We synthesize new texts from the corpora by altering them in two directions: more orderly texts are obtained by repeating increasingly smaller subsections of the text itself until a single word is repeated; whereas a completely disordered configuration is obtained by resampling increasingly smaller n-grams taken from the original text. The synthetic data represents, for each corpus, a way to explore the spectrum from total order to total disorder.

On all texts we compute type-token-ratio-based and surprisal-based metrics, including fluctuation complexity (Bates and Shepard, 1993). A preliminary visual inspection through boxplots shows that TTR-based metrics are sensitive only when there is an increase in order but not

---

\*Speaker

†Corresponding author: paolo.brasolin@eurac.edu

in disorder. Surprisal-based measures, on the other hand, do show interesting Kolmogorov (entropy) or Gell-Mann behavior (normalized entropy and fluctuation complexity). Moreover, when used in combination, they mutually enhance their interpretability. Our results indicate that surprisal-based measures could be a useful addition to the toolkit of the linguistic complexity researcher, since they are adherent to both the theoretical and observational underpinnings of usage-based accounts of language learning.

References:

Bates, J.E., & Shepard, H.K. (1993). Measuring complexity using information fluctuation. *Physics Letters A*, 172(6), 416-425.

Gries, S.T., & Ellis, N.C. (2015). Statistical Measures for Usage-Based Linguistics. *Language Learning*, 65(S1), 228-255.

Jaeger, T., & Snider, N. (2008). Implicit Learning and Syntactic Persistence: Surprisal and Cumulativity. *CogSci30*. Austin:TX.

Pallotti, G. (2015). A simple view of linguistic complexity. *Second Language Research*, 31(1), 117-134.

**Keywords:** complexity theory, lexical complexity, information theory

# Learner corpora of German as a second language in newly migrated students – methodological limits for the design of L1 reference corpora

Aylin Braunewell \* <sup>1</sup>, Julia Schlauch \* † <sup>1</sup>, Jana Gamper ‡ <sup>1</sup>

<sup>1</sup> Professur für Deutsch als Zweitsprache, Justus-Liebig-University Giessen – Germany

Germany is not only de facto a country of immigration, but is also urgently dependent on immigration due to its demographic change and economic requirements (cf. Fuchs et al. 2019). The integration of school-age children and young people who immigrate with little or no knowledge of German is therefore a key task for the education system. Apart from compulsory foreign language training, the German education system is monolingual: German is the language of education and a prerequisite for graduation, vocational training and thus for general participation in society. In several federal states, schooling for newly migrated students consists of intensive language support measures, which in most cases precede attendance in a regular class.

Although the integration of newly migrated students is not a new topic, there is still a considerable desideratum for research on the acquisition of German as a second language for education (cf. Goltseva / von Dewitz 2023).

This lack of research is mainly due to methodological problems: Firstly, the group of learners to be focused on is highly heterogeneous in terms of various factors such as age, L1, country of origin, school experience, and others. It is therefore not precisely a learner group in the sense of established SLA research, but rather a diverse group of people who are constructed as a group by only two characteristics: their status as recent immigrants and their lacking proficiency in the German language (cf. Massumi / von Dewitz 2015). A second challenge for SLA research with a focus on newly migrated students is the operationalisation of language development and the ultimate goal of a language acquisition for educational purposes, as described by Schleppegrell (2001, among others) as "language of schooling". Construction of L1 reference corpora is a common method to address this problem (cf. Wisniewski 2022).

In our presentation we would like to discuss why this practice is not ideal for research on SLA in newly migrated students. Based on a longitudinal learner corpus of newly migrated students that is currently developed at the University of Giessen (SeiKo, cf. Schlauch 2022, Schlauch et al. in prep.), we first want to reflect on how an ideal reference corpus (for SeiKo and similar corpora) could be designed and what challenges this poses (e.g., when dealing with variation, cf. Shadrova et al. 2021). We examine who might be suitable participants for such corpora and how a longitudinal perspective on L2 acquisition and length of exposure could be taken into account. In a second step we assess the suitability of current German corpora from educational settings as reference corpora for our research. In addition to the structural designs, we also consider the availability and linguistic preparation and annotation of the data. Thereby, we aim to shed light

---

\*Speaker

†Corresponding author: Julia.S.Schlauch@germanistik.uni-giessen.de

‡Corresponding author: Jana.Gamper@germanistik.uni-giessen.de

on the question which methodological approaches might be more promising for modelling SLA in the highly heterogeneous context of newly migrated students.

**Keywords:** L2 German, learner corpus, longitudinal, L1 reference corpora, language of schooling, newly migrated students



# Variation in target language use by foreign language teachers: effects of online vs. offline language education

Marco Bril \*<sup>1</sup>, Mieke Cremers<sup>1</sup>

<sup>1</sup> Utrecht University – Netherlands

Target language use (TLU) by foreign language (FL) teachers has been shown to stimulate language learning as it provides FL learners with comprehensible input, situations for language production and interaction (e.g., (3)). However, the amount of TLU by FL teachers is largely affected by personal and teaching practice-related factors ((2)). To date, previous studies on TLU have only focused on in-class situations and not on *online* FL education. Since online education negatively affects interaction between teacher and students (e.g., (4)), one may hypothesize that TLU is negatively affected by online education.

The present study provides new insights into the effect of online FL education on TLU by FL teachers. We focused on the amount of TLU by teachers of French ( $n = 37$ ;  $M$ years experience = 7.4 years;  $SD$ years experience = 7.2 years; speaking proficiency range = B2 – C1) and English ( $n = 23$ ;  $M$ years experience = 12.0 years;  $SD$ years experience = 9.1 years; speaking proficiency range = C1 – C2) during grammar, writing, reading, speaking and listening instruction. We controlled for self-reported speaking proficiency, pedagogical skills, years of experience, TLU with colleagues and the overall students' proficiency. The amount of TLU was reported by the teachers for pre-COVID-19 in-class education, online homeschooling during COVID-19 and post-COVID-19 in-class education by means of a 10-point Likert-scale (1 = never; 10 = always) per type of instruction. Pedagogical skills were measured with the Perceptions of Knowledge and Skills of Teaching Test ((1)) and TLU with colleagues by means of a questionnaire measuring the time per week spent to TLU with colleagues. Data were analysed using a mixed ANCOVA per type of instruction with Teacher (French vs. English) as between-participant variable, Time of Measurement (pre- vs. during vs. post-COVID-19) as within-participant variable and the control variables as covariables. The results showed a (near to) significant effect of Time of Measurement when controlled for years of experience in TLU during writing ( $F(2,106) = 7.53$ ;  $p = .001$ ;  $\eta^2 = .12$ ), reading ( $F(2,106) = 2.92$ ;  $p = .058$ ;  $\eta^2 = .05$ ) and listening instruction ( $F(2,106) = 3.13$ ;  $p = .048$ ;  $\eta^2 = .06$ ), indicating less TLU in online education as compared to the pre- and post-homeschooling situation. For TLU during speaking instruction a significant effect of Time of Measurement was found when controlled for TLU with colleagues ( $F(2,106) = 5.30$ ;  $p = .006$ ;  $\eta^2 = .09$ ). However, for TLU during grammar instruction no significant effects were observed. Overall, teachers of English used the target language more often than teachers of French, except during listening instruction.

We concluded that online FL education does not foster target language pedagogy with respect to all language competences, which may be explained by reduced interaction in online education (cf. (4)), that the extent to which TLU is reduced in online education is influenced by the teacher's teaching experience (writing, reading and listening instruction) and TLU with colleagues (speaking instruction), and that after a period of online language education teachers use the target language to the same extent as before.

---

\*Speaker

**Keywords:** target language use, foreign language teachers, online education

# Home literacy predicts lexical, not grammatical, L2 skills and mediates effects of social status on young learners' phonological awareness

Ann-Christin Bruhn <sup>\*† 1</sup>, Kristin Kersten <sup>‡ 1</sup>

<sup>1</sup> University of Mannheim – Germany

Research indicates a robust predictive influence of socio-economic status (SES) on children's linguistic and cognitive development (overview in Kersten, 2020). This is largely attributed to parental interactions, support and engagement at home, factors also intricately linked to SES (Hackman et al., 2015). The role of the home literacy environment (HLE) has received special attention in this context (Niklas et al., 2016). Phonological awareness is closely linked to reading and writing skills, and previous research underscores the crucial role of HLE for literacy skills. It is therefore plausible, and indeed has often been found, that HLE affects children's phonological awareness (Kang, 2012).

Numerous studies have also demonstrated the positive influence of HLE on language development (Sénéchal, 2015). While effects on the development of a second language have recently gained more attention (Dong & Chow, 2022), its effects on foreign language learning have been studied less frequently (Lau & Richards, 2021; Zhang et al., 2021). In a dataset similar to the current study, Kersten (2023) found that parental HLE and language support predicted L2 lexical skills, but not L2 grammatical reception in 93 German primary school learners of L2 English. The study also revealed an SES effect on phonological awareness and L2 grammar, but not on L2 lexicon.

To further investigate these relationships, in the current project we looked at a sample of third graders in German primary schools. We investigated (1) whether HLE and SES predict receptive L2 skills and phonological awareness, and (2), if so, whether this effect is mediated by HLE.

Our sample consisted of 104 German learners of L2 English in third grade for the PA tests, and a sub-sample of 57 learners (restricted due to the pandemic) for the L2 tests. All skills were assessed using standardised instruments at school. Receptive L2 English lexicon (BPVS III, Dunn et al., 2009) and receptive L2 grammar (ELIAS Grammar Test II, Kersten et al., 2012) were elicited as group tests in class. For PA, we used the P-ITPA (Esser & Wyschkon, 2010), carried out as individual test at school. For SES, operationalized as highest occupational status, and home literacy environment (HLE), we used a parental questionnaire. Following Niklas et al. (2016), we operationalised it as degree of importance of reading at home, general enjoyment of reading in the family, the child's interest in reading and being read to, as well as discussions of reading contents.

A series of regression analyses with SES and HLE as predictors (1) shows that PA is predicted significantly by both independent variables ( $r^2 = .161$ , HLE  $\beta = .257^*$ , SES  $\beta = .226^*$ ). L2

---

\*Speaker

†Corresponding author: ann-christin.bruhn@uni-mannheim.de

‡Corresponding author: kristin.kersten@uni-mannheim.de

lexicon is only predicted by HLE ( $r^2 = .165$ , HLE  $\beta = .349^*$ , SES  $\beta = .107$ ), and L2 grammar by none of the predictors. An analysis using SPSS PROCESS (2) revealed that HLE significantly and positively mediated the effect of SES on PA. Results thus show that parents can promote their children's phonological awareness skills and help them acquire a second language by providing a supportive home literacy environment.

**Keywords:** phonological awareness, home literacy environment, social status, receptive L2 skills

# The Impact of Exposure to Different English Accents on EFL Children’s Word Perception

Zhuohan Chen \* <sup>1</sup>

<sup>1</sup> Zhuohan Chen – United Kingdom

Second language (L2) speakers are likely to produce an accent (the degree of deviation or adherence to a language variety’s pronunciation). Although the notion of ‘standard’ has long been contested, the intelligibility of L2 accents (from both one’s own and others) is crucial for communication efficacy and perception of identities (Lin et al., 2023).

Previous findings showed 30 years of evidence that **adult** L2 learners improve the production of L2 phonemes when they receive High-Variability Phonetic Training (HVPT), i.e., using natural speech stimuli produced by multiple talkers in a variety of phonetic contexts (Sakai & Moorman, 2018; Zhang et al., 2021; Zhang et al., 2023). The effectiveness of HVPT has been extensively examined in consonants (e.g., Japanese learners’ acquisition of English (r) and (l)) but less so in L2 English vowel acquisition (Saito et al., 2022), despite evidence showing that enhanced variability could diminish learning effects for highly confusable vowels (Wade et al., 2007; Zhang et al., 2021). However, little research on HVPT has focused on children. The present few targeting children also showed inconclusive results of high variable input in terms of its generalisability from perception to production (Hu., 2021; Giannakopoulou et al., 2017; Zhang et al., 2023).

However, children, especially sequential EFL learners, are worth researching. Firstly, in practice, their L2 pronunciation learning can face input diversity: Teachers, parents, textbooks, and the ‘digital wilds’ may possess varied accents; how children recognise the words and map the varied phonological forms to the meaning is intriguing. Secondly, the line of Global Englishes research has long been focusing on adults. Little is known about how children react to different L2 accents, despite some evidence showing accents-based social grouping among L1 children (Jeffries, 2022; Paquette-Smith, 2018). Whether global Englishes awareness can be cultivated from a young age is under-explored.

Thus, not joining the ‘native-likeness and intelligibility’ debate, this research looks at how exposure to varied L2 accents impacts children’s word learning by exploring the following:

- To what extent do children notice the accentedness in the L2 input?
- To what extent do children recognize the accented L2 words?

The study will first conduct a systematic review to explore what previous research on children has found about the impact of exposure to different L2 accents on word learning. As for the empirical study, around 90 L1 Mandarin primary schoolers learning English in Hangzhou, China’s new first-tier city, will be recruited with consent from parents. The potential confounding of dialects and socio-economic status will be controlled by selecting participants from a mainstream primary

---

\*Speaker

school. They will listen to an English story twice in one of three accent conditions (single-talker, multi-talker, and control). Pre- and post- referent selection tests (before and after the 1st and the 2nd exposures to the story) will assess their (re)interpretation of stimuli words of four kinds (heard words, unheard words, non-words, and novel words). Children's phonological decoding, L1 and L2 proficiency, vocabulary knowledge, English exposure and so forth will also be assessed.

**Keywords:** L2 pronunciation, Children, Accents

# Exploring the role of segmental and suprasegmental features in L2 speech: Insights from native and non-native listener judgments on comprehensibility and accentedness

Aliaga-Garcia Cristina <sup>\*† 1,2</sup>, Ingrid Mora-Plaza <sup>\* ‡ 2</sup>, Mireia Ortega Duran <sup>\* § 2</sup>, Natalia Fullana <sup>¶ 2</sup>

<sup>1</sup> Universitat de Barcelona (UB) – University of Barcelona Gran Via de les Corts Catalanes, 585 08007 Barcelona, Spain

<sup>2</sup> Universitat de Barcelona – Spain

Second language (L2) speech research has evidenced a shift in focus towards prioritizing comprehensibility and intelligibility over nativelikeness in L2 pronunciation instruction (Levis, 2005, 2020; Munro & Derwing, 2020). Building on a decade of research examining perceptual dimensions for native listeners (e.g. Isaacs & Trofimovich, 2012; Saito et al., 2017), studies on non-native listeners are still fewer (e.g. Mora, 2022; Saito et al., 2019). Inconsistencies in previous research raise questions about whether global evaluations of L2 speech are perceived differently as a function of listener group (native vs. non-native listeners). Some studies have revealed little or no difference in global ratings between native and non-native listeners (Derwing & Munro, 2013; Crowther et al., 2016; Julkowska & Cebrian, 2015), suggesting comparable comprehensibility and accentedness ratings across listener groups. Conversely, other studies have revealed distinct patterns in the perception of these constructs between native and non-native listeners (Ludwig & Mora, 2017), proposing that non-native listeners may exhibit processing advantages in L2 speech, resulting in higher comprehensibility judgements for accented English compared to native listeners. Moreover, no research to date has examined the role of segmental and suprasegmental features of L2 speech and their relative functional load in both native and non-native listeners' ratings of L2 speech comprehensibility and accentedness.

The present study addresses this gap by exploring (1) the differential contribution of segmental and suprasegmental features to explaining variance in L2 speech comprehensibility and accentedness, and (2) the extent to which their relative functional load varies as a function of listener group (native vs. non-native listeners). Eighty-two Spanish/Catalan EFL learners completed a monologic problem-solving task, from which 45-second speech excerpts were selected for analysis. Segmental accuracy measures included voice onset time (VOT) of stressed voiceless plosives and counts of phonetic, allophonic, and syllable structure errors (Trofimovich & Isaacs, 2012; Suzukida & Saito, 2021). Suprasegmental accuracy was examined both (a) objectively through counts of word stress errors, raw- and rate-normalized rhythm metrics (White & Mattys, 2007), stressed-to-unstressed vowel duration ratios, and pitch range (Trofimovich & Isaacs, 2012) and (b) impressionistically in terms of native English listeners' judgments of learners' intonation,

---

\*Speaker

†Corresponding author: cristinaaliaga@ub.edu

‡Corresponding author: imoraplaza@ub.edu

§Corresponding author: m.ortega@ub.edu

¶Corresponding author: nataliafullana@ub.edu

rhythm and discourse richness (Crowther et al., 2018). In addition, the speech excerpts were rated for comprehensibility and accentedness by 20 native and 20 non-native English listeners. Data analysis is currently underway in an attempt to shed light on the intricate relationship between certain segmental and suprasegmental features and the perceived degree of pronunciation proficiency in L2 speech acquisition by both native and non-native listeners.

**Keywords:** L2 SPEECH PERCEPTION, COMPREHENSIBILITY, ACCENTEDNESS, NATIVE SPEAKERS, NON, NATIVE SPEAKERS



# Vocabulary knowledge and vocabulary use in writing: A cross-sectional comparison of L2 English and French

Eva Caltabellotta \*<sup>1</sup>, Elke Van Steendam<sup>1</sup>, Ann-Sophie Noreillie<sup>1</sup>, Elke Peters<sup>1</sup>

<sup>1</sup> KULeuven – Belgium

Although previous research has indicated that receptive vocabulary knowledge significantly predicts L2 writing performance (Lin et al., 2015; Miralpeix & Muñoz, 2018; Stæhr, 2008), the role of productive vocabulary knowledge in L2 writing is less clear. Laufer (1998) found no correlation between productive vocabulary knowledge and lexical sophistication (the proportion of low-frequency words). Additionally, comparison of learners in two consecutive grades indicated that vocabulary knowledge improved as formal instruction increased, but lexical sophistication did not. Given the limited sample size in Laufer’s study ( $n = 48$ ) and the lack of large-scale cross-sectional analyses of productive vocabulary knowledge and vocabulary use, more research on this topic is needed. Further, few studies have compared productive L2 vocabulary knowledge and use across learning settings. The Dutch-speaking region of Belgium presents an interesting context for a cross-linguistic comparison of English and French, as these two foreign languages differ considerably in length of instruction, out-of-school exposure and attitudes, important predictors of L2 vocabulary knowledge and writing performance (Peters et al., 2019; Tahmouresi & Papi, 2021).

The present study aims to extend Laufer’s (1998) study on productive vocabulary knowledge and lexical sophistication in writing. First, we investigated two foreign languages and increased the sample size substantially to boost the generalizability of findings. Second, we administered two writing tasks in each language to reduce topic effects (Yoon, 2017), thereby improving the methodology. Third, in addition to lexical sophistication (use of low-frequency words), we also included a measure of lexical diversity (variety of words). In a cross-sectional within-subject design, 437 learners from grades 10 and 12 (15-16 years old and 17-18 years old) completed a frequency-based form recall test and two informal writing tasks in both English and French. Lexical sophistication was measured using Compleat Web VocabProfile (Cobb, 2023) based on BNC and Lonsdale & LeBras frequency lists. Lexical diversity was measured using a Python script (Kyle et al., 2021) to calculate the moving average type-token ratio (MATTR).

Preliminary analyses ( $n = 188$ ) suggest a moderate correlation between productive vocabulary knowledge and lexical diversity of L2 writing in English ( $r = 0.456$ ;  $p < .001$ ) and French ( $r = 0.450$ ;  $p < .001$ ). A small to moderate correlation was found between English productive vocabulary knowledge and lexical sophistication ( $r = 0.344$ ;  $p < .001$ ), whereas no correlation was found for French ( $r = -0.042$ ,  $p = .569$ ). Further, a cross-sectional comparison shows that in both languages, grade 12 learners scored higher on the form recall test than grade 10 learners (for English  $d = -0.419$ ,  $p < .001$ ; for French  $d = -0.535$ ,  $p < .001$ ). They also used more varied vocabulary in writing than their younger counterparts (for English  $d = -0.688$ ,  $p < .001$ ; for French  $d = -0.406$ ,  $p = .006$ ). However, the proportion of low-frequency vocabulary only increased to some degree in their English writing ( $d = -0.312$ ,  $p = .034$ ), and did not rise in

---

\*Speaker

their French output ( $d = 0.148$ ,  $p = .311$ ). Our results only partially corroborate Laufer's (1998) findings. The presentation will conclude with some pedagogical implications.

**Keywords:** vocabulary, writing, L2 English, L2 French, productive vocabulary knowledge, lexical complexity, lexical diversity, lexical sophistication

# Pinning down the positive effects of translanguaging-based pedagogies on students' metalinguistic abilities

Valentina Carbonara <sup>1</sup>, Jacopo Torregrossa \* <sup>2</sup>

<sup>1</sup> University for Foreigners of Siena – Italy

<sup>2</sup> Goethe-University Frankfurt – Germany

Recent studies have shown that the implementation of translanguaging pedagogies (TPs) at school benefits students' achievements, including narrative (Carbonara et al., 2023) and metalinguistic abilities (MAs; Leonet et al., 2020) and foreign language learning (Hopp & Thoma, 2021). It has been proposed that TPs encourage students to reflect on differences/similarities between the languages of their repertoire, which enhances MAs (Cenoz & Gorter, 2021). However, no study so far really examined why TPs benefit MAs. This study fills this gap considering the effect of language proficiency and language-learning aptitude on students' MAs in the school language (Roehr-Brackin, 2018 for the relevance of these variables).

Sixty-seven students from four classes of a secondary public school in Italy (M-age:13;7) participated. The target group ( $n=32$ , 12 with a language of origin (LoO) different from Italian) was exposed to TPs for one school year, while the control group ( $n=37$ , 6 with LoO different from Italian) to monoglossic pedagogies (i.e., classroom activities only in Italian). Before the intervention, the participants took a test tapping into literacy skills in Italian (designed by the National-Education-Evaluation Institute), which did not reveal any difference across the groups. After the intervention, we tested participants' MAs in Italian with the TAM-2 test (Pinto et al., 2003), vocabulary proficiency in Italian with the PMA test (Rubini & Rossi, 1982) and language aptitude with LLAMA-v.3 (Rogers et al., 2023).

First, we run a generalized-linear-mixed-effects analysis with MAs in Italian (whether or not participants provided a metalinguistic explanation to each trial) as binary dependent variable and, as fixed effects, the interaction between group (target vs. control) and vocabulary, group and LLAMA-score, and LoO (Italian vs. other). Then, we conducted two causal mediation analyses with MAs as outcome variable and group as treatment. Vocabulary was the mediator in the first model and LLAMA-score in the second.

The GLMER-analysis revealed a significant positive effect of LLAMA-score and vocabulary and an effect of group, whereby the target exhibited better MAs. In the first mediation model, we found a direct effect of both groups on MAs. This effect was mediated by vocabulary in both groups, with the mediated effect being stronger in the control (based on the observation of the model's estimates). In the second model, there was a direct effect of both groups on MAs, which was mediated by the LLAMA-score, especially in the target.

The study showed that TPs enhanced students' MAs. The use of causal-mediation analyses demonstrated why this happened. Bearing in mind that both the direct and mediated effects were significant across both groups in both models, the magnitude of the estimate effects suggested that the mechanisms underlying the development of MAs differed across the two groups. In the target, TPs enhanced students' language-learning aptitude, which in turn boosted their

---

\*Speaker

MAs in Italian. Therefore, TPs seemed to enhance students' learning outcomes in the school language over and above the improvement of their proficiency in it. In the control group, development of MAs were more related to the development of proficiency in Italian.

**Keywords:** translanguaging pedagogies, metalinguistic abilities, language, learning aptitude, language proficiency, causal mediation analysis

# Cross Linguistic Influence in L3 acquisition: Do dominance and recency play a role?

Chloe Castle <sup>\*†</sup> <sup>1</sup>, Anna Skalba <sup>2</sup>, Marit Westergaard <sup>1,3</sup>

<sup>1</sup> The Arctic University of Norway [Tromsø, Norway] – Norway

<sup>2</sup> Adam Mickiewicz University, Poznań – Poland

<sup>3</sup> Norges teknisk-naturvitenskapelige universitet - NTNU – Norway

The literature is divided regarding whether dominance plays a role in crosslinguistic influence (CLI) in L3A. While some scholars have found that dominance is a factor (e.g., Fallah & Jabbari 2018), others have found that it is not (e.g., Lloyd-Smith et al 2021). This paper focuses on morphosyntactic CLI at the beginning stage of acquisition, before input-induced and overt learning has occurred. We also explore the effect of little-investigated factor of recency, operationalised as the language of instruction.

Participants are L1 Polish L2 English speakers in Poland or the UK, who do not know Norwegian. They are exposed to thirty-six Norwegian words (with pictures) as many times as desired, followed by a picture-sentence matching task. They proceed to the main experiment if they obtain at least 80% accuracy on the lexical items. The main experiment is a forced-choice judgement task, wherein participants must choose between two sentences in Norwegian using the words they learned – one with Polish-like and the other English-like morphosyntax. The constructions are ditransitives, articles, number agreement and semantic gender.

Dominance is operationalised using the Language and Social Background Questionnaire (Anderson 2017). Participants are also tested on their performance on the relevant structures in English and Polish. We predicted that participants dominant in Polish would choose Polish-like sentences, and those dominant in English would choose English-like sentences. These effects would be mitigated by the recency effect, i.e., Polish-dominant Polish-recency participants would choose the most Polish-like constructions, followed by Polish-dominant English-recency participants. Participants are expected to perform differently for different properties, as the difference in complexity and saliency for different grammatical constructions can affect CLI (Jensen et al 2023). However, it is unclear how the response variable would differ for these constructions.

We tested the combined effects of dominance and recency on participants' Polish-like choices using a generalised linear mixed effects model. No main effects were found for LSBQ score ( $p = 0.44$ ) or recency ( $p = 0.79$ ). Random effects accounted for most of the variance (conditional  $R^2 = 0.68$ , marginal  $R^2 = 1.13e-03$ ). We then conducted exploratory analyses, including construction as a fixed effect. In the Polish recency group, more Polish-dominant participants made more English-like choices in the article ( $p = 0.0002$ ) and ditransitive ( $p = 0.001$ ) conditions. For the number agreement condition, however, dominance in Polish led to more Polish-like choices ( $p = 0.02$ ). These differences were primarily driven by the Polish recency group in Poland. There may be less cognitive cost for these participants as they complete the experiment in their L1 and dominant language and are thus able to focus more on linguistic structures. Each of the significant constructions involve a morpheme which is 'added' – and participants more dominant in Polish opt for the sentence with an 'added' structure. For articles and ditransitives, this is a free morpheme (en 'a', til 'to'), and for number agreement it is a bound morpheme (-e

---

\*Speaker

†Corresponding author: chloe.castle@uit.no

PL). Additionally, we find that experiment language and construction choice impact participant behaviour, emphasizing their crucial role in experimental design.

**Keywords:** Cross Linguistic Influence, language dominance, instruction language, experiment design, L3 acquisition

# Examining the effect of contextual diversity on the processing and learning of new vocabulary

Radha Chandy \* <sup>1</sup>, Raquel Serrano \* † <sup>2</sup>, Ana Pellicer-Sánchez <sup>3</sup>

<sup>1</sup> Universitat de Barcelona – Spain

<sup>2</sup> Universitat de Barcelona (UB) – University of Barcelona Gran Via de les Corts Catalanes, 585 08007 Barcelona, Spain

<sup>3</sup> UCL, Institute of Education – United Kingdom

Incidental vocabulary learning requires multiple exposures to the same target words (Webb, 2007). There is not much evidence, however, as to whether novel words are learned more successfully if those multiple exposures appear in the same text read several times or in different texts. Encountering words in varied contexts seems to support the establishment of richer semantic associations (e.g., Ferreira & Ellis, 2016). Additionally, it has been claimed that reading the same text multiple times could lead to boredom and decreased attention, resulting in fewer vocabulary gains (Liu & Todd, 2016). The present study contributes to this line of research by being the first one to examine online processing during repeated reading of the *same* versus *different* texts through eye tracking, with two main goals. The first goal is to explore changes in L2 learners' attention as measured by eye fixations on repeated exposures to novel words in the two conditions (*same* vs. *different*), while the second one is to analyze the relationship between eye fixations and subsequent learning and retention of new vocabulary.

Forty-two adult Spanish EFL learners participated in two repeated-reading sessions involving short narratives in English that included two sets of 10 pseudowords. One set appeared in the *same* text read three times and another one in three *different* texts in a counterbalanced fashion. The pseudowords appeared a total of six times (twice in each text) and replaced high-frequency nouns. Online processing of the pseudowords was measured by total fixation duration, first fixation duration, average fixation duration and fixation count, as recorded with a Tobii Pro Spectrum eye tracker. Participants' knowledge of the target vocabulary was assessed after the last reading and one week later through form-recognition, meaning-recall, and meaning-recognition tests.

The results of the GLMMs suggest that repeated reading of the *same* text led to a more significant decrease in processing time across repeated encounters of the target pseudowords, as observed through the eye-tracking measures. The relationship between eye fixations on the pseudowords and their learning and retention was weak overall, but more significant in the *same* condition. The results will be discussed considering previous research on contextual diversity and vocabulary learning and some pedagogical suggestions will be offered.

Ferreira, R. A., & Ellis, A. W. (2016). Effects of contextual diversity on semantic decision and reading aloud: Evidence from a word learning study in English as a second language. *Studies in Psychology*, 37, 162–182

---

\*Speaker

†Corresponding author: raquelserrano@ub.edu

- Liu, Y.-T., & Todd, A. G. (2016). Implementation of assisted repeated reading techniques for the incidental acquisition of novel foreign vocabulary. *Language Teaching Research*, 20(1), 53–74.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.

**Keywords:** incidental vocabulary learning, repeated reading, processing, contextual diversity, eye tracking



# Network science: A new lens for studying L1 lexical attrition

Adel Chaouch-Orozco <sup>\*† 1</sup>, Fernando Martín-Villena <sup>2</sup>

<sup>1</sup> The Hong Kong Polytechnic University [Hong Kong] – Hong Kong SAR China

<sup>2</sup> Universitat Pompeu Fabra [Barcelona] – Spain

The weakening or loss of L1 lexical-semantic abilities, known as L1 lexical attrition, is often linked to reduced L1 exposure and/or L2 interference (Schmid et al., 2019). Semantic fluency tasks, where participants provide exemplars of a given semantic category within a timeframe, are central to L1 attrition research. However, traditional analyses have yielded inconsistent results (Schmid & Köpke, 2009). Some studies reveal that L1 attriters name fewer items than functional monolinguals, while others find no significant differences (Jarvis, 2019). Our study moves beyond traditional analyses, employing a novel network science approach to assess the bilingual lexicon’s structural properties, thus offering a new perspective in L1 attrition research. We gathered data from two semantic fluency tasks involving 94 immersed and 80 non-immersed late bilinguals with comparable L2 proficiency (as measured by the OQPT). In their L1 Spanish, participants provided exemplars of the fruits and vegetables category, whereas in L2 English, they named animals. Following previous studies (Christensen & Kenett, 2021), we built networks with nodes representing words and links capturing co-occurrences. Our analysis centred around three network measures reflecting the lexicon’s structural organisation, which may be critically impacted during lexical attrition (Gallo et al., 2021): Clustering coefficient (CC), the degree to which nodes group together; average shortest-path length (ASPL), the average distance between node pairs; and modularity (Q), the degree to which the network comprises distinct communities. Notably, higher CC is associated with better semantic organisation in bilinguals and monolinguals (Cosgrove et al., 2021; Feng & Liu, 2023); lower ASPL with faster navigability within the lexicon (Siew & Guru, 2023); and lower Q with more efficiently organised knowledge networks (Siew & Guru, 2023).

We examined whether L2 immersion results in L1 lexical attrition, manifested by shifts in the native semantic system’s structure, reflected by lower CC, higher ASPL, and Q. Employing bootstrap analyses, we generated partial networks from randomly selected node subsets, replicating these processes 1000 times for subsets containing 50% to 90% of the original network nodes. This created a sampling distribution for each measure and network. Through analyses of covariance (ANCOVA), significant differences emerged between the L1/L2 semantic networks of immersed and non-immersed bilinguals. The L1 networks of the immersed participants showed signs of attrition, such as reduced connectivity and greater sparsity. Conversely, their L2 networks displayed a more efficient organisation compared to their non-immersed counterparts. In-depth analyses examining language use and proficiency further validated these findings. Crucially, these trends were absent in traditional analyses like response count.

Our results demonstrate the efficiency of the network science approach in illuminating the complex dynamics of bilingual semantic memory systems. Furthermore, our analysis indicated that the erosion of the native semantic system unfolds gradually, first impacting network interconnectivity, with information flow and community structure being less affected initially. Drawing from

---

\*Speaker

†Corresponding author: adchaor@gmail.com

these insights, we introduce the Lexical Attrition Foundation (LeAF) framework, introducing a network-based developmental perspective on lexical attrition and laying the groundwork for future research.

**Keywords:** lexical attrition, network science, semantic representation, L2 immersion

# English exposure in the classroom and vocabulary development. Longitudinal data from a CLIL project

Luca Cilibrasi \* <sup>1</sup>, Tina Skočilová \*

2

<sup>1</sup> University of Siena – Italy

<sup>2</sup> Charles University – Czech Republic

This study delves into the longitudinal impact of Content and Language Integrated Learning (CLIL), a teaching method wherein students are instructed in some subjects using their second language (L2). Here we focus on English language development in Czech primary school children. The research spans five months and centres on two distinct groups of children: one group engaged in CLIL instruction for their mathematics and history classes, while the other received conventional English language classes only. Each group comprised eighteen participants aged 11-13, who underwent two rounds of testing. The assessments consisted of a 110-question vocabulary test adapted from Paul Nation's Picture Vocabulary Size Test (PVST) and the completion of 3 brief reading comprehension passages from YARC (Snowling et al., 2009). The same test was repeated five months later. The vocabulary test encompassed four sections: part I mirrored Paul Nation's original PVST test (Anthony & Nation, 2017), part II encompassed words related to everyday vocabulary, and parts III and IV focused on mathematics and history vocabulary. Following the completion of both rounds of testing, the two groups were compared, and statistical analysis was employed to assess the progress made in each group. The analysis revealed that while there were no significant differences in the overall vocabulary scores between the two testing sessions, the CLIL group demonstrated marginal improvements in everyday vocabulary, a progress not observed in the control group. Furthermore, the progress in reading comprehension skills was significantly more pronounced in the CLIL group.

This study is presented within the context of the existing second language literature that addresses issues related to language exposure and is integrated into the broader debate surrounding the optimal conditions for second language (L2) acquisition, as discussed by scholars such as Pfenninger & Singleton (2017) and Birdsong (2018). Our findings suggest that studying a subject in a second language does not yield dramatic enhancements in technical vocabulary, at least within a relatively short 5-month timeframe. However, this approach does play a crucial role in bolstering general L2 skills, which is reflected in improved reading comprehension abilities and, to a certain extent, in the expansion of general vocabulary.

## References

Anthony, L., and Nation, I.S.P. (2017). Picture Vocabulary Size Test (Version 1.2.0.) (Windows 10). Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/>

---

\*Speaker

Birdsong, D. (2018). Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in psychology, 9*, 81.

Pfenninger, S. E., & Singleton, D. (2017). *Beyond age effects in instructional L2 learning: Revisiting the age factor* (Vol. 113). Multilingual Matters.

Snowling, M. J., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., ... & Hulme, C. (2009). YARC York Assessment of Reading for Comprehension Passage Reading.

**Keywords:** vocabulary, reading, CLIL

# How well can LD measures predict speaking proficiency, given text length differences and different analysis units?

Jon Clenton \* <sup>1</sup>, Takumi Uchihara <sup>2</sup>, Simon Fraser <sup>1</sup>

<sup>1</sup> Hiroshima University – Japan

<sup>2</sup> Tohoku University [Sendai] – Japan

Using indices of lexical diversity (LD) to gauge a speaker’s lexical proficiency, an essential aspect of fluency and overall language proficiency, and has been in practice for some time. One issue, however, with many commonly used indices of lexical diversity (e.g., TTR) is that they vary as a function of text length. A further issue with such indices is that they also vary as a function of analysis unit (e.g., lemma or flemma; Myint Maw et al., 2022). The current study aims to investigate potential correlations between lexical diversity (using varied text lengths and analysis units), fluency of speech, and overall speaking proficiency. Our aim is to present findings from a study involving around 150 participants from a university in the United Kingdom, who were classified into three distinct levels of IELTS proficiency. We investigate a corpus of undergraduate presentations delivered by the participants. The purpose of our investigation is to analyze and contrast the impacts of flemma, lemma, and simple count on LD measure predictability in speaking. We employ three basic LD measures and three sophisticated measures for our analysis. Following earlier research (Treffers-Daller 2013; Treffers-Daller et al, 2018), we also examine different text lengths. By investigating whether results mirror recent findings for writing (Myint Maw et al., 2022) that both flemmatization and lemmatization influence LD scores, we examine potential relationships between LD and our other factors, and the extent to which LD is a predictor of general speaking proficiency when simple, flemma, and lemma counts are applied. Furthermore, we offer analysis on the impacts of different spoken text lengths. The focus of our discussion is on the influence of analysis units and text length on the predictions of utterance fluency at varying levels of general speaking proficiency.

References

Myint Maw, T.W, Clenton, J., & Higginbotham, G. (2022). Investigating whether a flemma count is a more distinctive measurement of lexical diversity. *Assessing Writing* (53). <https://doi.org/10.1016/j.asw.2022.100444>

Treffers-Daller, J. (2013). Measuring lexical diversity among L2 learners of French. In S. Jarvis M. Daller (Eds.), *Vocabulary Knowledge: Human ratings and automated measures* (pp. 79-104). John Benjamins.

Treffers-Daller, J., Parslow, P., & Williams, S. (2018). Back to basics: How measures of lexical diversity can help discriminate between CEFR levels. *Applied Linguistics*, 39(3), 302-327. <https://doi.org/10.1093/applin/amw009>

**Keywords:** Lexical Diversity, Flemma, Lemma, Speaking Proficiency

---

\*Speaker

# Looking into cognate recognition in intercomprehension: an eye-tracking experiment

Larissa D'angelo <sup>1</sup>, Ilaria Borro \* <sup>1</sup>, Saturno Jacopo \*

1

<sup>1</sup> Università degli Studi di Bergamo – Italy

The present study aims to investigate cognate recognition in the context of intercomprehension, a communicative mode in which the proximity of the respective languages allows interlocutors to understand each other despite their speaking their own L1s (European Commission 2012).

The experiment relies on an eye-tracking protocol, a technique that so far has been hardly applied to the study of intercomprehension. Indeed, while most existing research (e.g. Berthele 2011; Möller & Zeevaert 2015; Saturno 2023) is mainly concerned with the final output of the phenomenon, the present study rather aims to describe the process, with the goal to investigate whether cognate recognition in the context of intercomprehension activates a sequential or holistic reading mechanism (Grainger 2022). In the former scenario, target items are compared to similar items from known languages in search of correspondences and divergences. One would thus expect longer fixation times on those letters of the target item that diverge from similar items belonging to other languages; at the same time, different types of interlingual correspondences are expected to lead to different recognition rates. In the holistic reading scenario, on the other hand, fixation times should equally distribute across all letters of the stimulus item, with no observable difference attributable to the specific type of interlingual correspondences.

An artificial language was developed based on the assumption that the recognition of a word is significantly hindered by modifications concerning consonants, but not vowels (Schmandt et al. 2022; Berrebi et al. 2023). The stimuli were designed on the basis of 12 Italian words, selected from the PhonItalia corpus (Goslin et al. 2014) in such a way as to balance overall frequency and phonological neighborhood size and frequency. These words were then modified with respect to one, two or three articulatory traits (place of articulation, manner of articulation, voicing) of the initial consonant, e.g. It. *domanda* > *\*tomanda*, *\*momanda* etc. 18 distractors were also devised in which no systematic relation to the Italian counterpart could be identified.

The planned experiment will be carried out as follows. The stimuli (target items and distractors) will be presented to 20 L1 Italian participants in a sufficiently large font size for the eye-tracker (Tobii Pro Nano - 60 Hz) to detect fixations on individual letters. Participants will be asked to translate the unknown item into Italian. Fixation times (first fixation duration, first pass reading time, total reading time, total fixation count) will be computed for three regions of interest, i.e. a) the modified letter, b) the whole word, and c) the whole word with the exclusion of the first letter. Data collection and analysis will be conducted with iMotions 9.3.0.5 software. The poster discusses the rationale of the experiment and presents its preliminary results. It

---

\*Speaker

is expected that longer fixation times will correspond to initial consonants modified with respect to a greater number of articulatory parameters. Should it be confirmed, this observation would bring support to the hypothesis that cognate recognition in intercomprehension relies on sequential reading.

**Keywords:** cognate recognition, intercomprehension, eye, tracking, artificial language, orthographic processing

# Morphophonological ambiguities and automatic assessment of spoken L2 lexical forms for pedagogical purposes: a pilot study among Japanese learners of French

Sylvain Detey <sup>\*† 1</sup>, Verdiana De Fino <sup>‡ 2,3</sup>, Lionel Fontan <sup>\* § 4</sup>

<sup>1</sup> Waseda University (SILS GSICCS) – Japan

<sup>2</sup> Institut de Recherche en Informatique de Toulouse – Institut de recherche en informatique de toulouse, IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France – France

<sup>3</sup> Archean Labs – Archean Labs – France

<sup>4</sup> Archean Labs – Archean Labs – France

Until the recent emergence of AI-based tools on the pedagogical market, classic spellcheckers and grammar correctors were commonly used by L1 and L2 writers to automatically assess, correct, and improve their writing, both at the orthographic and at the grammatical levels. Thanks to such tools, L2 learners have been able to improve their writing skills in an autonomous manner through various activities. Speaking skills training, on the other hand, remains difficult. One of the reasons behind the lack of available tool for autonomous L2 speech assessment is linked to the variability between the phonetic, the phonological and the lexical levels of analysis that may be attached to the forms produced by the learner. This becomes even more challenging for automatic speech processing systems when the L2 contains cohorts of phonological neighbors (e.g., *mangent* (m) vs *mangeons* (m)) and homophones (e.g., *mangé* vs *manger* both as (me)). In the case of Japanese learners of French, the phonetic errors induced by interphonological features in learners' speech may lead to linguistic ambiguities that standard automatic speech recognition systems may not be able to solve (e.g. does (mleokola) correspond to "J'aime le chocolat" ('I like chocolate') with a phonetic substitution between (e) and () or "J'aime les chocolats" (instead of "J'aime le chocolat") with a morphophonological substitution between the plural and the singular form of "le chocolat"?). In this presentation we describe an attempt to tackle this challenge by using a computer-interfaced translation task using a double output modality (spoken and written) to assess the possibility of designing methodological guidelines to bypass the technical hurdle. In our study, 12 sentences containing typical morphophonological learning difficulties for Japanese learners of French were selected, and 40 learners were asked to translate them from Japanese to French, both in a spoken and in a written form, in a constrained double-task recording protocol. The collected data was then automatically (for the written production) or manually (by human evaluators) assessed and the results were compared to determine whether such method could help to build L2 speech learning applications allowing for translation-based automatic L2 spoken lexical form evaluation. Our results support our approach since our protocol allows us to identify the nature of the substitution between the target element and the learner's production, i.e., whether they are phonetic or morpholexical substitutions. This constitutes a methodological basis for pedagogical developments in L2 speech learning applications, with automatic L2 speech assessment at the lexical level.

---

\*Speaker

†Corresponding author: detey@waseda.jp

‡Corresponding author: verdiana.defino@irit.fr

§Corresponding author: lfontan@archean.tech



## References

De Fino, V., Fontan, L., Pinquier, J., Ferrané, I. & Detey, S. (2022). Prediction of L2 speech proficiency based on multi-level linguistic features. *Proceedings of Interspeech2022*, 4043-4047.

Detey, S., Fontan, L., Le Coz, M., Barcat, C. & Kawaguchi, Y. (2022). Identifying segmental substitutions in spontaneous speech of L3-French/L1-Japanese learners: a corpus-based pilot study. *Journal of the Phonetic Society of Japan* 26(3): 124-134.

Detey, S., Fontan, L., Le Coz, M. & Jmel, S. (2020). Computer-assisted assessment of phonetic fluency in a second language: a longitudinal study of Japanese learners of French. *Speech Communication* 125: 69-79.

**Keywords:** Automatic L2 speech assessment, Morphophonological ambiguities, Japanese learners of French, Translation, Phonological lexicon, Language Learning Technology

# Awareness of orthographic transparency benefits the lexical encoding of mid-vowel contrasts in Portuguese L2 words

Isabelle Darcy \*<sup>1,2</sup>, Hunter Brakovec<sup>1</sup>

<sup>1</sup> Indiana University Bloomington – United States

<sup>2</sup> Université Grenoble-Alpes, Laboratoire LIDILEM – Université Grenoble Alpes – France

Non-native speakers often have difficulty accurately encoding sound contrasts in lexical entries (1). Lexical storage may still be fuzzy even when learners can accurately perceive the difference between sounds (2). Previous novel-word learning studies suggest that orthography may improve learners' lexical encoding; however, findings are mixed (3). We explore the potential effect of (traditional) learners' attention to orthography on lexical encoding, rarely examined previously.

In Brazilian Portuguese (BP), mid-vowels can be close (/e/-/o/) or open (/E/-/C/), and they pose difficulties for English learners of BP in perception and production (4). Some words carry an accent mark (*férias* /E/ 'vacation'; *êxito* /e/ 'result'), transparently indicating vowel quality (close: < ê, ô > ; open: < é, ó > ), but for unmarked words (), vowel quality is opaque (*festa* /E/ 'party'; *maleta* /e/ 'suitcase'). If accent marks help learners encode words with the expected vowel quality at the time of learning, learners' lexical encoding accuracy will be higher on marked than unmarked words, but only for those who are aware of the accents' reliability.

We tested 40 participants (20 native BP speakers; 20 L2 learners) on a speeded auditory lexical decision task (LD) and obtained measures of perceptual discrimination (AX task) and orthographic awareness (OA; accent mark identification). LD-stimuli were familiar words (/fEsta/) and non-words with a switched mid-vowel (Target Contrast, \*/festa/). Control conditions contained unrelated real words (distractor) and non-words (Unrelated Contrast, \*/fEspa/). 50% of selected words carry an accent ('marked'); the others are 'unmarked'. No spelling information was displayed during the LD task; OA tasks were administered later. If an effect is due to OA of accent marks, learners with higher OA should display more accurate lexical encoding on words with accent marks than those without, whereas learners with lower OA may show no effect of marking. We preregistered the study on OSF and will make the data and analysis accessible.

Both groups discriminated between open and close vowels well (about 80% accuracy), and their d' scores indicated medium sensitivity to the open/close distinction for all vowels. In LD, learners had difficulty rejecting non-words with the opposite vowel (about 20% accuracy) but performed > 75% accuracy in all other conditions. There was no overall effect of marking – but as predicted, an interaction between OA group and marking: the "high-awareness" learners were more accurate for marked (38%) than unmarked items (23% accuracy). We discuss implications for vocabulary learning and teaching, and how phonotactic predictability additionally interacts with non-word rejection.

---

\*Speaker

## References

- Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, 34, 269-284.
- Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *Mental Lexicon*, 8(3), 372-420.
- Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, 54(3), 297-326.
- Diaz-Granado, M. (2010). *L2 and L3 acquisition of the Portuguese stressed vowel inventory by native speakers of English* (Unpublished doctoral dissertation). University of Arizona.

**Keywords:** phonolexical representations, orthographic awareness, L2 Portuguese, L1 American English, mental lexicon, phonology

# Word Segmentation and Reading Comprehension among French Learners of Chinese as a Second Language

Rachel Daveluy \* <sup>1</sup>

<sup>1</sup> Université Paul-Valéry Montpellier 3 - Faculté de Langues et cultures étrangères et régionales –  
Université Paul-Valéry - Montpellier 3 – France

Most alphabetic writing systems use a blank space to mark the boundary of a word, enabling rapid segmentation of a sentence into lexical units. Chinese, however, has a non-alphabetic writing system with no delimiter between words. A Chinese word can consist of one or more characters. Although about 6% of Chinese words are single-character words (Gu et al 2022), most words in modern Chinese are composed of two characters (Alleton 2008). Those characters can appear in any position in the word, and it is possible to write a very large number of words using a small number of frequently used characters (Allanic 2017). A Chinese written sentence comprises a regularly interspaced string of characters. Therefore, a chain of characters can be segmented in different ways and the reader must associate characters together to form words depending on context, lexical and syntactical knowledge. However, Chinese native readers as well as Chinese L2 learners do not agree among themselves on how to segment the sentence into lexical units (Li et al 2013; Deng 2020). Thus, in Chinese, reading is a two-stage process, involving character identification and lexical unit recognition (Allanic 2017). The L2 reader then faces two major challenges regarding the segmentation of sentences into words: word ambiguity and lack of lexical knowledge (Huang and Wue 2012; Li and Yuan 1998). For L2 teachers, the challenge is to teach learners how to better identify lexical units.

Given the specificities of the Chinese writing system, we seek to discern the challenges encountered by French students with a beginner level in L2 Chinese reading comprehension activities. We conducted 19 interviews with Chinese teachers in France (10 teaching in secondary schools, 4 at university and 5 teaching both in secondary schools and university), aiming at shedding light on teaching practices and strategies related to the identification of lexical unit boundaries during reading comprehension activities. We also distributed surveys to 203 French secondary schools-students from first year to fourth year of L2 Chinese and 72 second year university L2 Chinese learners to better understand the students' issues during the reading comprehension activities. The purpose of those surveys was to investigate the perception of the word segmentation process among beginners and non-beginner learners.

Based on those interviews and surveys we try to answer the following questions: (a) What are the teaching practices among French L2 Chinese teachers regarding word segmentation during the reading comprehension activities? (b) What are the common issues faced by French L2 Chinese learners regarding the word segmentation process and the understanding of a Chinese text? (c) How can we improve the teaching practices and strategies to facilitate reading comprehension among French L2 Chinese learners at different proficiency levels?

---

\*Speaker

**Keywords:** L2 chinese learners, word segmentation, identification of lexical unit boundaries, teaching practices

# On the interpretation of null and overt anaphoric pronouns: a pilot study with Italian native, second language and heritage speakers

Elisa De Cristofaro \* <sup>1</sup>

<sup>1</sup> Universiteit Gent – Belgium

Anaphoric pronoun resolution is a vulnerable interface phenomenon in second language and bilingual acquisition (Montrul, 2012; Sorace & Filiaci, 2006; Tsimpli et al., 2004). Existing literature highlights that (various profiles of) bilinguals typically align with monolingual controls with null pronouns but tend to overuse and over-interpret overt pronouns (Di Domenico et al., 2020; Rinke & Flores, 2019 a.o.). These disparities can be attributed to several language-internal and language-external factors, (Belletti et al., 2007; Carminati, 2002; Filiaci et al., 2014; Giannakou, 2023; Keating et al., 2016; Lozano, 2018), including differences in the quality and quantity of the received linguistic input, dominance (Torregrossa & Bongartz, 2018), and age of onset (Di Domenico & Baroncini 2019).

This study examines anaphoric pronoun interpretation by comparing Italian monolingual native speakers (L1S) with two profiles of bilinguals with different ages of onset: heritage speakers (HS) and L2 speakers (L2S) with advanced or intermediate proficiency (first/community language: Belgian-Dutch). Participants realized an offline sentence comprehension task. Bi-clausal forward and backward anaphora sentences were followed by a question on the interpretation of either a null or an overt pronominal subject, which could ambiguously refer to the subject (Sbj), the complement (Compl) of the main clause or – in the case of overt pronouns - to an external (Ext) referent in the discourse (cf. Contemori et al., 2019).

The analysis of the results through descriptive statistics shows that the Sbj antecedent is preferred for interpreting a null pronoun across groups of speakers. With an overt pronoun, L1S strongly prefer the Compl antecedent, while L2S lack a strong preference, opting for either Compl or Sbj (in line with Tsimpli et al. 2004). Advanced L2S more frequently choose the Ext antecedent than intermediate L2S. Interestingly, HS exhibit a pattern similar to L2S, lacking a clear preference for Compl or Sbj, and favor the Ext antecedent more often than advanced L2S. Relevant differences emerge in backward anaphora with an overt pronoun: in this condition, monolinguals prefer the Ext antecedent, while HL and L2S both favor the Sbj antecedent (cf. Belletti et al., 2007). However, HS select the Ext antecedent more than L2S and, among the L2 group, advanced speakers prefer it more than intermediates.

The observation that bilinguals and monolinguals exhibit greater similarity in the null condition than in the overt pronoun condition aligns with numerous of the studies mentioned above, confirming that both L2S and HS are sensitive to the null vs. overt distinction on a pair with controls. Contrary to the production findings of Di Domenico & Baroncini (2019), our interpretation data does not reveal a clear-cut distinction between speakers exposed to Italian from birth (L1S and HS) and late learners (cf. Kaltsa et al., 2014). However, the greater variation shown by HS in the overt pronoun condition calls for further investigation into factors such as

---

\*Speaker

age, language exposure, and dominance. This work serves as a pilot for a more comprehensive study in which, combining interpretation and production data through correspondence analysis, the interplay between these variables will be addressed.

**Keywords:** anaphora resolution, Italian heritage language, second language acquisition, age of onset

# Acquiring nasal vowels in French: an acoustic and perceptual study on Belgian Dutch learners' pronunciation

Hanna De Haes <sup>\*† 1</sup>, Ellen Simon <sup>1</sup>, Peter Lauwers <sup>1</sup>

<sup>1</sup> Department of Linguistics, Ghent University, Blandijnberg 2, 9000 Ghent, Belgium – Belgium

*Background and research aims* — Several models have been proposed to study the production/perception of non-native (L2) phonemes by L2 learners (e.g., L2LP, Van Leussen & Escudero, 2015; SLM-r, Flege & Bohn, 2021). These models generally assume that category creation for an L2 sound partly depends on the acoustic-phonetic differences and similarities between L1 and L2 categories. In this light, the production of nasal vowels by L2 learners is a particularly interesting subject, because unlike nasal consonants, nasal vowel phonemes are relatively uncommon in the world's languages. In fact, the phonemic contrast between oral and nasal vowels exists in only 22,6 % of them (Maddieson & Precoda, 1990). To our knowledge, little research has focused on the production of nasal vowels by L2 learners (but see e.g., Detey, 2012; Inceoglu, 2016). This paper aims to contribute to our understanding of the acquisition of nasal vowels by investigating how Belgian Dutch learners produce French nasal vowels. While Dutch has no nasal vowels in its inventory, present-day Hexagonal French has three, being ton, teint and temps. Several scholars have indeed reported that L1 Dutch learners of French have difficulties producing those vowels (e.g., Berns & Nouveau, 2016).

*Method* — This study is innovative in that we focus on both listener-based and acoustic measures, thus shedding light on two different dimensions of production accuracy, i.e. vowel intelligibility and phonetic nativelikeness. Vowels were produced in high-frequency monosyllabic French words by 20 L1 Belgian Dutch speakers. First, 71 L1 French listeners performed an on-line identification task that assessed listeners' actual understanding of non-native accented nasal vowels (cf. vowel intelligibility). Listeners represented two French-speaking regions, namely the Île-de-France region, traditionally considered as the center of the French-speaking world, and Liège, a peripheral region often associated with a pronunciation that is distinct from the hexagonal norm. Using a mixed-effects model, we examined the intelligibility of Dutch-accented nasal vowels. Secondly, the learners' vowel productions were acoustically analyzed and compared to Carignan's (2014) data on native productions of French nasal vowels (cf. phonetic nativelikeness).

*Results and discussion* — The results of the perceptual experiment first indicated that listeners' overall understanding of non-native accented nasal vowels was rather high, except for the temps-vowel. Further analysis revealed that more closed and thus less canonical productions were less often identified. This result is associated with an ongoing sound shift affecting nasal vowels in present-day hexagonal French, leading to the reduction of the phonological contrast temps-ton (Hansen, 2001). Moreover, the two groups of listeners did not differ in their actual understanding of non-native accented nasal vowels, except for the teint-vowel, which was identified significantly more often by listeners representing the Liège region. Two likely causes for the observed difference are the impact of (i) L1 regional variation and (ii) Belgian listeners' familiarity with Dutch-accented speech on speech perception. Finally, the study has implications

---

\*Speaker

†Corresponding author: hanna.dehaes@ugent.be



for pronunciation teaching, as it reveals the particularities of Dutch-accented nasal vowels both from an acoustic and perceptual perspective.

**Keywords:** L2 phonology, nasal vowels, production, French, Dutch

# Impact of cross-linguistic divergences on the acquisition of it-cleft sentences in L2 Italian and L2 French: an experimental study

Bianca Maria De Paolis \* <sup>1,2</sup>

<sup>1</sup> Structures Formelles du Langage (SFL) – CNRS : UMR7023 – France

<sup>2</sup> Università degli studi di Torino = University of Turin – Italy

Cleft sentences, such as "C'est Marie qui achète le journal" and the Italian equivalent "è Maria che compra il giornale", are part of the inventory of marked word-orders in Italian and French. These structures fulfill specific functions, mostly related to focus (De Cesare, 2017 and Katz, 2000 among others), and they cross-linguistically show asymmetries as per their frequency with different types of constituents, delineating distinctions among subjects, objects, and adverbials (see Samo & Merlo, 2021 and references therein).

Despite these shared traits, several studies highlight noteworthy variations in the overall frequency of it-clefts between Italian and French (Roggia, 2008). Scholars widely attribute this dissimilarity to divergent levels of word-order flexibility (Lambrecht, 2001) and typological parameters, such as pro-drop vs non pro-drop tendencies (Belletti, 2008). In addition, insights from recent studies (Garassino, 2022 and De Paolis, 2023) have highlighted another layer of diversity: in French, it-clefts emerge as the default strategy for focus in both identification and correction contexts, while in Italian, their specialization leans more towards correction.

Motivated by these linguistic disparities, our study delves into the impact of these differences on the acquisition of cleft sentences by learners of L2 French and L2 Italian. The investigation encompasses three key facets: the overall frequency of it-clefts (H1); the distribution of it-clefts based on the type of target constituent (subject, object, or adverbial) (H2); and the prevalence of it-clefts relative to focus subtype (identification or correction) (H3).

To conduct this analysis, we built a corpus of task-elicited speech, drawing from two native groups-15 native French speakers and 15 native Italian speakers-as well as two learner groups-15 French speakers of L2 Italian and 15 Italian speakers of L2 French. Both non-native groups are made up of adult speakers, acquiring their L2 through non-guided learning; L2 proficiency levels span from beginners (A1-2) to advanced (C1-2). The resulting dataset comprises a total of 1500 utterances (375 for each group).

Our results show that both learner groups demonstrate an approximation to the target language as for the overall frequency of it-clefts (H1). However, nuances surface for hypotheses (H2) and (H3), indicating an influence from the L1s. Learners, in fact, exhibit a discernible sensitivity to the constraints of their native languages, particularly regarding the role of focus subtype (identification vs correction).

Adding a layer of granularity to the analysis, the examination of sub-groups based on different proficiency levels reveals intriguing subtleties. This finer exploration offers insights into how proficiency levels impact the learners' adoption of it-cleft structures, outlining a non-linear

---

\*Speaker

path: as one moves towards more advanced levels, the behavior of learners is more homogeneous and compact, but not necessarily more target-like. In early stages, cleft structures' complexity may in fact hinder usage, regardless of their pertinence to the information context. Proficiency development diminishes this complexity barrier, leaving space for cross-linguistic influences to emerge in terms of frequency and context of use of clefts.

**Keywords:** cleft structures, L2 Italian, L2 French, information structure, focus

# Pronunciation and discrimination develop independently whatever the explicit or implicit mode of teaching.

Nathalie Dherbey Chapuis \* <sup>1,2</sup>, Raphael Berthele<sup>†</sup> <sup>3,4</sup>

<sup>1</sup> Institut de plurilinguisme – Switzerland

<sup>2</sup> Université de Fribourg = University of Fribourg – Switzerland

<sup>3</sup> Université de Fribourg = University of Fribourg – Switzerland

<sup>4</sup> Institute of Plurilingualism – Switzerland

This study compares the influence of two teaching methods on the development of auditory discrimination and pronunciation in French as a foreign language in teenagers. On the one hand, training of pronunciation, explicitly (Kissling, 2018) or communicatively (Saito & Wu, 2014) was shown to enhance perception of second language (=L2) words. On the other hand, auditory training of L2 sounds can induce little improvements in their pronunciation but the relationship between perception and production progress is not significant (Sakai & Moorman, 2018).

In this study in French as a foreign language, we compare the effects of explicit teaching (i.e., focus on forms with deductive elaboration of rules and practice) and implicit teaching (i.e., focus on meaning with systematic conversational recast) on the learning of pronunciation and discrimination. The two targeted French phonemes don't exist in the learner's repertoires. The nasal vowel // and the voiced fricative // are known to be difficult to learn for L1 German speakers.

Participants were 82 Swiss German teenagers (mean age 12;6) learning French as a foreign language in the compulsory state school curriculum in Switzerland (level A1-A2 in French). In each of the six classes, the learners were randomly assigned to one of the two groups that followed either explicit or implicit instruction for six weeks (total lasting of instruction 4 hours).

The pronunciation and the discrimination tasks were administered at pretest and as an immediate post-test. In the pronunciation task, participants had to repeat stimuli immediately after hearing them. Comprehensibility of their pronunciation was evaluated by three raters on a 5-points Likert scale. Progress in pronunciation was operationalized as the difference between pre- and post-test in the score in comprehensibility (Poisson distribution). In the discrimination task (=AX), participants had to decide immediately after hearing them whether the two stimuli were the same or different. The variable "progress in discrimination" is the difference between pre- and post-test in dprime (according to signal detection theory: Makowski, 2018; DeCarlo, 2012).

The link between pronunciation and discrimination was explored with pronunciation's progress as the dependant variable and progress in discrimination as independent variable, and vice versa. The response data for pronunciation were analysed with Robust linear mixed models (Koller, 2016) and Poisson Bayesian models (Bürkner, 2021), and for discrimination with Robust linear mixed models and Gaussian Bayesian models (Bürkner, 2021).

---

\*Speaker

<sup>†</sup>Corresponding author: raphael.berthele@unifr.ch

Overall, the results indicate that participants made significant progress in pronunciation and discrimination, with no significant influence in either direction between development of pronunciation or discrimination. Furthermore, the mode of teaching has a differential impact on the progress in discrimination and pronunciation depending on the targeted phoneme. For //, implicit teaching led to a significantly greater improvement in discrimination. For //, phonological form-focused teaching led to a significantly greater improvement in pronunciation.

To sum up, the pronunciation and discrimination seem to develop independently whatever the explicit or implicit teaching methods. However, both modes of teaching are necessary to enhance the development of these two competencies in the two targeted phonemes.

**Keywords:** pronunciation, discrimination, teaching, explicit versus implicit methods, French as a foreign language, teenagers

# The (not so) Simple View of Reading in bilingual and monolingual children: the impact of vocabulary, verbal working memory and literacy

Ifigenia Dosi <sup>\*†</sup> <sup>1</sup>, Evgenia Karadimoula <sup>\*</sup>

<sup>1</sup>, Klisilda Loubonia <sup>\*</sup>

<sup>1</sup>, Zoe Gavriilidou <sup>\*</sup>

1

<sup>1</sup> Democritus University of Thrace – Greece

The aim of the present study is twofold: to investigate reading skills (a) under the light of *Simple View of Reading* (SVR) in bilingual and monolingual children and (b) considering the impact of other factors. According to SVR, reading consists of decoding and language comprehension. Several studies have shown that reading comprehension in bilinguals is lower compared to monolinguals. These differences are not due to decoding (an ability that may be more developed in bilingual speakers) but due to language comprehension, and mostly due to their lower vocabulary size in each language (Melby-Lervåg, & Lervåg, 2014). Previous studies exhibited that the quantity of words a speaker knows correlate with reading comprehension in monolingual and bilingual children (Harkio & Pietilä, 2016). Research, so far, revealed that verbal working memory (VWM) affect vocabulary and morphosyntactic development in monolingual children and thus reading skills are also affected. Contextual factors, predominately (early) literacy, affect both vocabulary development and reading comprehension (Papastefanou et al., 2021). From the above we deduce that reading comprehension is a skill that is affected by many factors and seems to work in a similar, but in different, way at the same time, in bilingual and monolingual speakers. To date, our knowledge is limited about the interconnection of the aforementioned variables, particularly in bilingual children. Forty children (9-12 years old) took part in the present study. Twenty of them were Greek-speaking monolingual children (ML) and twenty were Greek-Turkish bilingual children (BL), who live in Greece. All participants were assessed by means of a fluid intelligence task (Raven et al., 2008), an expressive vocabulary task (Vogindroukas et al., 2009), a VWM task and a reading test (Padeliadou & Antoniou, 2007). Information about bilingual speakers' input and literacy practices were collected through a background questionnaire. Our results revealed that ML had larger vocabularies than BL. Nevertheless, they did not differ in their non-intelligence and VWM abilities. In relation to their reading skills, ML read more words per minute and they uttered them more accurately than BL. Moreover, ML exhibited higher language comprehension than their BL peers. Correlations within each group were found between (a) accurate word reading and vocabulary and (b)

---

\*Speaker

†Corresponding author: idosi@helit.duth.gr

vocabulary and accuracy in answering in comprehension questions in both groups. In addition, word reading per minute correlated with VWM and reading comprehension but only in the ML group. In the bilingual group, emergent literacy in Turkish was found to relate with reading comprehension and current literacy in Greek is correlated with accurate word reading. From our findings we may infer that although ML read faster and more accurate vocabulary and both aspects of reading (decoding and language comprehension) are interconnected in the same way in ML and BL speakers. VWM seem to be more involved in ML, compared to BL who may encounter an overload in their VWM by decoding, retrieving word meanings and updating context. Finally, reading skills are affected by early and current literacy practices, verifying that skills can be transferred from one language to the other.

**Keywords:** bilingualism, reading, vocabulary, working memory, literacy

# Speeded C-Test - a better predictor of oral proficiency?

Anastasia Drackert \* <sup>1,2</sup>, Franziska Möller \* <sup>† 3</sup>, Anna Timukova \* <sup>‡ 4</sup>

<sup>1</sup> Ruhr University Bochum = Ruhr-Universität Bochum – Germany

<sup>2</sup> Society for Study Preparation and Test Development (g.a.s.t. e.V.) – Germany

<sup>3</sup> Society for Study Preparation and Test Development (g.a.s.t. e.V.) – Germany

<sup>4</sup> Society for Study Preparation and Test Development (g.a.s.t. e.V.) – Germany

C-Tests as highly efficient and reliable tools for measuring global language proficiency have recently gained much attention in SLA research (e.g. Grotjahn & Drackert, 2022; Norris, 2020). In a C-Test, testees are required to fill the gaps in partially deleted words in short texts. The new time-reduced Speeded C-Test (S-C-Test) appears to promise even more in terms of efficiency, but has not yet been sufficiently investigated. In particular, many questions regarding the underlying construct of a S-C-Test remain unanswered. It is hypothesized that, in contrast to a canonical C-Test with a generous time limit of five minutes primarily measuring the amount of learners' declarative and procedural knowledge, a S-C-Test additionally gauges the level of automaticity of their skills (Grotjahn, 2010) and should thus be better at predicting learners' oral proficiency. So far, no evidence is available to back up this claim.

In the presentation, we report on partial results of a large study that aims to investigate the role of the time variable in the construct of computer-administered C-Tests across different levels of proficiency in English, German and Russian.

A total of over 200 L2 English learners, 180 L2 German learners and 50 L2 Russian learners took a canonical and a Speeded C-Test along with eight instruments measuring their declarative and procedural knowledge, their general oral language proficiency (oral elicited imitation test, OEIT) and typing skills. We are going to present findings for the English and German samples on three of our five research questions: How does the time variable influence the reliability of computerised C-Tests? How does the time variable influence learners' scores depending on their proficiency level? How does the time variable influence correlations between a C-Test and an integrated measure of oral proficiency? Answers are gained through ongoing reliability and correlation analyses as well as ANCOVA.

For the English sample, we found a slightly higher reliability coefficient for the S-C-Test than the C-Test (0.915 vs. 0.904). As expected, learners' scores were lower on the S-C-Test than on the canonical C-Test: a statistical significance was found for the mean differences in the total score also when typing speed (words per minute) is accounted for. Further analyses that will be conducted within the next months will show whether the difference is moderated by the learners' level of proficiency and C-Test-text difficulty. We are also going to correlate the score on the two differently timed C-Tests with the score on an OEIT to test the hypothesis that the S-C-Test is better at predicting learners' oral skills. The same analyses are also going to be conducted for the German sample.

Thus, the study contributes not only to our understanding of the C-Test construct but also

---

\*Speaker

<sup>†</sup>Corresponding author: moeller@gast.de

<sup>‡</sup>Corresponding author: timukova@gast.de



to our understanding of new constructs and assessment formats measuring different linguistic sub-skills in different languages.

**Keywords:** general language proficiency, c, test, elicited imitation test

# Unraveling the ‘dynamic turn’ in L2 motivation research: A systematic review of L2 motivation change in SSCI-indexed journals from 2000-2023

Chenxi Du \* <sup>1</sup>

<sup>1</sup> University of International Business and Economics – China

In the last decade, L2 motivation research has witnessed a shift towards the ‘dynamic turn’ because L2 motivation is complex, context-bound, and goes through constant changes. Against the backdrop, a number of empirical studies revolving around L2 motivation changes have been conducted in the Second Language Acquisition (SLA) field. Despite the significance of L2 motivation change no systematic reviews pertaining to L2 motivation research have been made and the theoretical framework, methodology and research foci in L2 motivation change research differ significantly from those in L2 motivation research. Therefore, a systematic review on L2 motivation change is needed to obtain a deeper understanding of the “motivational turn” in L2 motivation research. To address the research gaps, the present study makes a systematic review of 115 articles published in SSCI-indexed journals between 2000 to 2023 to summarize the trends of theoretical framework, methodology of the existing research on motivation changes. In addition, the systematic review makes a thematic analysis of the studies’ findings to summarize the individual-related, context-related and environment-related antecedents, patterns and impacts of L2 motivation changes. Based on these findings, the research summarizes the research gaps and future directions in L2 motivation changes research. The study has found that L2 motivation change research adopt a diverse theoretical framework (CDST, PCR, L2MSS, SDT), multiple methodology (quantitative including Latent Growth Curve Modelling, Cross-lagged analysis, qualitative including Q method and mix-methods design), features an over-emphasis on college students in EFL contexts. The review also finds that L2 motivation changes can be attributed to individual-related factors such as personal experience, basic need satisfaction/ frustration and context-related experience such as the change of school years or semesters, class task as well as environmental factors such the change of education system. In addition, the review has demonstrated L2 students show complex and idiosyncratic patterns in their motivation changes, which may be associated with their learning achievement, learning engagement. The study provides important insights for L2 motivation research as it unravels the complex nature of L2 motivation and deepens the understanding of the longitudinal perspective in L2 motivation research.

**Keywords:** dynamic turn, L2 motivation change, SSCI, indexed journals, systematic review

---

\*Speaker

# What linguistic features distinguish and predict CSL writing quality?

Yueming Du \*<sup>1,2</sup>, Shue Sum Leung \*

2

<sup>1</sup> Peking University – China

<sup>2</sup> University of Cambridge [UK] – United Kingdom

**Introduction.** Lexical knowledge, an important construct in language acquisition, plays a critical role in assessing the language proficiency of second language (L2) learners (Laufer & Nation, 1995; Read, 2000). Among numerous features employed to assess L2 learners' lexical competence, lexical richness has received considerable attention. Particularly, English-as-a-Second-Language (ESL) writing studies have developed a range of indicators to represent lexical richness. Conversely, research on lexical richness in Chinese-as-a-Second-Language (CSL) writing is still scarce and with limitations. First, past studies only used small-scale corpora, casting doubt on the generalizability of the results. Second, they extracted lexical features manually, resulting in small feature sets which may not suffice to represent the intricacy of lexical richness. **Aims.** The current study attempts to examine the relationship between CSL writing quality and lexical richness by constructing a feature set covering the four dimensions of lexical richness based on Read (2000), which are diversity, sophistication, density and errors. We seek to answer the following questions: (1) What lexical richness features would vary across different levels of writing scripts? (2) Which lexical richness features can best predict L2 writing quality of scripts?

**Methodology.** For this study, we compiled CSL writing data from the HSK Dynamic Composition Corpus, which consists of graded and error-coded writing scripts by CSL candidates sitting the official Chinese Proficiency Test from 20 different countries/regions. After determining the mark boundaries for scripts for three different levels (high: 80-95; mid: 60-75; low: 40-55), we randomly selected 1,000 writing samples for each level, resulting in 3,000 texts for further analysis. The texts were then pre-processed by LTP 4 (Che et al., 2021) and analysed with a Python programme to extract 64 features corresponding to the four aspects of lexical richness (diversity: 16; sophistication: 36; density: 6; error: 6).

**Results.** One-way ANOVAs revealed that 61 measures discriminated scripts of different levels (i.e. low vs. mid and/or mid vs. high). Key findings were: (i) in 11 indices of lexical diversity (e.g. the number of word types), high-level scripts were significantly greater than mid-level scripts; (ii) in 18 indices of lexical sophistication (e.g. the portion of low-frequency words), mid-level scripts displayed a significant difference from low-level scripts; (iii) 4 lexical density measures (e.g. functional word per lexical word) significantly differentiated mid-level scripts from both high-level and low-level scripts; (iv) 5 measures related to lexical error (e.g. wrong vocabulary usage per word) exhibited significant differences between high-level and mid-level scripts. A further stepwise regression against writing marks with these features found that 8 variables (including the number of word types and the portion of high-frequency vocabulary) could explain 44.14% of the variance.

---

\*Speaker

Discussion & Conclusion. These results confirm a significant role of lexical richness on CSL writing quality. They further reveal differential predictive power of different types of lexical richness indices for different levels of examination scripts, and that the number of types contribute the most to writing quality. The implications for the teaching and assessment of CSL writing are considered.

**Keywords:** lexical knowledge, lexical richness, writing quality, lexical acquisition

# Implicit Learning of a Semi-artificial Language by Chinese L1 Learners: Exploring the Acquisition of Pragmatics

Xuewen Duan \* <sup>1</sup>

<sup>1</sup> Department of Education, University of Oxford – United Kingdom

Implicit learning is a critical topic in the field of second language acquisition. While researchers have found that learners could implicitly learn the linguistic forms and establish form-(semantic) meaning mappings following incidental exposure to the linguistic input (e.g., Williams, 2005; Rebuschat & Williams, 2009; Li et al., 2013), implicit pragmatic learning remains an under-explored area (Ellis & Roever, 2021; Luo, 2021). This study aimed to investigate whether learners could learn honorific constraints on the usage of pseudo verbs under incidental learning conditions. Given the mixed results regarding the relative effectiveness of explicit and implicit L2 learning, this study also assessed whether learners who were aware of honorific constraints performed better than those who remained unaware. A semi-artificial language which consist of 4 pseudo verbs (hu, ke, li, na), Chinese lexis and novel syntactic structures was adopted in the study. The pseudo verbs all indicated the action of "transfer", and two of them were honorific forms while the other two were plain forms. A total of 90 Chinese L1 learners (experiment group 1=30, counterbalanced experiment group 2=30, control group=30) participated in the study. These learners did not know any language with highly developed honorific systems (e.g., Japanese, Korean). In the training session, participants were asked to attend to the linguistic forms (syntax, and the novel verbs' pronunciation and orthographic form) as well as the semantic information which influenced the use of novel verbs. Therefore, learners had no intent to learn the pragmatic cues. An aural, time-pressured Discourse Completion Task (DCT) and an interlocutor-judging task (judging the interlocutors of conversation based on pragmatic cues embedded in the novel verbs) were used to assess whether learners have acquired the honorific cues and other knowledge. Learners' pragmatic awareness was measured by trial-by-trial confidence ratings and the source attribution system (Dienes & Scott, 2008). The results showed that the experimental learners did not outperform the control learners. Despite an overall learning effect detected for the semi-artificial language, there was no evidence that learners had acquired the pragmatic cues implicitly. It was likely that the learning was impaired by minimal use of attentional resources and affected by various levels of concept availability (Williams, 2013; Jiang, 2016). That is, not all concepts were equally cognitively salient and learnable. Additionally, pragmatic awareness only seemed to contribute to marginally better performance in these tasks. A possible explanation was that learners were not fully aware of the fact that the usage of novel verbs was shaped by honorifics. Instead, the learners who became pragmatically aware recognised concepts such as "social status" and "age"-the most cognitively salient honorific-related concepts based on learners' reports. Inadequate exposure to the semi-artificial language could also have contributed to these results. More research is needed in terms of implicit pragmatic learning. Standardised tests for implicit pragmatic knowledge are urgently required in the field as well.

---

\*Speaker

**Keywords:** Implicit learning, pragmatics, pragmatic processing, honorifics

# Cross-linguistic gender congruency effects in sentence context depend on gender transparency and L2 proficiency

Kamil Długosz \* <sup>1</sup>, Anna Olszewska <sup>1</sup>

<sup>1</sup> Adam Mickiewicz University, Poznań – Poland

A substantial body of research has observed that the grammatical gender values of L1 interact with those of L2 in language processing, a finding referred to as the "gender congruency (GC) effect" (see Sá-Leite et al., 2020, for discussion). GC effects have been explored with respect to the recognition of nouns in isolation (e.g., Paolieri et al., 2020; Sá-Leite et al., 2023), but not in sentence context. However, evidence for cross-linguistic activation of gender at the sentence level is needed to increase the ecological validity of research on GC effects. Furthermore, although it is now established that gender transparency, i.e., the degree to which the gender of a noun can be predicted from its form, plays a pivotal role in gender processing (e.g., Caffarra et al., 2014), it is yet to be understood how it interacts with GC in L2 comprehension. Finally, previous studies present an inconsistent picture of the association between L2 proficiency and the degree of GC effects (L2 proficiency has never been included as a factor in the analyses of any study on GC effects).

To address these issues, we conducted a self-paced reading task with 40 adult L1 Polish L2 German learners, whose L2 German proficiency ranged from A2 to C1 (mean Dialang score = 49.4 out of 75). We created four conditions with eight trigger nouns each by crossing the factors of gender congruency and gender transparency. The conditions were matched for the length, frequency (SUBTLEX-DE), cognate status, and concreteness of the trigger nouns. The trigger nouns were embedded in low-constraint, grammatical sentences that followed the same syntactic structure. Participants read them for comprehension in a linear, non-cumulative procedure with word-by-word segmentation. They also completed a background questionnaire, the Dialang placement test, and a control gender assignment task in L2 German.

We analysed the reading time (RT) of the trigger nouns using generalised linear mixed effects models with congruency (congruent, incongruent) and transparency (transparent, non-transparent) as fixed effects and L2 proficiency as a covariate, including their interactions. The random-effects structure of the model was determined based on the Akaike and Bayesian information criteria. The model returned significant interactions between congruency and transparency ( $\chi^2 = 4.55$ ,  $p = .033$ ) as well as between congruency, transparency, and L2 proficiency ( $\chi^2 = 5.63$ ,  $p = .018$ ), but no main effects. Simple effects analysis revealed shorter RTs for gender-congruent ( $M = 580$ ,  $SE = 40.4$ ) than gender-incongruent ( $M = 658$ ,  $SE = 45.7$ ) nouns, but only if they had a transparent ending ( $estimate = 0.13$ ,  $p = .043$ ). This GC effect decreased with advancing L2 proficiency.

Our study shows that (i) L2 learners activate gender information in L1 when reading nouns embedded in sentences in L2 providing that the nouns carry formal gender cues, and (ii) GC effects in sentence context are negatively modulated by L2 proficiency. We interpret these findings as evidence that the resting level of activation of gender nodes in L2 is higher for transparent

---

\*Speaker

than non-transparent nouns, and that low-proficiency L2 learners access gender through the L1 lexicon.

**Keywords:** grammatical gender, coactivation, language processing



# The distributed practice effect in L2 vocabulary learning: A close replication to look at long(er)-term retention

Amanda Edmonds \*<sup>1</sup>, Shona Whyte<sup>1</sup>, Katerina Palasis<sup>1</sup>, Emilie Gerbier<sup>1</sup>

<sup>1</sup> Université Côte d'Azur – Université Côte d'Azur (UCA) – France

The distributed practice effect refers to the impact of the temporal organization of learning episodes on both short-term learning and retention over time. A large body of research in the field of cognitive psychology has shown that wider spacing of learning episodes generally leads to lower levels of initial learning but better results in long(er)-term retention (Gerbier & Topino, 2015). The way learners and teachers organize learning sessions across time has begun to receive attention from second language (L2) acquisition researchers, who have sought to determine whether the distributed practice effect holds in ecological L2 learning situations. This body of research reports mixed findings concerning the potential benefit of more widely-spaced learning episodes on ultimate retention (e.g., Kim & Webb, 2022), highlighting the importance of exploring findings from laboratory settings in more pedagogically valid learning situations (see Dunlosky et al., 2013).

One potential explanation for inconsistent findings across L2 studies concerns the different temporal intervals selected by researchers. Two different types of temporal intervals are important in distributed practice research: the time between learning sessions (i.e., the intersession interval, or "spacing") and time between the end of learning and final testing (i.e., retention interval). Whereas intersession intervals have ranged from a few seconds (Nakata & Webb, 2016) to almost two months (Bahrack et al., 1993), retention intervals have varied between one day (Moinzadeh et al., 2008) and five years (Bahrack et al., 1993). Such variety, as well as other sources of variability, greatly hamper comparison across studies. One approach that has the potential to aid in clarifying past results involves carrying out replication studies (Plucker et al., 2021). The current project belongs to this line of research.

We conducted a close replication "whereby *only one* major variable is modified each time ... and all others are kept as constant as possible to be better able to single out the kind of influences each has on the dependent variable" (Porte & McManus, 2019, p. 73). The major variable that we modified was the retention interval. In Authors (xxxx), 78 francophone university learners of L2 English completed three 30-minute learning sessions focusing on the meaning, form, and use of 16 English verbs at intersession intervals of 2 or 7 days. All learning and testing sessions were delivered with a dedicated online platform. Lexical development was measured using a modified version of the Vocabulary Knowledge Scale (Paribakht & Wesche, 1993), and the retention interval was 2 weeks. Results provided support for the distributed practice effect, with significantly better retention 15 days after learning for participants who followed the 7-day intersession schedule. In the present close replication, we changed only the retention interval from 2 to 4 weeks, allowing us to examine even longer-term retention. A total of 147 francophone university learners completed this study. Results from this replication will be discussed and compared with previous work on the distributed practice effect in L2 vocabulary learning.

---

\*Speaker

**Keywords:** distributed practice, replication, English, vocabulary learning, technology

# The Role of L1/L2 Transfer in Morpho-Syntax from a Processability Perspective: A Longitudinal Study of Learners of English as an Additional Language (LX) in Lower Secondary Education in Austria

Katharina Egger \* <sup>1</sup>

<sup>1</sup> Leopold Franzens Universität Innsbruck - University of Innsbruck – Austria

My ongoing PhD project focuses on the role and influence of German and Bosnian/Croatian/Serbian as an L1/L2 on the morpho-syntactic acquisition of English as an additional language (LX) in the Austrian educational context. Furthermore, I will investigate the development of L1/L2 transfer effects on student performances from a longitudinal perspective.

The extent of L1/L2 transfer, especially in the areas of morphology and syntax, is still highly debated among SLA researchers (Lenzing & Håkansson, 2022). The theoretical focus of this PhD project concerning language transfer draws on the principles of *Processability Theory* (PT) and the related *Developmentally Moderated Transfer Hypothesis* (DMTH). PT assumes that morpho-syntactic structures can only be produced when learners are developmentally ready to process them (Pienemann, 1998). In turn, the DMTH implies that language transfer is constrained by the architecture of the human language processor. Consequently, L1/L2 features are claimed to be only transferred when the respective processing procedures are processable in the LX (Lenzing & Håkansson, 2022).

Research on language transfer from a DMTH perspective requires studies that focus on different L1s/L2s as well as their influence on the acquisition of the same LX (Lenzing & Håkansson, 2022). So far DMTH-motivated transfer research has mainly focused on languages such as German, Swedish, and English (e.g., Lenzing, 2013; Pienemann et al., 2005) or Turkish and Russian (e.g., Habertzettl, 2005), but has not yet systematically explored the influence of L1/L2 Bosnian/Croatian/Serbian and German on the acquisition of LX English. Furthermore, the Austrian educational context has not been considered in that respect.

In my presentation, I will present the theoretical considerations and the methodological approach of my ongoing PhD project. Based on the theoretical assumption derived from PT and the DMTH, I hypothesise that L1/L2 transfer plays a marginal role in the acquisition of English as an LX, as it is developmentally constrained. I also assume that no bulk transfer of L1/L2 morpho-syntactic structures will be noticeable in the LX English performances and that L1/L2 transfer phenomena will solely be visible as soon as the linguistic structure is processable in the learner's LX English. I further hypothesise that the results of the collected learner speech data will show that the route of acquisition of LX English is not altered by L1/L2 influence. Moreover, the minor role of L1/L2 transfer will not change longitudinally.

---

\*Speaker

To test these hypotheses, I will collect oral speech data of 30 English LX learners (grade 5, approximately 10-11 years old) in Lower Secondary Classes in Austria using communicative tasks. With respect to the longitudinal study design, I plan to collect learner data at three points in time: towards the beginning and end of the school year as well as at the beginning of the following school year. I then establish linguistic profiles on the basis of distributional analyses of the learner data and additionally analyse the learner data in terms of morpho-syntactic L1/L2 transfer and its possible development over time.

**Keywords:** Developmentally Moderated Transfer Hypothesis, longitudinal design, L1/L2 transfer, English as an LX, morphology, syntax, route of acquisition

# What has Automatic Speech Recognition to offer for the study of learners' corpora?

Sarra El Ayari \*<sup>1</sup>, Marzena Watorek<sup>1</sup>

<sup>1</sup> Structures Formelles du Langage – CNRS : UMR7023, Université Paris VIII - Vincennes Saint-Denis  
– France

Providing useful tools for corpus linguistics is an important aspect of Second Language Acquisition. Recent technological advances raise new methodological questions (Benazzo & Watorek 2021). While the results obtained for Automatic Speech Recognition (ASR) in general are very encouraging (Radford & al. 2023), we still need to be able to evaluate their performance in detail on non-standard languages such as the interlanguage of foreign language learners (Selinker 1972). This project is therefore at the crossroads of linguistics and computer science, and aims to carry out a study of ASR on French learners' corpora of different native languages - data which have already been manually transcribed - in order to compare and quantify the results obtained, as well as to finely assess which elements are causing difficulties.

Speech transcription is the first step of the study of any oral corpora and as such it implies theoretical choices. We assume that "learner varieties are not imperfect imitations of a 'real language' - the target language - but systems in their own right, error-free by definition" (Klein & Perdue 1997). As such, two important points to be addressed are comparative fallacy (Bley-Vroman 1993) and proximity fallacy (Perdue 1993), which are bound to cause difficulties for automatic tools.

We used *Whisper* ASR tool developed by Open-AI(1). The speech-to-text functionalities provide the transcription of the audio files, with audio alignment. This study is conducted on the French data of the ESF corpus (Perdue 1993), which focused on the spontaneous SLA of forty adult migrants living in Europe. This corpus has already been transcribed and studied (Perdue 1993, Klein & Perdue 1997), fundamental point for the comparison of the results between the two transcriptions: automatic *vs* manual.

The comparison process is allowed by the platform Sarramanka (El Ayari & Watorek 2021, El Ayari 2022) created to transcribe, annotate and explore oral and written corpora. The interface provides a compared view of the two versions of one audio file's transcription. An annotation schema of the errors has been implemented in order to evaluate the results initially, and furthermore to precisely characterize the nature of the problems encountered by *Whisper*.

Our preliminary conclusions are that, even if ASR is not sufficiently trained on SLA data in general, using the results as a basis for manual correction could still be an interesting feature to implement in our Sarramanka platform, especially as it provides also audio alignment. We would then need to conduct a study to measure with accuracy how time consuming this process could be, in order to be able to fully evaluate the benefit of such a functionality. Our guess is that the results would be correlated with learner variety, and therefore could provide indications about the SLA stages of the learner. This study is the starting point of a wider project, where we would like to train ASR tools on French learners' corpora of different native languages and acquisition stages.

---

\*Speaker

(1) Link to Whisper on GitHub: <https://github.com/openai/whisper>

**Keywords:** transcription, automatic speech recognition, second language acquisition, learners corpora

# Assessing Students' Spanish Oral Proficiency Development after One Semester Abroad in Spain: A group and individual analysis

Sara Fernandez Cuenca <sup>\*† 1</sup>, Minerva Rojas <sup>\*</sup>

<sup>2,3</sup>, Anna Attal <sup>\*</sup>

1

<sup>1</sup> Wake Forest University – United States

<sup>2</sup> Université Côte d'Azur – Université Côte d'Azur (UCA) – France

<sup>3</sup> Bases, Corpus, Langage (UMR 7320 - UCA / CNRS) – Université Nice Sophia Antipolis (1965 - 2019), Centre National de la Recherche Scientifique, Université Côte d'Azur – France

The rapidly growing body of study abroad (SA) research suggests that short and long-term immersion in a foreign language is beneficial for second language (L2) development, particularly oral proficiency skills (Marijuan & Sanz, 2018; Serrano et al., 2012). This study aims to contribute to SA research by investigating the longitudinal development of Spanish L2 learners' fluency, accuracy, and complexity after a semester studying abroad in Spain, and how initial proficiency and degree of immersion in the L2 may modulate changes over time. More recently, SA research presented at the Residence Abroad and Language Learning: Where Are We Now? Conference (October, 2023) points to individual differences as the determining factor that can explain the variable outcomes different L2 learners exhibit as the result of studying abroad. With this in mind, our study expands on previous SA research by investigating how initial Spanish proficiency and total number of exposure in the L2 while abroad may play a role in L2 oral proficiency outcomes at an individual and group level.

A total of 15 Spanish L2 learners took part in a 4-month long semester abroad program in Spain. Learners' speech was elicited using the ACTFL Oral Proficiency Interview (OPI) method, which was conducted at the beginning and at the end of the semester. Proficiency was assessed with a standardized elicited imitation task (EIT), and fluency, accuracy, and complexity data was coded using the audio annotation software EXMARaLDA (Schmidt, 2012). The transcription protocol we used was based on MacWhinney (2000)'s CHAT conventions. Fluency was measured through computing the average duration of filled and silent pauses. To track changes in accuracy, the number of morphological, lexical, and syntactical errors per utterance was calculated. Participants were asked to self-report how many hours of L2 exposure they had had in the L2 using a biweekly questionnaire, and this was how the degree of L2 immersion was determined (Issa et al., 2020).

Preliminary results from 5 participants revealed that the overall fluency of the participants

---

\*Speaker

†Corresponding author: fernans@wfu.edu

increased, as seen through a decrease in the average duration of pauses (both filled and silent) from session 1 to session 2, however this difference did not reach significance ( $t = 1.56$ ,  $p = 0.19$ ). As a group, participants' number of errors (accuracy) also decreased from session 1 to session 2, but this differed greatly at the individual level with one participant showing the opposite pattern. Neither initial proficiency nor degree of immersion in the L2 seemed to predict oral development over time at the group or individual level, despite previous research pointing to this initial proficiency as a strong predictor within a SA context (Zalbidea et al., 2020). Complexity analyses are in progress and are being measured through the mean length of utterance (MLU) and the ratio of simple and complex utterances.

**Keywords:** learner corpus, spanish L2, study abroad, complexity, accuracy, fluency



# The Effects of Textual Enhancement and Individual Differences on Young Learners' Vocabulary Acquisition Through Captioned Cartoons

Rebeca Finger-Bou <sup>\*† 1</sup>, Carmen Muñoz <sup>1</sup>

<sup>1</sup> University of Barcelona – Spain

Research suggests that exposure to captioned audiovisual material (subtitles in the L2) can facilitate L2 vocabulary acquisition, as learners can readily recognize written words (Authors, 2022). Some studies have effectively directed learners' attention by enhancing specific parts within the captions, thereby increasing their saliency (Puimège et al., 2021). However, the potential advantage of textually enhanced captions over regular captions remains unclear (Montero Perez, 2022). Moreover, the impact of textual enhancement on the learning experiences of young learners is largely unknown, given the scarcity of studies on primary school L2 learners. It has been suggested that the processing of captions and young learners' capacity to learn from them might be influenced by both first language (L1) and L2 skills and their language learning aptitude (Author, 2023; Sparks & Dale, 2023). The present study aimed, first, to analyse the effects of textual enhancement on vocabulary acquisition by 39 L1-Spanish/Catalan primary school learners of English through exposure to 11 episodes of the British animated children's television series *Charlie & Lola*. A second aim of the study was to examine the extent to which individual differences, such as previous vocabulary knowledge, L1 and L2 literacy skills, namely text segmentation and reading efficacy, and MLAT-EC Part 4 (only for 5th graders), were associated to vocabulary gains. Four groups from 3rd and 5th grade were formed: one from 3rd grade and one from 5th grade were provided with regular captions, whereas the other two groups watched the episodes with enhanced captions, where target words were bolded and in yellow. Vocabulary gains were assessed through pre- and post-tests that tapped into form recall and meaning recognition. Results show higher gains in the enhanced captions group only for meaning recognition in 3rd grade and for both form recall and meaning recognition in 5th grade, which might indicate that textual enhancement had a distinctive impact on learning, especially for older children, whose previous vocabulary knowledge was significantly higher. As for the analysis of individual differences, results suggested that 3rd graders relied more on their L2 abilities for form recall and on their L1 abilities for meaning recognition. On the other hand, 5th graders relied on their L2 abilities as well as their L1 reading skills for both form recall and meaning recognition. Associative memory, as measured by the MLAT-EC Part 4 subtest, also showed a significant effect on both levels of knowledge. Several hierarchical multiple regression analyses showed that L1 skills accounted for the highest amount of variance in subjects' gains for meaning recognition for both grades. Overall, these findings endorse the potential advantage of textually enhanced captions also for young beginner learners and provide supportive evidence for the importance of the role played by L1 and L2 literacy skills in L2 learning.

---

\*Speaker

†Corresponding author: rfingerb@ub.edu

**Keywords:** multimodal input, primary school learners, vocabulary acquisition, textual enhancement, captions, individual differences

# The effect of machine translation access on L2 writers' cognitive processes across proficiency levels

Margot Fonteyne <sup>\*† 1</sup>, Maribel Montero Perez <sup>1</sup>, Joke Daems <sup>1</sup>, Lieve Macken <sup>1</sup>

<sup>1</sup> Ghent University – Belgium

The linguistic knowledge of L2 learners, especially beginners, is often limited and lacks automatization (De Larios et al., 1999). This increases learners' cognitive load during lower-order writing processes (vocabulary retrieval, grammatical encoding, and local revision), leaving them with fewer resources to allocate to other, higher-order processes (generating, organising, and revising content). Emerging technologies, however, may help writers with cognitive offloading (Overstreet, 2022). For example, survey results suggest that machine translation (MT) can help L2 learners to reduce their cognitive load during writing (Ahn & Chung, 2020).

In this study, we investigate how MT access affects both higher- and lower-order cognitive processes of L2 writers. Specifically, we compare MT-assisted writing processes to those where L2 writers had access to an online bilingual dictionary. Twenty-seven L1 Dutch learners of Swedish were assigned four writing tasks, with two tasks completed in each condition (MT or dictionary). While writing, participants' keystrokes, screen activity, and eye movements were recorded. We also gathered stimulated recall data for each fourth task and measured participants' proficiency through a placement test. A series of linear mixed effects and simple regression analyses will be conducted on various measures extracted from these data sources to compare participants' lower- and higher-order processes across the two conditions.

We hypothesise that participants' lower-order processes will be accelerated in the MT condition, resulting in more engagement in higher-order processes. This is expected to be apparent in participants' writing speed, pausing behaviour (frequency and duration of pauses between different textual units), revision behaviour (frequency, length, and type of revisions), eye movements (frequency and length of saccades), and stimulated recall comments (percentage of comments related to different processes). We anticipate these effects to be most prominent among the least proficient.

Preliminary analyses of speed and pausing measures extracted from the keystroke logs show that the participants in the MT condition produced more words per minute, paused for a shorter time within words, and took fewer long pauses between words. No interactions were found between the tool participants had access to and their proficiency level. These effects suggest that MT does speed up L2 learners' lower-level processes during writing. However, this acceleration is not necessarily more pronounced for less proficient writers. We will triangulate these first results with the results of the analyses based on the revision, eye-tracking, and stimulated recall data. Additionally, we will discuss our findings in relation to previous research on the impact of MT on L2 writing and reflect on their pedagogical implications.

---

\*Speaker

†Corresponding author: [margot.fonteyne@ugent.be](mailto:margot.fonteyne@ugent.be)

Ahn, S., & Chung, E. S. (2020). Students' perceptions of the use of online machine translation in L2 writing. *Multimedia-Assisted Language Learning*, 23(2), 10–35. <https://doi.org/10.15702/mall.2020.23.2.10>

De Larios, J. R., Murphy, L., & Manchon, R. (1999). The use of restructuring strategies in EFL writing: A study of Spanish learners of English as a foreign language. *Journal of Second Language Writing*, 8(1), 13–44. [https://doi.org/10.1016/S1060-3743\(99\)80111-8](https://doi.org/10.1016/S1060-3743(99)80111-8)

Overstreet, M. (2022). Writing as extended mind: Recentering cognition, rethinking tool use. *Computers and Composition*, 63, 102700. <https://doi.org/10.1016/j.compcom.2022.102700>

**Keywords:** writing process, writing tools, machine translation, bilingual dictionaries, eye tracking, keystroke logging, stimulated recall

# You're better than you think: conceptual changes of caused motion events in Chinese learners of L2 English: a continuous flash suppression study.

Xingyi Fu <sup>\*† 1</sup>, Leah Roberts <sup>1</sup>, Norbert Vanek <sup>2</sup>

<sup>1</sup> University of York – United Kingdom

<sup>2</sup> University of Auckland – New Zealand

Linguistic relativity (Whorf, 1956) predicts that learning a second language will alter the learners' underlying cognition, but only limited evidence has been found to support this claim (Flecken et al. 2014), especially in the domain of motion. This might be due to the difficulty in disentangling the relationship between language and cognition. For instance, it is difficult to prevent participants from using language silently during non-verbal experiments. The present study attempts to tackle this issue by introducing a novel paradigm, breaking continuous flash suppression (b-CFS, Jiang et al., 2007; Vanek & Fu, 2023) which detects low-level visual processing, to examine the conceptual changes in caused motion events experienced by Chinese L2 learners of English. The major cross-linguistic difference between Chinese (equipollently-framed) and English (satellite-framed) is the semantic features used to describe Manner and Path components in motion events. For example, in '*A man is carrying a suitcase into a room*', English uses the main verb (*carry*) and its subordinate satellite element (e.g., a verb particle: *into*) to express Manner and Path respectively (Talmy, 2000), whereas Chinese uses two verbs to illustrate those two semantic components (a 'serial verb construction', Chen & Guo, 2009). Therefore, English speakers were expected to use different amounts of time when processing Manner and Path salient stimuli in the b-CFS experiment, in contrast, Chinese speakers should process them similarly (Vanek and Fu, 2023).

Native Chinese (N = 55) and English speakers (N = 52) were tested, together with 52 Chinese L2 learners. The experimental settings were identical to participants across three language groups except that the L2 learners were assigned into two sub-groups (monolingual vs. bilingual context) randomly to examine the effect of immediate language use on processing motion events in the b-CFS. Each participant firstly viewed a prime showing a sequential photo of a motion scene, and then detected a target stimulus which was visually suppressed by a flickering mandarin mask (see detailed procedure in Fu et al., 2023). The target stimuli were either the same or partially different compared with the prime photo. The time they spent on the detection was recorded and analysed. The results supported the predictions for the natives: English speakers were significantly faster to break through the suppression and detect the target stimuli with Manner alteration (*manner mismatch* condition) compared to those with Path alteration (*path mismatch* condition), whereas the native Chinese speakers showed no such difference. The L2 learners did not significantly resemble the processing patterns of shorter time in the *manner mismatch* compared to the *path mismatch* condition, but they were significantly different from Chinese but not English speakers regardless the given linguistic contexts. More importantly, the L2 learners were significantly different from both native Chinese and English speakers in

---

\*Speaker

†Corresponding author: xingyi.fu@york.ac.uk

the *manner mismatch* condition. In general, the results demonstrate cross-linguistic effects on event cognition, specifically, that the L2 learners converged concepts of motion event featured in both English and Chinese, and they were able to shift from L1 to L2 (Jarvis & Pavlenko, 2010) at the intermediate-level.

**Keywords:** breaking continuous flash suppression, linguistic relativity, motion events, Chinese learners of L2 English, conceptual changes

# Implicit learning with(out) UG: An extraction constraint, case marking, and displacement

Junya Fukuta \*<sup>1</sup>, Takayuki Kimura<sup>2</sup>, John Matthews<sup>1</sup>, John Williams<sup>3</sup>,  
Boping Yuan<sup>3</sup>, Yuyan Xue<sup>3</sup>, Shigenori Wakabayashi<sup>1</sup>

<sup>1</sup> Chuo University – Japan

<sup>2</sup> Utsunomiya University [Utsunomiya] – Japan

<sup>3</sup> University of Cambridge [UK] – United Kingdom

This study investigates the learnability of constraints on a syntactic operation using a semi-artificial language, exerting control over language experience (e.g., input distribution and frequency) and manipulating learners' first language. We created an artificial language 'Janglish,' an amalgamation of Japanese and English with inflectional morphology modeled on Japanese and lexis taken from English. Grammatical morphemes, including case markers, postpositions, and classifiers, function as they do in Japanese.

In numeral expressions in Janglish, a classifier is placed after a numeral, forming a Classifier Phrase (ClassP) (e.g., ((2NUM *keku*CL)ClassP *book*)NP "two volumes of book(s)"). A noun phrase can be extracted from ClassP, resulting in numeral quantifier floating (NQF). However, the extraction is not freely applied but is constrained by a general principle of Universal Grammar (UG): Extraction of a noun phrase (out of ClassP) is permitted if it is in a subject/object (i.e., argument) position, but not if it is present within a locative phrase (i.e., adjunct).

Native speakers of English with no knowledge of Japanese were presented with Janglish input and then tested on their ability to judge novel strings for acceptability. The experiment tested participants in two stages: a Learning Stage (LS) and a Testing Stage (TS). In the LS, participants were taught the mappings of classifiers and their referents and the functions of the case markers. Then, they practiced the mappings through computer-based activities. In the TS, participants judged the acceptability of novel sentences with "grammatical" and "ungrammatical" word orders not seen in the LS. The rules for word order were not explicitly taught in LS, and all word orders that appeared in TS were different from the input sentences received in LS.

The goal was to uncover the knowledge participants draw on to make grammaticality judgments following this initial exposure. As the underlying constraint on NQF was not taught explicitly, and no ill-formed sentences appeared in the LS, participants received no input/instruction on (un)grammaticality; if learners with no experience of Japanese distinguished grammatical from ungrammatical word orders in the TS, they provide evidence for knowledge of implicit innate principles of grammar in second language acquisition.

Results of a pilot study with English native speakers showed that participants were: (i) insensitive to differences in the constraints on extraction from ClassP, (ii) sensitive to the differences among case markers and postpositions, and (iii) sensitive to the number and distance of movement operations.

Based on these data, we infer that UG may not constrain learner grammars and that FULL

---

\*Speaker

ACCESS to UG from the initial point is not supported. Moreover, L2 learners may not rely exclusively on the L1 system. They are likely to utilize a linguistic (but not L1) grammatical system in the acceptability judgment. Lastly, the number of movement operations may provide participants with intuition for their acceptability judgments. This again suggests that L2 learners utilize their "linguistic knowledge" (unconscious knowledge) for acceptability judgment.

**Keywords:** semi, artificial language learning, implicit learning, universal grammar, learnability constraints



# Why some aspectual structures are more difficult to acquire than others: Insights from complex predicates in L2 Romance

Birgit Füreder \* <sup>1</sup>

<sup>1</sup> University of Salzburg – Austria

Aspectual distinctions have proven to be particularly challenging in non-native language acquisition (e.g. Bardovi-Harlig 2000, Cadierno 2000, Salaberry/Shirai 2002, McManus 2015, Vallerossa et al. 2023) – even more so with increasing (typological) differences between the source and target language(s). On the one hand, this is due to the variety of aspectual shades typically expressed in the individual languages, on the other hand, to the different linguistic means used to express these shades: While Romance languages, for instance, show fairly differentiated aspectual systems predominantly expressed by verbal morphology or complex predicates, the German system appears rather limited, relying mostly on lexical or derivational means. In order to investigate which aspectual structures are most problematic for German-speaking learners, a self-paced reading study was administered to a L1 German experimental group (proficiency level in Spanish B2-C2) and a L1 Spanish control group (see Marsden et al. 2018 for a methodological review). The stimulus material consisted of sentences in Spanish containing complex verbal structures from 12 aspectual categories (80 items in total, 38 critical ones), which were presented on a computer screen using the moving-window technique (Just et al. 1982), immediately followed by an acceptability rating and a comprehension task.

First results (n=17) suggest that gerundival structures (except for the progressive) are most challenging – for learners and natives, interestingly – surfacing in highly divergent acceptability ratings and reaction times. Apart from that, some infinitival structures expressing habitual, inchoative and terminative values show varied results among learners and natives. Drawing on reference grammars, corpus research and retrospective interviews with the participants, some of these divergences may be explained by (i) frequency of use, (ii) regional variation, (iii) combinatorial preferences and (iv) instructional input. All of these factors coincide in the Spanish complex verbal structures *ir/venir/andar* + gerund (literally ‘go/come/walk around’ + gerund), expressing cumulative and distributive values (often with modal shades), making them especially vulnerable in non-native language acquisition. These results suggest that particular instructional effort is required for such structures in order to help learners establish adequate form-meaning pairs.

References:

Bardovi-Harlig, Kathleen (2000): *Tense and aspect in second language acquisition. Form, meaning, and use*. Malden/Massachusetts: Blackwell.

Cadierno, Teresa (2000): "The acquisition of Spanish aspectual distinctions by Danish language learners". In: *Spanish Applied Linguistics* 4, 1-53.

---

\*Speaker

Just, Marcel/Carpenter, Patricia/Woolley, Jacqueline (1982): "Paradigms and processes in reading comprehension". In: *Journal of Experimental Psychology* 111 (2), 228-238.

Marsden, Emma/Thompson, Sophie/Plonsky, Luke (2018): "A methodological synthesis of self-paced reading in second language research". In: *Applied Psycholinguistics* 39, 861-904.

McManus, Kevin (2015): "L1/L2 differences in the acquisition of form-meaning pairings: A comparison of English and German learners of French". In: *The Canadian Modern Language Review* 71 (2), 155-181.

Salaberry, Rafael/Shirai, Yasuhiro (eds.) (2002): *The L2 acquisition of tense-aspect morphology*. Amsterdam/Philadelphia: Benjamins.

Vallerossa, Francesco/Gudmundson, Anna/Bergström, Anna/Bardel, Camilla (2023): "Learning aspect in Italian as additional language: The role of second languages". In: *International Review of Applied Linguistics in Language Teaching* 61 (3), 1139-1171.

**Keywords:** Aspectual distinctions, complex predicates, Romance languages

# The Predictive Validity of IELTS Scores: A Meta-Analysis

Tomlin Gagen <sup>1</sup>, Farahnaz Faez <sup>\*† 1</sup>

<sup>1</sup> Western University – Canada

Increases in international student enrolment has led to an increased need to examine the language tests used to ensure they have the necessary English skills to succeed in university. One of the most popular of those tests is the International English Language Testing System, or IELTS. Thousands of education institutions worldwide rely on IELTS scores as criteria for accepting international students whose first language is not English. The question then arises: is the IELTS test capable of predicting academic success?

Individual studies have found varying degrees of strength of correlations and conflicting results between IELTS scores and academic success. These conflicting results were examined through a meta-analysis, while also investigating multiple moderating variables: research funding bias, individual skill scores, level, field, & country of study, presence of additional English courses, and GPA timepoint. Results from 18 studies show an approaching-small effect size of  $r = .227$  for the relationship between IELTS scores and post-secondary GPA. The majority of macro skills (listening, writing, and speaking) do not reach the small effect size, however reading approaches it with an effect size of  $r = .215$ .

Additional examination is made into if the relationships are stronger for: different majors, different levels of study (graduate or undergraduate), different countries, presence of extra English courses, and results at different times. Results show that the IELTS test has a relatively small predictive effect for GPA in postsecondary programs. No evidence of bias was found in publishing source, and overall scores were much more predictive than any one individual skill such as listening or writing, although reading shows a slight relationship. Most moderator analyses were inconclusive owing to the small amount of studies, but potential differences from individual studies are examined and discussed.

**Keywords:** IELTS, Meta, Analysis, Predictive Validity, GPA

---

\*Speaker

†Corresponding author: ffaez@uwo.ca

# L1 Attrition in instructed settings: L1 Spanish-L2 English bilinguals' relative clause attachment preferences

Elena García-Guerrero \* <sup>1</sup>, Cristóbal Lozano <sup>1</sup>

<sup>1</sup> Universidad de Granada = University of Granada – Spain

Traditional second language acquisition (SLA) research has focused on bilinguals' second language (L2) as modulated by their first language (L1). However, more recent studies challenge this unidirectional influence, suggesting that the L1 may also undergo adjustments due to extensive L2 use and exposure (Chamorro et al., 2016; Dussias & Sagarra, 2007), a phenomenon termed L1 attrition. This evidence mostly comes from studies on long-immersed bilinguals in naturalistic settings, but there is scarce data on the potential influence of other L2 exposure types like instructed settings (i.e., classroom settings). This study aims to bridge this gap by investigating potential attrition in bilinguals living in an L1 environment but extensively exposed to formal L2 instruction.

The present investigation explores L1 attrition related to relative clause attachment (RCA) ambiguities, given the crosslinguistic variation observed in monolingual speakers of different L1s. For instance, the relative pronoun *who* in (1) may refer to either *el alumno*, i.e. high attachment (HA), or *la científica*, i.e., low attachment (LA). Native Spanish shows a tendency for HA, whereas native English speakers seem to favour LA (Cuetos & Mitchell, 1988; Dussias, 2003; Fernández, 2002; Jegerski et al., 2016).

(1) *Mira al alumnoi de la científicaj **quei/j** lee un libro atentamente*

*Look at the studenti of the scientistj **whoi/j** reads a book carefully*

This study explores: (1) the RCA disambiguation preferences of L1 Spanish-L2 English bilinguals in their L1 as modulated by extensive L2 instructed exposure and (2) whether attrition occurs solely in processing or also in offline comprehension. To do so, three groups were tested: advanced L1 Spanish-L2 English bilingual students of English Studies in Spain (N = 44), Spanish functional monolinguals (N = 50), and English functional monolinguals (N = 49). Data were collected from a picture selection task, an auditory sentence-picture verification task, and a visual-world eye-tracking experiment to address both offline comprehension and online processing preferences. All participants were tested in their L1, and target sentences were manipulated to create three conditions (i.e., HA-bias, LA-bias and ambiguous) by manipulating the relative pronoun gender as in (2).

(2) *Mira al alumnoi de la científicaj **el cuali / la cualj / quei/j** lee un libro atentamente*

*Look at the studenti (masc) of the scientistj (fem) **whoi/j** reads a book carefully*

Based on results from the picture selection task and participants' reaction times, L2 instructed

---

\*Speaker

exposure seems to influence bilinguals' comprehension and processing of RCA in their L1 Spanish. While Spanish monolinguals show a straightforward HA preference, bilinguals show significantly more optionality, which implies higher acceptance of the L1 dispreferred strategy (i.e., low attachment). Additionally, with ambiguous sentences, bilinguals are significantly slower than Spanish monolinguals in the final interpretation of the sentence. These findings shed light on L1 attrition in an overlooked bilingual population, contributing to our understanding of this phenomenon.

References: [https://docs.google.com/document/d/1txGIZW\\_GQIS6n6-ThbTCj1v6VaJrd74Ajn30U3HahxY/e](https://docs.google.com/document/d/1txGIZW_GQIS6n6-ThbTCj1v6VaJrd74Ajn30U3HahxY/e)

**Keywords:** L1 attrition, relative clauses, instructed setting, L1 Spanish, L2 English

# What factors hinder or favor intelligibility in English Second Language Pronunciation? A scoping review of 25 years of research

Marie Garnier <sup>\*† 1</sup>, Saandia Ali <sup>\*</sup>

<sup>2</sup>, Linda Terrier <sup>\*</sup>

1

<sup>1</sup> Centre for Anglophone Studies – Université Toulouse Jean Jaurès – France

<sup>2</sup> Cognition, Langues, Langage, Ergonomie – Université Toulouse - Jean Jaurès – France

The domain of L2 English pronunciation was marked by a major turning point almost three decades ago when Munro and Derwing published their seminal paper "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners" (1995). They reported the results of an experiment showing that foreign-accented English speech was not necessarily judged unintelligible and could be, on the contrary, understood perfectly by L1 English listeners. The researchers opened new research perspectives on the constructs of foreign accented speech and intelligibility, as well as comprehensibility. They suggested an operational definition of intelligibility in L2 pronunciation as "the extent to which a speaker's message is actually understood by a listener (...)." (Munro et Derwing 1995 : 76-77). More recently, Derwing and Munro called for "more research on factors that contribute to intelligibility" (2015 : 198). Our current research aims at answering this question by pinpointing the factors that have been identified as hindering and favoring intelligibility in the research on L2 English pronunciation.

To that aim, we have produced a scoping review of 25 years of research in this field. The corpus of this review includes 485 articles published between 1995 and 2020, documenting experimental research on L2 English pronunciation. The articles come from 35 major journals in the fields of phonology/phonetics (PHON), second language acquisition (SLA), and second language learning and teaching (SLLT). The sources selected are international journals published online and in English. The corpus of articles was annotated manually, taking into account three main items of information: the participants in the experiments, the phonological features examined in the studies and the importance of the construct of intelligibility. The corpus is balanced, as 56% of the articles focussing on L2 pronunciation were published in the domains of L2 research (SLA and SLLT) while 44% were published in journals of phonetics and phonology.

As our poster presentation for the 33rd EuroSLA conference will emphasize, the scoping review shows a rise in the proportion of articles taking into account the intelligibility construct between 1995 and 2020, with a bump in interest for research focused on this construct in the mid-2000s. More articles taking intelligibility into account are published in the PHON domain (71%) than in the SLA/SLLT domain (58%). We measured the importance of the construct of intelligibility in research by running a keyword search for the item "intelligibility" in each article, as well as looking at its context of use. Articles were classified in 4 categories on the basis of this quantitative and qualitative data: A. Absent, B. Marginal, C. Important, D. Central. By looking at

---

\*Speaker

†Corresponding author: marie.garnier@univ-tlse2.fr

the articles in the Marginal, Important and Central categories, we identified which factors were the focus of research in the literature, as well as whether these factors were associated with an increase or a decrease in intelligibility. This poster will present more detailed results following a full exploration of the corpus.

**Keywords:** second language acquisition (SLA), second language learning and teaching (SLLT), phonetics, phonology, second language pronunciation, intelligibility

# Auditory processing ability predicts non-native word learning: Evidence from online (eye-tracking) and offline measures

Yuxin Ge \*<sup>1,2</sup>, Magdalena Kachlicka<sup>3</sup>, Kazuya Saito<sup>4</sup>, Patrick Rebuschat<sup>1,5</sup>, Padraic Monaghan<sup>1</sup>

<sup>1</sup> Lancaster University – United Kingdom

<sup>2</sup> NOVA University Lisbon – Portugal

<sup>3</sup> Birkbeck University of London – United Kingdom

<sup>4</sup> University College London – United Kingdom

<sup>5</sup> University of Tübingen – Germany

Infants and children can rapidly pick up novel words from the surrounding linguistic environment. However, this task is considerably more challenging for adults learning a non-native language (L2) because they may encounter sounds that do not exist in the inventory of their native language(s). Difficulty perceiving and representing L2 sounds has been found to interfere with word learning (Escudero et al., 2022; Ge et al., under review). However, there is also evidence that this domain-specific perceptual difficulty may be modulated by domain-general auditory processing abilities (i.e., the ability to encode acoustic features of sounds) (Saito et al., 2020). In this study, we investigate whether and how individual differences in auditory processing ability interact with L2 word learning. To better capture the word learning process and outcome, we employed an online eye-tracking measure in addition to an (offline) accuracy measure.

We trained 53 English-native speakers with 16 Mandarin tonal pseudowords via a cross-situational word learning task (CSWL) (Monaghan et al., 2015, 2019). Each CSWL trial contained ambiguous word-picture mappings: participants were presented with two referent pictures and one pseudoword. Their task was to decide, as quickly and accurately as possible, which picture the spoken word referred to, and no feedback was provided. Throughout the task, we carefully selected the two pictures, target and distractor, to manipulate the degree of phonological similarity of the pseudowords associated with them. In some trials, the two pictures were mapped to two minimal pair words (e.g., *pa1mi1* vs *pa4mi1*, numbers referring to Mandarin tone 1 and tone 4), and in others, non-minimal pairs (e.g., *pa1mi1* vs *lu4fa1*). We recorded participants' online eye-gaze fixation while they completed the CSWL task, using the built-in eye-tracking function in the Gorilla platform (<https://gorilla.sc/>). After the CSWL task, we tested participants' auditory processing ability by means of a pitch discrimination task (Wong & Perrachione, 2007) and a melody reproduction task (Saito & Tierney, 2022).

We found that both the online and offline measures revealed clear learning effects. In the final block of the CSWL task, participants were significantly more likely to fixate on the correct referent and to select the correct target (compared to chance level). The two measures of learning aligned and thus provide strong evidence of successful learning in L2 contexts where words contain non-native phonological features. Analyses of the auditory processing measures showed that melody reproduction ability predicted participants' overall percentage fixation at target ( $Z = 2.310$ ,  $p = .021$ ) and picture selection accuracy ( $Z = 2.138$ ,  $p = .032$ ) at the end of CSWL,

---

\*Speaker



whereas pitch discrimination ability predicted participants' fixation at target in non-minimal pair trials ( $Z = -2.230$ ,  $p = .026$ ).

Our findings contribute to the understanding of individual differences in L2 word learning. More precise processing of domain-general acoustic information was associated with better learning of words with unfamiliar sounds. Our study also has methodological implications for word-learning research in terms of the measurement of learning performance. The online and offline measures provided highly consistent results, suggesting that web-based eye-tracking techniques can be used reliably in picture-word mapping paradigms.

**Keywords:** Implicit learning, statistical learning, cross, situational word learning, non, native phonology, lexical tone, individual differences, eye, tracking

# Reference in children with Greek as a heritage language – When production and comprehension skills draw a different picture

Angelika Golegos \* <sup>1</sup>, Theodoros Marinis <sup>1</sup>

<sup>1</sup> University of Konstanz – Germany

Research on Greek as a heritage language (HL) is a relatively new field of interest. To date, few studies have investigated children growing up in Germany with Greek as a HL and most of them addressed the development of narrative skills in terms of reference production not taking also into account reference comprehension within the same population. Hence, little is known if a child who shows good narrative skills will also perform well in comprehension. Our study aims to fill this gap and to shed light on the question of whether the two components show similarities or differences.

For reference comprehension, research with monolingual Greek children showed that 6-to-7-year-old children prefer a subject referent when encountering a null pronoun, similar to adults, whereas there is no clear object preference when encountering an overt pronoun, in contrast to adults who have an object preference for overt pronouns (Papadopoulou et al., 2015). Research with Greek as a HL in children growing up in Germany is limited to reference production, showing that children often prefer nominal phrases over null pronouns for subject reference, a pattern that is not found to the same extent in monolingual peers (Torregrossa et al., 2021; Andreou et al., 2015). To date, no studies have investigated pronoun production as well as comprehension within the same participants.

The present study investigates 6-to-9-year-old children with Greek as a HL and German as the majority language living in Germany as well as 6-to-8-year-old monolingually raised Greek children living in Greece. The data collection is ongoing. In a narrative production task, children listen to a model story while looking at accompanying images. Afterwards, children have to re-tell the story. In an online picture selection task, participants listen to sentences that contain two competing characters, followed by an ambiguous third-person singular masculine pronoun (*pro*, *aftos*), see Example (1).

1)  $\pi\iota\lambda$   
 $\tau\alpha\upsilon(j)$   
 $\theta\epsilon\tau\alpha\lambda\epsilon\iota\ \alpha\ \zeta\omega\gamma\rho\alpha\phi\sigma\epsilon\iota$   
 $\tau\alpha\upsilon$   
 $\tau\alpha\upsilon\eta\alpha$   
 $\tau\alpha\upsilon\rho(k)\ \mu\epsilon\ \alpha\ \mu\iota\rho\ \pi\iota\lambda.$   $\lambda\lambda\ \text{pro}(j)/\ \alpha\nu$   
 $\tau\alpha\upsilon(k)\ \epsilon\alpha\iota\ \pi\lambda\ \nu\rho\alpha\sigma\mu.$

The pilot(j) wants to draw the vet(k) with a small brush. But pro(j)/ he(k) is very tired.

*‘The pilot wants to draw the vet with a small brush. But he is very tired.’*

---

\*Speaker

While listening to the sentences, participants see three images on the screen, representing the competing subject and object as well as a distractor. After the end of the sentences, participants are asked a comprehension question that forces them to interpret the pronoun towards one of the two characters. Responses are given by clicking on the image of the respective character. Preliminary results show a clear discrepancy between production and comprehension skills in the HL children. While most of the HL children lack production skills during testing, leading to no narrative productions in Greek, their comprehension skills do not differ from those of their monolingual Greek peers. The results indicate that when investigating reference skills of HL children drawing conclusions based on either production or comprehension skills might be misleading. A combination of both measurements is needed to understand reference skills in HL children.

**Keywords:** child language acquisition, heritage language Greek, reference, comprehension, production

# The effect of etymological relatedness and perceived semantic relatedness of multiple meanings of words on L2 learners' vocabulary knowledge.

Beatriz González-Fernández \* <sup>1</sup>

<sup>1</sup> The University of Sheffield – United Kingdom

Most words in a language have multiple meanings (Hoshino & Shimizu, 2018). Lack of knowledge of the various meanings that a word can convey hinders language comprehension and use (Floyd & Goldberg, 2021; Logan & Kieffer, 2021), making it a crucial type of word knowledge for successful communication (Nation, 2020; Schmitt & Schmitt, 2020; Stahl & Nagy, 2006). Research suggests that learning the different meanings of words is one of the main challenges of L2 lexical development (Crossley et al., 2010; González-Fernández & Schmitt, 2020). However, few studies have examined which factors influence the acquisition of the various meanings of words in an L2. L1 research points to the etymological distinction of word meanings between polysemous (i.e., related meanings with the same origin) or homonymous (i.e., unrelated meanings with different origins) as a key factor determining their acquisition (Klepousniotou et al., 2012; Maciejewski et al. 2020). Yet, recent L2 research suggests that this etymological distinction might not be an influencing factor in L2-meaning learning (González-Fernández & Webb, under review). It is possible that learners' own perceptions of semantic-relatedness among the meanings of words is a better indicator of knowledge and acquisition of L2 word meaning. However, this hypothesis has not yet been examined.

The present study addresses this gap by explicitly comparing the effect of etymological relatedness (historical distinction between polysemous or homonymous) and perceived semantic relatedness (meaning similarity) on L2 learners' vocabulary knowledge. 75 adult L2-English learners with varying proficiencies participated in this study. One-hundred English multi-meaning words (50 homonyms (e.g., toast), 50 polysemes (e.g., chair)) were identified. 2-3 distinct meanings for each word were targeted, including a spread across higher and lower degrees of semantic relatedness. Target words' meanings were classified as polysemous or homonymous following the etymological information in the Oxford English Dictionary. Participants' perceptions of the semantic relatedness among the various meanings of the target words were captured using a self-report questionnaire, where learners were presented with each target word and its most frequent meaning. The word's additional meanings were then introduced, and learners had to rate how related they believed each of the additional meanings of the target word was to its main meaning, on a scale from 1 (completely unrelated) to 5 (completely related). Finally, learners' general receptive vocabulary knowledge was assessed via the LexTALE test (Lemhöfer & Broersma 2012).

The results from descriptives and regression analyses indicate that learners' overall perceptions of the semantic relationship between a word's various meanings tend to deviate from their etymological relatedness. In addition, the perceived degree of semantic relatedness was found to have a greater effect on learners' vocabulary knowledge than the etymological distinction of polysemy and homonymy, although this was moderated by L2 proficiency. Overall, findings

---

\*Speaker

suggest that learners' perception of semantic relatedness may be a more appropriate factor to explain the acquisition of multiple meanings in L2 English. The methodological and pedagogical implications of the findings will be discussed in detail.

**Keywords:** second language vocabulary acquisition, etymological relatedness, polysemy, homonymy, perceived semantic relatedness

# Are non-words really superior to real words to train L2 sounds?

Celia Gorba <sup>\*† 1</sup>, Miquel Llompart <sup>2</sup>, Pilar Prieto <sup>2</sup>

<sup>1</sup> Universitat Autònoma de Barcelona – Spain

<sup>2</sup> Universitat Pompeu Fabra – Spain

High variability phonetic training (HVPT) has proven its effectiveness regarding the improvement of L2 phonetic learning (1). Both non-words (NW) and real words (RW) have been used in phonetic training regimes, resulting in phonetic improvement, but some previous studies report greater gains when training with NWs as opposed to RWs, as the former allow learners to focus on the distinctive phonetic features while eliminating interference from the lexicon (e.g., 2, 3). However, the tasks typically used to test both RWs and NWs (i.e., phoneme identification and discrimination) effectiveness are purely phonetic, and recent studies found that accuracy in prelexical categorization does not necessarily translate into the establishment of robust distinctions at the lexical level (4, 5). These mismatches thus question the generalizability of NW training regimes to lexical processing.

The current study compares the effectiveness of an audiovisual HVPT regime using NWs to a comparable regime using RWs for gains at a phonetic as well as lexical level, both in perception and production. L1-Catalan/Spanish learners of English with an upper-intermediate proficiency level completed a total of 6 15-minute training sessions (60 trials each) involving the identification of English front vowels (/i/, //, // and /æ/). A subset of participants completed an audiovisual training with RWs, whereas the other half were trained using comparable NWs. All participants completed a series of pre- and post-training tasks which varied in the extent to which they required lexical processing. As for perception, trainees completed a word picture matching task (WPM) and an identification task involving RWs followed by NWs. Regarding production, trainees took part in a sentence reading task, followed by a word reading task and a delayed repetition task, the latter involving RWs first and then NWs.

So far, the data of 24 participants (NW group: 11; RW group: 13) has been analyzed and about 50 participants are predicted to finish the training this fall. Preliminary results show an overall higher accuracy in the identification tasks than in the WPM task both at pre- and post-test. In addition, improvement from pre-test to post-test also appears to be considerably larger for the former, which signals a greater gain at a phonetic than at a lexical level. Interestingly, regarding identification, the group trained with RWs improved to a greater extent than the NW group not only in the identification of RWs but also for NW stimuli. When it comes to production, results also show a general improvement in the distinction of the /i:/-// pair, and, to a lesser extent, the //-/æ/ contrast. However, this gain is again greater in the case of the RW group, particularly in the two reading tasks, which involve a greater degree of lexical processing than the repetition tasks. Thus, preliminary results do not show a superiority of NW over RW. Quite on the contrary, the RW group tends to show greater gains, especially in tasks that require more lexical activation. If confirmed with the larger sample, these results have important implications for the conceptualization of phonetic training regimes.

---

\*Speaker

†Corresponding author: [celia.gorba@uab.cat](mailto:celia.gorba@uab.cat)

**Keywords:** phonetic training, HVPT, real words, non words, generalization, lexical activation

# Subject and object pronominal drop in monolingual and bilingual children: Evidence from Russian and Hebrew

Galina Gordishevsky \*†<sup>1</sup>, Natasha Dvorina \* ‡<sup>2</sup>, Natalia Meir<sup>2</sup>

<sup>1</sup> Ben-Gurion University of the Negev – Israel

<sup>2</sup> Bar-Ilan University [Israël] – Israel

The study explored third-person subject and object pronoun production in monolingual and bilingual Russian-Hebrew-speaking children, considering the unique features of Russian and Hebrew regarding null pronouns. Hebrew syntactically licenses first- and second-person subject pro-drop, while third-person subject and object pro-drops are discourse-bound (Landau, 2018). Russian is a discourse-drop language (Franks, 1995). We aimed to assess the potential interaction between the two languages in bilinguals' choice of null and overt pronouns, recognizing the vulnerability of pronoun choice in bilingual acquisition due to (a) cross-linguistic influence and/or (b) a higher ambiguity tolerance threshold (Serratrice et al., 2004).

We recruited 66 children aged 4;2-8;0 and 40 adults: bilingual Russian-Hebrew-speaking (BiliCh, n=28), monolingual Russian-speaking (MonoCh-Ru, n=20), and Hebrew-speaking children (MonoCh-He, n=18), monolingual Russian-speaking (MonoAd-Ru, n=22), and Hebrew-speaking adults (MonoAd-He, n=18). The bilingual children were tested in both languages, their Heritage language, Russian, and the Societal language, Hebrew. We evaluated morpho-syntactic abilities of all children via LITMUS Sentence Repetition Tasks (Marinis & Armon-Lotem, 2015). Bilinguals showed lower morpho-syntactic abilities than both monolingual controls, and as a group, they were more dominant in Russian. Pronoun elicitation tasks tested third-person subject and object pronouns using pictures (Subject prompt: The boy is wet because... Target: he/∅ jumped into the puddle; Object prompt: The baby is laughing because... Target: the father is tickling him/∅).

The generalized linear mixed logistic regression included a random intercept for each item and a random intercept for each participant. The Group and Syntactic Position (Subject, Object) were entered as fixed effects, with the dependent variable coded as Null Pronoun=0, Overt Pronoun=1. First, our findings confirmed cross-linguistic differences between subject and object drop in adult controls: both monolingual adult groups used null subjects, yet to different extents (Russian: 97%; Hebrew: 18%); no null objects occurred in the adult data. Second, Hebrew-speaking monolingual children exhibited adult-like use of subject and object pro-drop, whereas Russian-speaking monolinguals were more likely to resort to overt pronouns in the subject position compared to adults. Thirdly, for subject-drop, bilingual children paired up with monolingual child controls in Hebrew, yet differed from Russian-speaking controls. For object drop, all groups converged. In bilingual children, age was related to subject drop: as children grow older, in Hebrew, they start using more overt pronouns, yet in Russian, they start using more null pronouns.

Our findings show that not only bilingual children might overuse overt pronominal subjects,

---

\*Speaker

†Corresponding author: galinag@post.bgu.ac.il

‡Corresponding author: n.dvorina@gmail.com



as predicted by Serratrice et al.'s (2004) hypothesis positing the vulnerability of the syntax-pragmatic interface, but also monolingual children might resort to overt pronouns when adults choose null elements. We believe that the data do not support the cross-linguistic influence account, as the differences were observed in the child's more proficient Russian rather than Hebrew. The data also question an increased intolerance to ambiguity in bilinguals since Russian-speaking monolingual children also overused overt pronouns. We suggest that the findings show that both monolingual and bilingual children might go through a protracted path in null-vs-overt pronoun acquisition (see also Iraola Azpiroz et al., 2017).

**Keywords:** first language acquisition, second language acquisition, pronoun use, Russian, Hebrew, pro drop, morphosyntax, pragmatics

# Motivation to learn a second foreign language in upper secondary school: A preregistered intervention study

Jonas Granfeldt \* <sup>1</sup>, Céline Rocher Hahlin<sup>†</sup> <sup>2</sup>, Joost Van De Weijer<sup>‡</sup> <sup>3</sup>,  
Alastair Henry<sup>§</sup> <sup>3,4</sup>

<sup>1</sup> Lund university, Centre for Languages and Literature (CLL) – Lunds universitet - Box 201 - 221 00  
Lund, Sweden

<sup>2</sup> Dalarna University – Sweden

<sup>3</sup> Lund University – Sweden

<sup>4</sup> University West – Sweden

## Background

A major reason supporting the conceptualization of L2 motivation as a self-system (Dörnyei, 2009) was the possibility to develop interventions based on possible selves. While reports of more than a dozen interventions can be found in the research literature (e.g. Safdari, 2021, Rocher Hahlin & Granfeldt, 2021), and while results are often promising, studies suffer from design weaknesses. In a critical overview of published studies, Vlaeva & Dörnyei (2021) found that despite recognition of the need to carry out and measure the effects of an intervention over longer timeframes, intervention periods are short, and delayed post-testing is rare. Moreover, methodological rigor lacks the standards accepted in experimental paradigms. Teaching in intervention/control groups is often conducted by the researcher, and effective randomization is frequently lacking. Lastly, outcome measures are generally limited to a restricted range of self-report measures. In addition to methodological and measurement limitations, studies have focused almost exclusively on English, and intervention activities have been poorly integrated into regular teaching.

## Study design

To overcome these weaknesses, and to better evaluate the utility of interventions based on possible selves, we designed a long-term intervention (6 months) to be simultaneously carried out at 2 schools, for 3 second foreign languages (French, German, Spanish), and involving 12 classes (6 intervention and 6 control) taught by 6 collaborating teachers (each teaching one intervention and one control class). The project was designed as research–practice partnership (Penuel & Hill, 2019) where the design of intervention activities took place in close collaboration with the teachers. The main study was preregistered, and the first wave of data collection took place in September 2023.

## Method

---

\*Speaker

<sup>†</sup>Corresponding author: crc@du.se

<sup>‡</sup>Corresponding author: joost.van\_de\_weijer@humlab.lu.se

<sup>§</sup>Corresponding author: alastair.henry@hv.se

The study is a 2 x 4 mixed factorial experiment where the main factors are the intervention (two levels, between-subjects) and time (four occasions, within subjects). Classes were randomly designated as intervention or control. Background variables included (e.g) self-reported multicultural personality, visualizing capacity and previous foreign language learning experiences. Self-report outcome measures included (e.g.) ideal L3 self, engagement and self-efficacy. Observed measures were of writing and vocabulary.

## Discussion

In this presentation we present and discuss the study's rationale, design, and the role of the collaborating teachers. We will also present initial findings from pre and post-test measures of self-reported outcome variables.

## References

- Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9-42). *Multilingual Matters*.
- Penuel, W. R., & Hill, H. C. (2019). Building a knowledge base on research-practice partnerships: Introduction to the special topic collection. *AERA open*, 5(4).
- Rocher Hahlin, C., & Granfeldt, J. (2021). Strengthening L3 French motivation: The differential impact of vision-enhancing activities. *Languages*, 6(1), 47.
- Safdari, S. (2021). Operationalizing L2 motivational self system: Improving EFL learners' motivation through a vision enhancement program. *Language Teaching Research*, 25(2), 282-305.
- Vlaeva, D., & Dörnyei, Z. (2021). Vision enhancement and language learning: A critical analysis of vision-building in an English for Academic Purposes programme. *Language Teaching Research*, 25(6), 946-971.

**Keywords:** motivational self, system, intervention, L3, research, practice collaboration

# The processing of passive sentences in German and French, L1 and L2: Evidence from an eye-tracking study with adults

Cyrille Granget <sup>\*† 1</sup>, Isabel Repiso <sup>\* ‡ 2</sup>, Pierre-Vincent Paubel <sup>3</sup>, Martin Haiden <sup>4</sup>, David Imbert <sup>4</sup>

<sup>1</sup> Laboratoire de NeuroPsychoLinguistique – Université Toulouse - Jean Jaurès – France

<sup>2</sup> Paris-Lodron-Universität Salzburg – Austria

<sup>3</sup> Cognition, Langues, Langage, Ergonomie – Ecole Pratique des Hautes Etudes, Université Toulouse - Jean Jaurès, Centre National de la Recherche Scientifique, Toulouse Mind Brain Institut – France

<sup>4</sup> Laboratoire de Linguistique de Nantes – Centre National de la Recherche Scientifique, Nantes Université - UFR Lettres et Langages, Nantes Université - UFR Lettres et Langages – France

While Ferreira (2003) and Lee & Doherty (2019) show that processing passive sentences in English and Spanish L1 respectively represents a greater cost for the adult than processing active utterances, Cristante and Schimke (2019) show that in German, passive utterances are no more difficult to process than active ones. These results invite us to extend research on the processing of passive sentences to other languages and examine the influence of L1 on L2, where no study exists apart from that of Lee and Doherty (2019).

Our study thus focuses on the role of L1, French versus German, in the online processing of active and passive sentences in L2, German and French, and the apprehension of the visually related events. Based on previous studies, we hypothesize a German advantage over French in the processing of passive sentences in L1 and in French L2 and a French L1 disadvantage in the processing of passive sentences in French L1 and German L2. Therefore and following Sauppe and Flecken (2021), we hypothesize that processing passive sentences activates representations, changes the apprehension of events and the proportion of first fixations in the agent AoI.

To examine these hypotheses, we used behavioral and eye-tracking data: 38 participants (22 German L1 and 16 French L1) took part in an online comprehension task involving 28 utterances including 16 targets in L1 and L2, French vs. German. Participants heard an utterance (active or passive), then saw an image (congruent vs. incongruent, depending on whether the left-right position of the characters corresponded to the order of the utterance’s constituents) (1000ms) and indicated whether the image represented the expected sentence. Our analyses focused on 16 target trials corresponding to 4 actions (washing, styling, hitting, pinching) x 2 syntactic conditions (active vs. passive) x 2 image orders (agent in front left vs. agent behind right), for which we analyzed the accuracy rate, reaction times and proportion of first fixations in the agent AoI.

Initial results invalidate the hypothesis of a German advantage: the accuracy rate in L1 German is not significantly better than in L1 French: responses are more than 97% correct in both L1 French and German. The accuracy rate in French L2 (German L1) is as high as in German L1 and French L1. On the other hand, results in German L2 are significantly lower than in German L1 and French L2, which can be explained by the lower score of participants in the German L2 level test. Generally speaking, the analyses also show a significant effect of syntactic priming

---

\*Speaker

†Corresponding author: [cyrille.granget@univ-tlse2.fr](mailto:cyrille.granget@univ-tlse2.fr)

‡Corresponding author: [isabel.repiso@sbg.ac.at](mailto:isabel.repiso@sbg.ac.at)

on the rate of correction, which is lower with a passive sentence, and no effect of congruence. Reaction time analysis corroborates these results. On the basis of the analyses carried out so far, this study does not show an influence of the L1 on the processing of passive sentences in L2 but suggests an effect of developmental stage in L2.

**Keywords:** passive sentences, French, German, eye, tracking, crosslinguistic influence

# The relationship between the situation model building ability, the episodic buffer and L2 writing process and product

Zhixing Han <sup>\*† 1</sup>, Rasmus Steinkrauss <sup>1</sup>, Marije Michel <sup>1</sup>

<sup>1</sup> University of Groningen [Groningen] – Netherlands

Writing is a complex process where writers negotiate meaning and deploy linguistic and extralinguistic resources to achieve a specific communicative goal (Durrant et al., 2021). In the process of writing, writers engage in knowledge-constitution, in which the ideas are gradually constructed over repeated synthesizing and knowledge-transformation, in which ideas are manipulated and organized into a coherent whole to align with the rhetorical goal (Galbraith & Baaijen, 2018). To a large extent both processes rely on retaining ideas and concepts and organizing them into a coherent mental model.

The ability to build coherent mental representations can be measured by the situation model building ability. A situation model is defined as the overarching understanding of the text or situation, consisting of the concepts and the relationship between them, which makes reasoning and inferences possible (Kintsch & van Dijk, 1978). This ability to build situation models has been associated with L2 reading comprehension (Raudszus et al., 2019). However, although constructing mental models also happens in writing, whether this ability can also affect L2 writing processes and products remains unknown.

While organizing concepts and ideas coherently requires the situation model building ability, the episodic buffer is in charge of retaining those ideas and concepts and their relationships. The episodic buffer is the component of working memory which stores the bindings of single- and cross-modality information (Baddeley, 2012). The situation model, in the form of bound coherently organized phonological and visuospatial information, is assumed to be retained by the episodic buffer, which therefore can play a role in L2 writing. However, to the best of our knowledge, no study on working memory in L2 writing has included the episodic buffer.

We investigate the following two research questions: 1) What is the relationship between the situation model building ability and L2 writing processes and products? 2) What is the relationship between the episodic buffer capacity and L2 writing processes and products?

We aim to recruit 30 L2 English learners as participants. They will be tested on their situation model building ability and the episodic buffer capacity alongside performing a writing task. The participants' situation model building ability will be evaluated using the pathfinder network approach (Raudszus et al., 2019). The participants will read an essay, and then, organize the keywords of the essay. Then, this data will be transformed into a network that represents the situational model built out of the text. The network will be compared with an expert network to generate a similarity index. We will gauge episodic buffer capacity by a test targeting the storage of the product of binding (Kofler et al. 2018), adapted by combining phonological and visuospatial working memory tests in Rapport et al. (2008). We will monitor participants' writing processes by Inputlog 8.0 (Leijten & Van Waes, 2013) and score their written productions

---

\*Speaker

†Corresponding author: z.han@rug.nl

with the CAFIC rubric (Hou et al., 2016). By involving the situation model building ability and episodic buffer capacity, this study will expand our knowledge on the potential factors influencing L2 writing processes and products.

**Keywords:** keystroke logging, writing processes, individual differences, working memory, situation model

# Processing generic reference in French and English: an elicited imitation study

Coralie Hervé \* <sup>1</sup>

<sup>1</sup> Formes et représentations en linguistique, littérature et dans les arts de l'image et de la scène –  
Université de Poitiers – France

Differences between L1-L2 syntactic processing (i.e., grammatical violation detection) are attributable to cognitive limitations, L1 influence, age of acquisition, language dominance and L2 linguistic representation (N. Ellis, 2015, MacWhinney, 2018). Recent L2 research on generic and specific reference specifically has shown that French L2ers can display as much implicit and explicit representations of article use as English monolinguals during online and offline sentence written comprehension (Hervé & Lawyer, accepted). It also highlighted that the encoding of mass NPs as opposed to plural NPs is more ambiguous even for English native speakers. Yet, little is known about how French L2ers and English native speakers apply their language-specific linguistic representation of genericity and specificity during real-time aural sentence processing. In the present study, we set out to assess (i) whether the co-activation of the language-specific morphosyntactic forms (articles) depending on the semantic context (specific vs. generic) affect French-English bilinguals' implicit and explicit linguistic representation during aural sentence processing; and consider (ii) the role of language exposure and of the relative fluency on the direction and magnitude of cross-linguistic transfers. French is the most restrictive Romance language as it requires the projection of a determiner in argument position when English allows bare mass nouns and bare plural nouns in generic (and indefinite specific) contexts.

Thirty L1 French L2 English speakers (mean age = 24.47), 30 L1 English L2 French speakers (mean age = 24.47), and 30 English monolinguals (mean age = 20.30) took part in an elicited imitation task (i.e. tapping into their implicit knowledge) and an untimed acceptability judgement task (i.e. calling for their explicit knowledge) (Godfroid, et al., 2015). LEAP-Q questionnaires (Marian, et al., 2007) were used to measure the participants' level of proficiency. L2 grammatical knowledge was measured with an Oxford Placement test. The experimental design manipulated grammaticality in two conditions (generic; specific) within subjects, as well as group, and proficiency measures between-subjects. Language exposure was treated as a continuous predictor. The materials included 16 paired-sentences with a subject NP in generic context, 16 in specific context controlled for noun type (mass vs. count noun).

In the English AJT, all the participants rated higher grammatical than ungrammatical sentences in the generic condition. In the specific condition, the English native participants discriminated grammaticality violation when the French learners did not. The effect of noun type (plural vs. mass NPs) indicates that the French learners were more accurate with plural nouns than with mass NPs. However, there was no effect of language exposure or fluency. In the elicited imitation, the English native speakers overwhelmingly corrected ungrammatical sentences in the two conditions. However, the French learners corrected significantly fewer sentences than the native speakers. Furthermore, the French learners were significantly less accurate in the generic than in the specific condition – surprisingly the opposite pattern from that observed in the AJT. Plural NPs were corrected more often than mass NPs. Fluency measures predicted the likelihood of

---

\*Speaker



sentence correction. These results will be discussed with reference to the bilinguals' implicit and explicit knowledge.

**Keywords:** linguistic representations, elicited imitation, genericity, language exposure

# Learning from processing: abstract structural priming across grammatical structures and languages in early L2 development

Holger Hopp \* <sup>1</sup>, Sarah Schimke <sup>2</sup>, Freya Gastmann <sup>2,3</sup>, David Öwerdieck <sup>1</sup>, Gregory Poarch <sup>3</sup>

<sup>1</sup> TU Braunschweig – Germany

<sup>2</sup> LMU Munich – Germany

<sup>3</sup> RUG Groningen – Netherlands

This study explores recent proposals that language processing mechanisms do double duty as language learning mechanisms (Bovolenta & Marsden, 2021; Phillips & Ehrenhofer, 2015). We test whether cumulative structural priming leads to learning of non-canonical object-initial word orders in a pre-posttest priming study on relative clauses among 165 (136) low-proficiency L2 learners of English. Specifically, we investigate the scope of structural priming by assessing if priming with an earlier-acquired related L2 structure (questions) or relative clauses in the L1 also promotes learning of later-acquired (object) relative clauses in the L2. Based on the assumption that relative clauses and questions are related at the level of sentence processing (Stowe, 1986), we further examine if cumulative priming goes hand in hand with changes in processing, namely a reduction in the subject-first preference to interpret the first noun as the agent/subject of the sentence.

Three groups of adolescent L1-German L2-English learners took part in a pretest and posttest that established the interpretation and processing of relative clauses using picture selection during eye-tracking. These tasks sandwiched a priming task, in which the groups were presented with unambiguous form-meaning mappings in the prime sentences and relative clause targets, all of which used different lexical items from pre- and posttest.

In the priming task, the EN-RC group ( $n = 44$ ) received English relative clause prime sentences in which the verb cued the object interpretation ("watch" vs "clean"). The EN-Q group ( $n = 43$ ) heard corresponding English question prime sentences ("Which bear does the duck clean?"), and the GE-RC group ( $n = 49$ ) received German relative clause prime sentences. The EN-RC and the EN-Q group had substantial and comparable increases in the accurate comprehension of object relative clauses and questions from pretest to posttest (+17.5%), while accuracy of subject-initial orders slightly decreased (-4.5%). For the GE-RC group, increases for object orders were also significant (+7%), though at lower effect sizes. These results show that cumulative abstract structural priming in the priming phase leads to increased posttest accuracy of non-canonical orders not just for identical structures in the L2 (relative clauses), but also across related L2 grammatical structures (questions) and also from L1 to L2 relative clauses. Critically, increased accuracy was accompanied by reductions in the subject-first preference in initial processing of object-initial orders. At posttest, learners made fewer looks to the subject interpretation before disambiguation by the second noun. This indicates that cumulative priming went hand in hand with changes in parsing strategies across the grammatical

---

\*Speaker

structures to which they apply. Such correspondences between priming and processing suggest that the revision of general processing strategies, cued by unambiguous prime sentences, may drive grammatical learning of non-canonical word orders in a developing L2 via priming. The study shows that learning via priming generalizes across related grammatical structures in the L2 and the L1 even among learners at low L2 proficiency. The study bolsters claims that priming acts as an across-structure learning mechanism in L2 acquisition and is accompanied by processing changes. We discuss implications for FL teaching.

**Keywords:** grammar learning, priming, eye, tracking, non, canonical word order

# L2 Acquisition of Chinese Perfective Motion Events

Shanshan Hu \* <sup>1</sup>

<sup>1</sup> University of Cambridge – United Kingdom

This study investigates the argument structure of motion events and their interaction with perfectivity within the framework of the Interpretability hypothesis (Tsimpli 2003; Tsimpli & Dimitrakopoulou 2007). Focusing on Chinese native speakers and L2 learners with English as their L1, we examine the role of grammatical aspect (perfective vs imperfective) and semantic make-up (manner vs path) of verbs in the distribution and interpretation of spatial prepositional phrases (PPs).

Chinese and English differ in the distribution of spatial PPs and perfective/imperfective markings in relation to telic/atelic motion predicates. In Chinese, postverbal complement PPs (Goal-denoting) and perfectives only select telic motion with an inherent endpoint, while preverbal adjunct PPs (Location-denoting) and imperfectives associate with atelic situations. In English, the postverbal PP can function as both adjunct and complement and may be ambiguous between Goal and Location readings. It can appear in both perfective and imperfective contexts, although Goal-PPs are preferred in perfective contexts and Location-PPs are preferred with imperfective verbs.

Using Grammaticality Judgment Task and Sentence-Video Matching Task, this study will answer two questions: a) how do English speakers acquire the word order and distribution of Chinese location/goal PPs with telic and atelic motion verbs? b) How do English speakers use the perfective marker *le* and the durative marker *zhe* in Chinese across telic and atelic motion events, and how do they interpret PPs with different aspectual markings? The study will examine both L1-transfer effects and the semantics of verbs.

This study associates telicity and perfectivity to explore the argument structure of Chinese motion events, an area remaining relatively unexplored. Furthermore, this study aims to provide new evidence on the role of perfectivity in the argument structure in L2 learners, shedding light on the acquisition at the syntactic-semantic interface.

**Keywords:** Perfectivity, Motion Events, Chinese, L2 Acquisition

---

\*Speaker

# Reading-while-listening is efficient and enjoyable, but it does not improve comprehension

Bronson Hui \* <sup>1</sup>, Aline Godfroid \* <sup>†</sup> <sup>2</sup>

<sup>1</sup> University of Maryland System – United States

<sup>2</sup> Michigan State University – United States

Audiobooks enjoy widespread use in early childhood education as a useful pedagogical tool to support literacy development (e.g., Rasinski & Hoffman, 2003). In Second Language Acquisition, the practice of reading-while-listening to audiobooks has likewise been promoted because of its reading- and vocabulary-related benefits. One such belief is that reading-while-listening may lead to better comprehension than reading-only (Amer, 1997; Dhaif, 1990; Taguchi, Gorsuch, Takayasu-Maass, & Snipp, 2012; Woodall, 2010; but see Tragant-Mestres et al., 2019 and Diao & Sweller, 2007 for conflicting evidence).

In this registered report, we revisited this belief, and, adopting a psycholinguistic perspective, tried to advance our understanding of why exactly audiobooks could have favorable effects on comprehension and for whom they might be more beneficial. We hypothesized that reading-while-listening's advantage for comprehension could depend on the learner's specific decoding profile in reading (i.e., their orthographic decoding skills) and listening (i.e., their ability to segment the sound stream).

Eighty-four intermediate to advanced Chinese learners of English read (Reading-Only), read and listened (Reading-While-Listening), and listened (Listening-Only) to excerpts of the novel *Crime and Punishment*. They were exposed to the novel only once, but experienced different parts of the text under each of the three conditions (i.e., within participants). We measured participants' orthographic decoding (a key component of reading) using reading times obtained through eye tracking. We also measured their segmentation skills (a key component of listening) using a shadowing task that was designed based on the text.

The overall hypothesis was not supported: Reading while listening was not beneficial for comprehension. Rather, it led to a small, significant comprehension cost relative to reading only.

The primary benefits of reading-while-listening were in reading speed. Reading-while-listening caused readers to process the text faster compared to reading only. Those who possessed strong orthographic decoding skills (as shown by the eye tracking) were found to be better comprehenders, but critically, the benefits or, in this case, the costs of reading-while-listening did not depend on the participants' individual differences in reading profile. Segmentation skills were not related to comprehension in any of the three conditions. This suggested that reading-while-listening is primarily a form of reading, rather than a mixed reading-plus-listening activity.

The results strengthen our theoretical understanding of reading-while-listening's presumed pedagogical benefits. Reading-while-listening is an efficient and enjoyable way of presenting learners to new texts, but it does not improve comprehension, to the contrary.

---

\*Speaker

<sup>†</sup>Corresponding author: godfroid@msu.edu

**Keywords:** eye tracking, reading, while, listening, comprehension

# Undergraduates' Plurilingual and Pluricultural Competence in the Context of a European Internationalized University. Relationship with their EFL Self-Concept from a longitudinal perspective

Socorro Ibáñez \* <sup>1</sup>, Mireia Trenchs-Parera \* † <sup>1</sup>

<sup>1</sup> Universitat Pompeu Fabra – Spain

The Plurilingual and Pluricultural Competence Scale (PPCS) was designed to inform linguistic and intercultural education and policies given the increasing international mobility of university students across the world (Galante, 2020). Stemming from sociolinguistic theories, such as plurilingualism (CoE, 2018; Marshall & Moore, 2018; Moore & Gajo, 2009) and translanguaging (Galante, 2019; Piccardo, 2017), it intended to gather large pools of data on plurilinguals' language practices and pluricultural awareness. Successfully validated and implemented in two studies in multilingual cities in Canada (Galante, 2020), the PPCS suggests that language and culture are inseparable dimensions. However, the scale still lacks validation in other linguistic and cultural contexts and other instruments and methods need to be used together with it, as Galante (2020) herself highlighted, so that underlying reasons behind different or similar PPC levels could be captured. In order to contribute to these research gaps, we have replicated her study in Barcelona, Spain, , a context fit for the replication of the study, given the richness of plurilingual practices predominant in its territory (Newman, Trenchs-Parera & Corona, 2019). This time from a longitudinal perspective, with the addition of a scale about students English-as-a-Foreign-Language (EFL) Self-Concept (Iwaniec, 2014). This latter construct deals with the self-evaluation of abilities and feelings of worth associated to the learning and command of English (Mercer, 2011; Pajares & Shunk, 2005). Specifically, we wanted to investigate to which extent the PPC would act as a predictor of higher EFL Self-Concept values. At the beginning of the academic year, we recruited 520 first-year local volunteering bilingual/multilingual undergraduates. At the end of the year, 466 undergraduates answered our questionnaire, 126 of which had already participated in time 1. Exploratory factor analysis was carried out to examine factors in the PPC scale and PPC was explored longitudinally through independent samples t-tests. A linear regression among PPC and EFL Self-Concept data was performed for both data collection times. Our analysis in either time of data collection did not yield results that were comparable to Galante's, despite contextual similarities. Our longitudinal component yielded a statistically significant change towards higher PPC levels at the end of the academic year in two of the three factors in the scale. The factor describing positive traits obtained in the PPC scale showed a mild positive correlation with EFL self-concept, and the linear regression showed it could account for around 20-25% of the variance in EFL Self-Concept. Further research will be conducted in order to re-configure the scale so that PPC can be accurately measured in this context in the future and to allow for a more robust exploration of its seeming relationship with EFL Self-Concept.

---

\*Speaker

†Corresponding author: mireia.trenchs@upf.edu

**Keywords:** plurilingual and pluricultural competence, EFL Self, Concept, Higher Education, longitudinal study



# Lexical competence in French middle school learners of English (is a bee the same as a fly?)

Marie-Pierre Jouannaud \*<sup>1</sup>, Pascale Manoïlov<sup>2</sup>

<sup>1</sup> Université Paris 8 Vincennes-Saint-Denis - UFR Langues et cultures étrangères – Université Paris 8 Vincennes-Saint-Denis – France

<sup>2</sup> Université Paris Nanterre - UFR Langues et cultures étrangères – Université Paris Nanterre – France

The role of the lexicon in foreign language (FL) teaching in France has evolved since the adoption of the communicative method in the late 1980s, followed by the action-oriented approach at the start of the 21st century. This evolution was accompanied by the gradual disappearance of the lexical syllabus previously present in official instructions for foreign languages. Yet the lexicon is an essential tool to understand and express oneself, and to interact with other speakers. Numerous studies have shown that lexical knowledge is highly correlated with L2 listening (Vafaei, 2020), speaking (Hilton, 2008), and reading and writing (Milton, 2013). It also occupies a central place in psycholinguistic models of speech processing for language comprehension and production (Cutler, 2012; Levelt, 1989).

The research question guiding our study is as follows: given the secondary status of lexis in the official FL curriculum, do French students possess the lexical knowledge necessary to express themselves in English at the expected CEFR levels (A2 at the end of middle school and B2 at the end of high school)? To answer this question, we have taken advantage of the movement toward open access of public data (MEN, 2021). The French ministry of education conducts regular evaluations of schoolchildren's proficiency in FL: every 6 years, a sample of students is tested in listening, reading comprehension and writing. In the latest iteration (CEDRE 2021-2022), a written test of vocabulary recall was added, and these data were recently made accessible to researchers upon request.

The participants are 3,138 middle schoolers in 9th grade (14-15 years old). Item difficulty of the vocabulary test will be calculated with Item Response Theory (IRT), each student having taken only a subset of the 197-item test. We will try to characterize students' IRT vocabulary scores by reference to items of corresponding difficulty on the IRT scale, their frequency and lexical field. We will also present correlations between students' vocabulary knowledge and their results in listening, reading and writing. Finally, we will analyse the lexical richness of a sample of learners' written productions. This will allow us to obtain a more complete picture of students' lexical competence (Kojima & Yamashita, 2014).

Since Meara's 1980 call and the renewed awareness of the importance of the lexicon in foreign language teaching and learning (Schmitt 2008), there have been numerous studies on L2 lexical knowledge, but our study is different in two respects. On the one hand, we will present a historical perspective on the place of vocabulary instruction in French schools based on changing official curricula. On the other hand, our study focuses on learners that are rarely present in SLA studies in France, namely middle schoolers. Although researchers in other countries have published studies with this population (e.g. Canga Alonso, 2013; Cobb & Horst, 2011), we are

---

\*Speaker

not aware that any studies have been conducted on the English vocabulary knowledge of French secondary school children since Arnaud (1985), and consequently, little is known on the topic.

**Keywords:** instructed learning, lexical competence, foreign language curriculum, item response theory

# Metalinguistic transfer in Slavic intercomprehension: the case of verbal aspect

Saturno Jacopo \* <sup>1</sup>, Valentina Nosedà <sup>2</sup>

<sup>1</sup> Università degli Studi di Bergamo – Italy

<sup>2</sup> Università Cattolica del Sacro Cuore - Milano (-) – Italy

The present study evaluates the ability of a group of L1 Italian learners of L2 Russian to select the correct value of verbal aspect in L3 Ukrainian, a language they had no experience of, but which is lexically and grammatically close to Russian.

In addition to a significant proportion of the lexicon, Russian and Ukrainian (both belonging to the East Slavic group) share numerous grammatical traits, including the systematic and lexicalized codification of aspect. For any verbal form, the appropriate aspectual marking – imperfective (ipf) or perfective (pf) – needs to be selected on the basis of the context (Zaliznjak & Šmelev 2000). The opposition between pf and ipf in Slavic languages differs markedly from that attested in other languages (Janda 2004): in Italian, to exemplify, verbal aspect is tied to the opposition between verbal tenses in the past.

The proximity between Russian and Ukrainian represents fertile ground for intercomprehension, broadly defined as the possibility to use (i.e. understand and/or produce) a language that one has no direct experience of, but is closely related to a known language (European Commission 2012). The research question of the study is to what extent the knowledge of an L2 structure can be transferred to a new L3 (Marx & Mehlhorn 2010; Pančíková & Horák 2020; Saturno 2022). Specifically, the presence of positive transfer in the selection of verbal aspect would suggest that the learning of L3s can significantly benefit from existing knowledge of a related L2 (Bardel & Sanchez 2017), especially if the latter is typologically distant from the learners' L1 and its mastering took significant time and effort (as is the case with L1 Italian and L2 Russian).

92 L1 Italian learners of L2 Russian filled in an online questionnaire (in three different versions) containing a Ukrainian text in which several verbs had been removed. The participants' task was to select the form that they deemed most appropriate by answering a multiple-choice question presenting the imperfective and the perfective form of the verb. All gaps were verified with the help of a Russian-Ukrainian bilingual native speaker to make sure that only one aspectual value would be acceptable. To establish the baseline of the participants' skills in L2 Russian, the participants also performed the same task on the Russian version of the text. The results show that while the accuracy of the participants' responses in Ukrainian is significantly lower than in Russian, it is still well above chance, which suggests that indeed, the participants could transfer their knowledge of the Russian aspectual system to its Ukrainian counterpart.

---

\*Speaker

**Keywords:** transfer, Slavic, intercomprehension, aspect, L1 Italian, L2 Russian

# Linguistic and trait determinants of self-regulation in multilinguals' use of linguistic landscape

Przybył Jakub \*<sup>1</sup>, Danuta Wiśniewska \*

1

<sup>1</sup> Adam Mickiewicz University, Poznań – Poland

Since the learner-environment interaction directly affects the process of language acquisition, investigations relying on agent-based frameworks need to account for the language learning environment and its relationship with the agent (cf. Dörnyei, 2009). Subscribing to Kaplan's (2008: 480) view, according to which learners' metacognition is "nested in and affected by environments", and recognising the status of the learning environment as one of the pillars of self-regulated learning (Andrade & Evans, 2015), we investigated multilinguals' self-regulation (SR) in the use of linguistic landscape (LL).

While constituting a virtually limitless source of target language (TL) input, understood as "everything in the TL that a learner is exposed to and has the opportunity to either hear or read" (Rast, 2008: 4), in order to be used as a learning resource LL requires individuals to select the learning content from its stock (Winne, 2018) as well as to manage it in prospective learning cycles (Zimmerman & Schunk, 2011). Relying on the Strategic Self-Regulation Model (Oxford, 2017), we operationalised those cycles as learning sequences involving seeking opportunities (SO), paying attention to cognition (PA), creating mental linkages (ML) and organising further learning (O). In order to identify the determinants of multilinguals' SR in the use of LL (LL SR), we developed, piloted, and validated the Linguistic Landscape Self-Regulation Inventory (LL SRI), consisting of the SO, PA, ML, and O subscales, which we subsequently used in data collection from 100 young adults, speakers of L1 Polish and L2English/L3German (53) or L2German/L3English (47). Our investigation was guided by the research question: **How do linguistic and trait characteristics contribute to the variation in multilinguals' SR in their use of LL?** To answer it, we supplemented the insights from the LL SRI with personality measurement based on participants' responses to the International Personality Item Pool (IPIP; Goldberg, 2006) and their self-assessments of reading comprehension skills in English and German (RCE, RCG).

After accounting for the standardisation, normalisation, reliability, and validity of our research instruments, we conducted path analysis, which revealed an overall good fit to the data ( $\chi^2(17) = 21.01$ ,  $p = 0.226$ ), CFI = 0.98, RMSEA = 0.049, 90% CI = (0.000, 0.106), SRMR = 0.079). We found statistically significant direct effects of RCE and RCG on participants' SR in the use of English ( $\beta = .44$ ,  $p < .001$ ) and German ( $\beta = .59$ ,  $p < .001$ ) LL respectively. We also unveiled indirect effects of personality traits: Intellect on LL SR in English through RCE ( $\beta = .15$ ,  $p = .009$ ), and Conscientiousness on LL SR in German through RCG ( $\beta = .10$ ,  $p = .049$ ), in mediation analysis. These outcomes partly corroborate the results of earlier research into the role of personality traits in instructional language learning (cf. Piechurska-Kuciel, 2020, Przybył & Pawlak, 2023), yet, strikingly, they also suggest that in the specific context of learning from

---

\*Speaker

LL, rather than relying on some universal characteristics of good language learners, learners acting as agents may draw on specific aspects of their personalities depending on the target language.

**Keywords:** Linguistic landscape, language learning environment, self, regulated language learning, personality traits

# ”And let’s move on”: and as a discourse marker in the teacher talk of Croatian EFL teachers

Eva Jakupčević \* <sup>1</sup>

<sup>1</sup> Faculty of Humanities and Social Sciences, University of Split – Poljička cesta 35 21000 Split, Croatia

Effective interaction management in the EFL classroom depends on the learners’ ability to successfully follow the lesson and appropriately participate in activities, which, in turn, depends on their ability to interpret the demands of the classroom context (De Fina, 1997; Walsh, 2006). Discourse markers (DMs) can facilitate understanding for learners and help them manage the complex communicational context that is the EFL classroom (Walsh, 2013). The discourse marker use of English language learners (and, to a lesser degree, teachers) has been an increasingly popular topic, with a number of studies focusing on different contexts and various individual markers or groups of markers. However, studies focusing on the use of *and* as a discourse marker are scarce, especially when it comes to the EFL classroom context.

The present study is a qualitative analysis of the use of *and* as a DM in teacher talk of 8 Croatian EFL teachers with learners in primary (aged 12/13) and secondary schools (aged 16/17). The analysis is based on the framework of DM use in the pedagogical context developed by Fung and Carter (2007). The study provides a detailed overview of the functions of *and* as a valuable and versatile DM. *And* was utilized by the teachers to a) express relations such as contrast and consequence, b) reflect their thought process, and c) as a useful discourse management tool by marking topic changes and continuation, introducing questions and making continuity more salient. The analysis showed that the Croatian teachers used this simple and frequent DM in a variety of ways, supporting their classroom management efforts and facilitating interaction in the classroom context.

De Fina, A. (1997). An analysis of Spanish *bien* as a marker of classroom management in teacher-student interaction. *Journal of Pragmatics*, 28(3), 337-354.

Fung, L., & Carter, R. (2007). Discourse markers and spoken English: Native and learner use in pedagogic settings. *Applied linguistics*, 28(3), 410-439.

Walsh, S. (2006). *Investigating classroom discourse*. Routledge.

Walsh, S. (2013). *Classroom discourse and teacher development*. Edinburgh University Press.

**Keywords:** discourse markers, teacher talk, and

---

\*Speaker

# Observing language minoritization through language teachers' minds: example of ESL/EFL teacher beliefs and practices about Guadeloupean Creole

Ingrid Jasor \*† 1,2

<sup>1</sup> Université du Québec à Montréal = University of Québec in Montréal – Canada

<sup>2</sup> Université des Antilles (Pôle Guadeloupe) – Guadeloupe

Language minoritization (Bazner & Lopez, 2022; Blanchet, 2005) happens when the political and socio-economic status of a language is qualitatively undermined in relationship to another. In response to the multilingual turn (May, 2014) that has tremendously shaped the ESL/EFL field in the last decades, valorizing students' plurilingual linguistic repertoires, all the more with a the presence of a minoritized language, are keys to inclusive and meaningful teaching practices. In parallel, language teacher cognition (Borg, 2019) has shown how intertwined the relationship between teachers' beliefs and their practices is, including their beliefs about languages. This study focuses on Guadeloupean Creole, a historically minoritized language primarily to French, the only official language of the French overseas department of Guadeloupe. Evolving in a diglossic yet complementary linguistic environment between French and Guadeloupean Creole (Anciaux & Prudent, 2021), the territory is also host of many other languages, including English that is widely taught in the archipelago. Using a combined methodology of visual narrative inquiry (Bach, 2007) and classroom observations, 3-4 Guadeloupean ESL/EFL teachers in secondary education were interviewed, had to create audio-visual life stories, and were observed in their classrooms; all to identify their beliefs about Guadeloupean Creole and the way they are articulated within their teacher practices. Through an initial interpretative phenomenological analysis (Antoine & Smith, 2017), these preliminary results demonstrate that, while teacher beliefs are key to increasing ESL/EFL teachers' pedagogical awareness, other factors such as thorough teacher training and plurilingual pedagogical resources are also needed to implement teaching practices that really include and valorize Guadeloupean Creole.

**Keywords:** language minoritization, English as a second/foreign language, Guadeloupean Creole, teacher beliefs, teacher practices

---

\*Speaker

†Corresponding author: [jasor.ingrid@courrier.uqam.ca](mailto:jasor.ingrid@courrier.uqam.ca)



# On the early acquisition of null subjects in Catalan and German as heritage or majority languages: early multilingualism in Germany and Catalonia

Amelia Jiménez-Gaspar <sup>\*† 1</sup>, Laia Arnaus Gil <sup>2</sup>

<sup>1</sup> University of the Balearic Islands – Spain

<sup>2</sup> University of Wuppertal – Germany

Languages differ as to whether subjects must be overtly expressed and, therefore, accompany the finite verb. With the aim to systematically describe what distinguishes null-subject from non-null-subject languages (i.e., NSLs and NNSLs), a cluster of properties for Italian was identified that excludes NNSLs such as French (Chomsky 1981; Rizzi 1982). However, evidence suggests that various languages are characterized by some of these properties while not fulfilling others (Newmeyer 2005). This has led scholars to study (N)NSLs in greater detail, allowing them to account for the null subject property (NSP) with a greater degree of granularity (Biberauer, Holmberg, Roberts & Sheenan 2014; Roberts 2019; Müller 2023). Following this idea, Catalan is described as a consistent NSL, while German should be considered a partial NSL (Biberauer 2010).

Empirical studies on early language acquisition have observed that monolingual children acquiring NSLs quickly align to adult rates of subject omissions, while children acquiring NNSLs take a long way to get to adult-like subject realizations (Bel 2003, 2005; Hyams 2011). Concerning the acquisition of the NSP in early bilingualism, much work has been done for the language combination English-Romance (cf. e.g., Licerias & Fernández Fuertes 2016). These studies observe high rates of subject realizations in Romance and lower omission rates, although still not adult-like in the early stages, than English monolingual adults. For the combination of a consistent and a partial NSL (the focus of our study), the studies of Patuto (2012) and Daniel (2021) confirm previous findings for the acquisition of a consistent NSL in early bilingualism: bilingual children at younger ages show lower subject omission rates (around 50%) which increasingly align to adult monolingual peers with age. For the partial NSL German, these studies observe high overt subject rates (around 90%), which drastically differ from the monolingual children. Their results further show that (null) subjects are neither related to language dominance nor to MLU.

Fifty-two multilingual children (mean age 7;3) acquiring German and at least one consistent NSL (i.e., Catalan and/or Spanish) were recruited in Germany (N=30) and in Catalonia (N=22). The analysis on Catalan (null) subjects in Germany is already available, but the analysis on Catalan in Catalonia and German in both countries is still ongoing. Catalan preliminary results indicate that younger children produce similar null and overt subject rates, while older children exhibit null subject rates similar to those of monolingual children and adults. (Null) Subject rates are independent of language dominance, as previous studies have observed. Interestingly, when the type of overt subjects is considered, a different pattern emerges: balanced children show a strong preference for pronominal subjects (70%), matching adult patterns (Biró 2017, de Prada 2009),

---

\*Speaker

†Corresponding author: amelia.jimenez@uib.es

whereas Catalan and German dominant children's pronominal subjects are located around 35%. Assuming these observations are accurate, we expect similar results for Catalan (and Spanish) in Catalonia. Following the previous literature on the early acquisition of German (Patuto 2012), we expect an acceleration effect in the overt subject rates of the multilingual children, irrespective of age and language dominance.

**Keywords:** heritage language, majority language, Catalan, Spanish, German, null subject property, null subject parameter, language dominance, language proficiency

# Mediating university EFL learners' inferential reading development through computerised dynamic assessment

Can Jin \* <sup>1</sup>

<sup>1</sup> University of Cambridge [UK] – United Kingdom

Rooted in Vygotsky's sociocultural theory and the concept of Zone-of-Proximal-Development (ZPD), Dynamic assessment (DA) has been recently introduced into L2 instructional settings as an innovative assessment and pedagogical approach. By unifying assessment and instruction into a dialectical activity, DA emphasizes on providing learners with assisted feedback (mediations) pitched at their ZPDs during the assessment process for a more comprehensive diagnosis and promotion of their L2 learning potential (Poehner & Lantolf, 2005). However, owing to the time-consuming and labour-intensive nature of the dialogic interactionist DA approach, the practicality of DA implementation in real L2 classrooms, especially in EFL classrooms with larger proportion of learners, remains a problematic issue. Therefore, this study aims to address this issue by devising a computerised dynamic assessment (C-DA) approach to mediate learners' microgenetic development (Wertsch, 1985) in L2 inferential reading comprehension in the large-scale EFL classroom. As part of a larger research project, this study adopted a mixed-methods design, with two parallel classes of freshmen EFL learners (N=34 and 38) in a Chinese university serving as the experimental group (CDA group) and the control group (Non-CDA group) respectively. An 8-week classroom-based intervention programme was sandwiched between a pre-test and a parallel post-test on English inferential reading. During the intervention programme, the CDA group received online weekly reading tests with automatic, stratified and contextualized mediations provided through a web-based computerised software whereas the control group received the same tests but without any mediations. Also, participants in the CDA group were asked to record their thinking processes during the reading tests about how they reached the final answers through think-aloud protocols and to submit these protocols online during Week 2, Week 4, Week 6 and Week 8. All the online reading tests were selected from TOFEL reading passages with multiple-choice inferential reading questions and the difficulty level of each passage was determined based on a quantifiable selection standard. Quantitative data was collected regarding participants' numerical scores in the pre-test, the intervention tests and the post-test as well as a set of individualised scores including actual/unmediated scores, mediated scores, gain scores and the Learning Potential Score (LPS) generated by the computerised software for participants in the CDA group. Qualitative data was collected concerning participants' submitted online thinking protocols transcribed verbatim. Preliminary inferential data analysis revealed that with similar pre-test scores, the CDA group performed significantly better than the Non-CDA group in the post-test, indicating the effectiveness of the C-DA programme. Simultaneously, for the CDA group, the interplay between participants' actual performances and mediated performances along with their LPS scores across the 8 C-DA tests showcased their divergent developmental trajectories as well as the non-linear and uneven nature of their microgenetic development in L2 inferential reading, which was further corroborated by a micro-analysis of participant's transcribed thinking processes. Limitations of this study as well as its pedagogical implications on L2 practitioners and learners will also be discussed.

---

\*Speaker

**Keywords:** sociocultural theory, computerised dynamic assessment, L2 microgenetic development, second language reading, EFL classroom

# Perception and production of voicing contrast in French stops by Wu and Mandarin speakers

Yishan Jin <sup>\*† 1</sup>, Damien Chabanal<sup>‡ 1</sup>

<sup>1</sup> Université Clermont Auvergne – Laboratoire de Recherche sur le Langage - UR 999 – France

The acquisition of voicing contrast in French stops is one of the most significant challenges for Mandarin speakers learning French, due to the absence of voiced categories in their native language. Nevertheless, Wu distinguishes itself from other Sinitic languages in having a three-way contrast between voiceless aspirated, voiceless unaspirated and voiced stops. The voiced series in the initial position of words, however, becomes phonetically voiceless while incorporating breathy phonation (Chao, 1956; Cao & Maddieson, 1992). According to the Perceptual Assimilation Model (PAM) (Best, 1995; Best & Tyler, 2007), cross-linguistic perception is influenced by L1 phonological knowledge. Conversely, the Speech Learning Model (SLM) (Flege, 1995) postulates that performance in second language speech learning is contingent upon the surface phonetic similarities and dissimilarities between L1 and L2 sounds. Furthermore, SLM suggests that more L2 experience enhances learners' ability to perceive and produce L2 sounds more effectively. From these predictions arise the following questions: 1) Can Wu speakers demonstrate a more proficient acquisition of the voicing contrast in French compared to the Mandarin group, given that voiced and voiceless plosives are classified into distinct categories in their L1? 2) Is the voicing contrast discerned more easily by the Wu group in word-medial position, where phonologically voiced initial stops are truly voiced? 3) Does a longer period of French learning lead to a better perception and production?

Considering the limited number of research studies that compared the performance of Wu and non-Wu speakers in acquiring French stops (Qin, 2010; Wang, 2020; Jin & Chabanal, 2022), the present paper conducted two types of tasks: a categorizing task and a reading task including 30 word pairs. This involved 30 participants who were native speakers of Wu and Mandarin, each with different proficiency levels in French. After measuring the accuracy rate and several acoustic cues using Praat and VoiceSauce softwares, we applied a series of multilevel regression models in the R environment to examine the influence of four predictors, namely L1, consonant, syllable position and duration of learning French. The results show that Wu speakers exhibit a high proficiency level in perceiving and producing the voicing contrast of French stops. In particular, the voiced categories produced by the Wu group are characterized by a more negative VOT and a greater V-ratio. We also observed that some Wu and Mandarin speakers use the following strategy to differentiate the French voicing contrast: nasalization of voiced stops in initial position and aspirated production of their voiceless counterparts. Moreover, the intervocalic position does not facilitate the acquisition of voicing distinction among Wu speakers. In fact, the phonetic cues used by these learners are VOT and V-ratio instead of phonation differences (f0, H1-H2, H1-A1, H1-A2 and CPP). In addition, an extended duration of French language learning does not appear to correlate with an enhanced perceptual capacity but rather with improved production. Together, this study illustrates that both the PAM and the SLM

---

\*Speaker

†Corresponding author: Yishan.JIN@doctorant.uca.fr

‡Corresponding author: damien.chabanal@uca.fr

achieve only partial success in explaining our results.

**Keywords:** L2 perception and production, French stops, voicing contrast, Wu and Mandarin speakers

# Oral fluency in the Spanish language classroom: Studying teachers' beliefs and practices in France and Sweden

Victor Jorméus \* <sup>1</sup>

<sup>1</sup> Dalarna University – Sweden

In the communicative paradigm oral proficiency is fundamental and successful language learners are expected to express themselves with fluency (Granfeldt et al., 2023). Even though oral fluency is well defined in applied linguistics and SLA research as a part of oral proficiency (e.g. Segalowitz, 2010), few studies discuss what teachers think and how they work with oral proficiency in the foreign language classroom (Baker-Smemoe et al., 2014; Bardel et al., 2023). After English, Spanish is the most studied language in French and Swedish schools. Both countries' curricula are based on the functionalist view in the Common European Framework of Reference for Languages (CEFR) where oral proficiency and fluency are key concepts. However, neither CEFR nor the respective national policy documents define what fluency is or how it should be taught.

Drawing on ecological and complexity theories, languages are regarded as Complex Dynamic Systems (CDST) (Hiver et al., 2022). Languages are seen as non-linear, complex systems with a multitude of variables. Consequently, language learning is also non-linear: Learners experience backslides, jumps, and stagnations in their learning process (De Bot, 2008; Larsen-Freeman, 1997). Teacher beliefs are a key concept since they could adequately explain what happens in the classroom (Burns, 1992). Beliefs and practices are related concepts even though there is no consensus regarding the impact beliefs have on teachers' practices (Borg, 1998). Teacher beliefs and classroom practices are also seen as Complex Dynamic Systems; taking a CDST perspective, teacher beliefs are situated in a state of constant change depending on context (Burns & Knox, 2011).

To bridge the gap between SLA theory and teacher practice (Spada, 2022), this doctoral project focuses on teacher beliefs and practices regarding oral proficiency, with a special focus on fluency. Using a Case Study design, upper secondary Spanish teachers in France and Sweden are interviewed then observed and videorecorded during various classes. The recordings are then used for a stimulated recall interview (Gass & Mackey, 2000). The interviews are transcribed and analyzed thematically (Braun & Clarke, 2006).

The study aims to better understand Spanish teachers' beliefs and practices regarding oral proficiency in a European context, and consequently offer a possibility of more equity in the definition and assessment of learners' oral fluency. In this doctoral workshop I hope to discuss both methodological and theoretical choices and implications for the project.

## SELECTED REFERENCES

Borg, S. (1998). Teachers' Pedagogical Systems and Grammar Teaching: A Qualitative Study. *TESOL Quarterly*, 32(1), 9. <https://doi.org/10.2307/3587900>

---

\*Speaker

Burns, A., & Knox, J. S. (2011). Classrooms as Complex Adaptive Systems: A Relational Model. *The Electronic Journal for English as a Second Language*, 15(1). <http://tesl-ej.org/pdf/ej57/a1.pdf>

De Bot, K. (2008). Introduction: Second Language Development as a Dynamic Process. *The Modern Language Journal*, 92(2), 166–178.

Hiver, P., Al-Hoorie, A. H., & Evans, R. (2022). Complex Dynamic Systems Theory in Language Learning: A Scoping Review of 25 years of Research. *Studies in Second Language Acquisition*, 44(4), 913–941. <https://doi.org/10.1017/S0272263121000553>

**Keywords:** oral fluency, teacher beliefs, stimulated recall, ELE



# Dampening of the bilingual brain response to negative words extends to speech comprehension and production

Rafał Jończyk \* <sup>1</sup>, Marcin Naranowicz <sup>1</sup>, Katarzyna Jankowiak <sup>1</sup>, Paweł Korpala <sup>1</sup>, Katarzyna Bromberek-Dyzman <sup>1</sup>, Guillaume Thierry <sup>2,3</sup>

<sup>1</sup> Adam Mickiewicz University in Poznań – Poland

<sup>2</sup> Bangor University – United Kingdom

<sup>3</sup> Adam Mickiewicz University in Poznań – Poland

Electrophysiological evidence has shown reduced sensitivity to negative content in bilinguals when they operate in the second language (L2, Wu and Thierry 2012; Jończyk et al. 2016, 2019; Zhang et al. 2023). These findings are often interpreted within the framework of L2 disembodiment (Pavlenko, 2012), underpinned by a hypothetical protection mechanism, whereby the brain shuts down access to L1 representations when faced with negative emotional content in L2 (Wu and Thierry, 2012; Zhang et al., 2023). The available evidence, however, is limited to written language comprehension and it is unknown whether similar mechanisms apply when bilinguals hear and produce emotional content, which is a frequent occurrence in natural communicative settings.

Here, we conducted two EEG experiments to investigate the production of written (Experiment 1) and spoken (Experiment 2) words in highly fluent Polish (L1)-English (L2) bilinguals (N = 35 per experiment). In both experiments, participants first saw emoji task cues that were either black ('read aloud') or white ('translate'), with lightness-task association rotated between participants, and that were either sad or neutral, announcing the affective valence of an upcoming word. Participants were tasked with producing or translating the word in their other language, according to emoji cue lightness. Translation direction (L1-L2, L2-L1) was blocked and counterbalanced between participants.

In Experiment 1 (written words), besides the expected reduction in N400 amplitude for negative L2 words (Jonczyk et al., 2016), we found that late positive potential (LPP) amplitudes elicited by negative words at the stage of preparation for production were larger in L1 (L2-L1 translation) than L2 (L1-L2 translation), suggesting that negative words to be produced in L1 undergo greater amounts of re-evaluation.

In Experiment 2 (spoken words), we found a robust reduction of N400 amplitudes for negative L2 words, extending to the auditory modality previous EEG evidence of reduced semantic processing of negative content in L2. As in Experiment 1, LPP amplitudes increased for negative as compared to neutral words at the stage of preparation for production in L1, but also when the task was simply to say the words aloud in L2 (control condition).

Analysis of event-related spectral perturbations in the theta and beta frequency bands revealed larger theta event-related desynchronization for L1 negative than L2 negative words and L1 neutral words, further supporting the involvement of mechanisms regulating negative emotion in L2 (Sulpizio et al. 2020). We also found that event-related synchronisation in the beta band

---

\*Speaker

was larger for negative words in L1 than L2, which we interpret as an index of greater emotional response to negative content in L1 (Güntekin & Başar, 2014).

Together, our findings align with results from language comprehension research in bilinguals, showing reduced sensitivity to negative L2 words. But our results extend this evidence across the board, to encompass spoken language comprehension and production. We also provide first insights into the neural dynamics of and mechanisms underpinning the production of written and spoken emotional words in bilinguals, getting one step closer to processes at work in everyday bilingual communication.

**Keywords:** Bilingualism, Reading, Translation, Emotion regulation, Event, related potentials, Event, related synchronization / desynchronisation

# What does it actually mean? How L1 speakers and advanced L2 speakers of Swedish process and interpret pragmatic particles

Pia Järnefelt \* <sup>1</sup>

<sup>1</sup> Uppsala Universitet [Uppsala] – Sweden

Pragmatic particles are used to express speaker attitudes and to create common ground between interlocutors. They are polyfunctional, highly frequent, serve fundamental communicative purposes and are notoriously hard for L2 speakers to acquire (Hogeweg et al., 2016). This study examines how advanced L2 Swedish speakers, compared to L1 Swedish speakers, process and interpret the Swedish particles *egentligen* (‘really’, ‘actually’) and *faktiskt* (‘really’, ‘in fact’) when they modify color words in everyday conversations. 42 advanced Swedish L2 speakers and 42 L1 speakers were tested in a novel eye tracking experiment within the Visual World Paradigm (Tanenhaus et al., 1995).

Participants were told beforehand that they would overhear everyday conversations while selecting which object the conversation was about, although the object was never mentioned by name. The use of particles modify the hearers’ interpretation, as it reveals the speakers’ attitudes towards the propositional content. For instance, upon hearing ”it’s (*egentligen/faktiskt*) black”, the particle *egentligen* could indicate that the object used to be black but it is not anymore (e.g. a gray shirt), whereas the particle *faktiskt* should only confirm the blackness of the object (e.g. a black shirt). The conversations were always about pieces of clothing or furniture, and each target item contained a white or black object, and a competitor of the same shape but in a shade of gray. The linguistic manipulation consisted of three words: a) *egentligen*, b) *faktiskt*, and c) *väldigt* (an unambiguous adverb meaning ‘very’, which functioned as a control base line). 36 target items and 48 filler items were used in the experiment.

Three different measures that have been shown to correlate with higher cognitive load were collected – pupil size (Tromp et al., 2016), accumulated gaze patterns and reaction times (van Bergen & Bosker, 2018). Behavioral response accuracy and reaction time data were also collected.

Preliminary results show that both groups react similarly to *faktiskt*, but differ qualitatively when processing and interpreting *egentligen*. A majority of the L1 speakers hesitate upon hearing *egentligen*. Their eyes wander between the gray and black object, and they slow down. They choose the gray object over the black object to a larger extent. Meanwhile, a majority of the L2 speakers choose the black object, and they do not look at both objects, suggesting that they do not know the adversative meaning of *egentligen*, and therefore only see it as confirming the color of the object. The study contributes to the body of knowledge about online processing of pragmatic information and the development of L2 pragmatics. It also sheds light on the interpretation of these particles, as it is the first empirical study on how they affect speakers of Swedish. Since *egentligen* belongs to the 500 most common words in Swedish, it is concerning

---

\*Speaker

that such a qualitative difference is detected in advanced L2 learners. Lacking comprehension of these particles increases the risks of misunderstandings in L1-L2 communication.

**Keywords:** eye tracking, pragmatic particles, incremental sentence processing, SLA, L2 pragmatics

# Effects of targeted perceptual training on L2 suprasegmental cue weighting strategies

Magdalena Kachlicka <sup>\*† 1</sup>, Ashley Symons <sup>2</sup>, Yaoyao Ruan <sup>3</sup>, Kazuya Saito <sup>4</sup>, Fred Dick <sup>4</sup>, Adam Tierney <sup>1</sup>

<sup>1</sup> Birkbeck, University of London – United Kingdom

<sup>2</sup> Royal Holloway – United Kingdom

<sup>3</sup> University of Oxford – United Kingdom

<sup>4</sup> University College, London – United Kingdom

Speaking a L2 is a highly desirable skill. However, only some learners achieve the desired level of proficiency, while others struggle to understand L2 speech. These difficulties might arise from listening patterns inherited from learners first language (L1). Attention-to-dimension theories of speech perception suggest that L1 experience biases listeners attention toward acoustic dimensions which are especially informative in that language (Francis et al., 2000; Holt et al., 2018), so a strategy optimal for any given L1 might not be useful for subsequent languages. If learning a new language requires redirecting attention to L2-relevant acoustic dimensions, it should be possible to quickly boost L2 learning by enhancing learners ability to rely on those dimensions.

Here, we tested this hypothesis with Mandarin learners of English who had recently moved to the UK (length of residence < 18 months). We developed a novel training paradigm to help participants adopt more native-like strategies by enhancing their ability to use target acoustic cues during categorization while ignoring distractor acoustic cues. Since Mandarin speakers tend to overweigh pitch in L2 speech perception and production (Zhang et al., 2008) and have trouble disengaging attention from it (Jasmin et al., 2021), the prosody training featured duration as the target cue and pitch as the distractor uninformative cue. We gradually introduced task-irrelevant pitch variation and decreased the size of the duration cue to increase the difficulty of the tasks throughout the training. We predicted that after prosody training the participants would weight pitch and duration in a more native-like manner, relative to a control training group. The vocabulary control training was matched in length and intensity to prosody training (6 days x 30 minutes of exercises daily) and involved non-auditory English vocabulary learning tasks. 60 learners completed the training (30 prosody, 30 vocabulary). Both training paradigms comprised trial-by-trial feedback. We measured participants selective attention to pitch and duration (detection of repetitions within the attended dimension; Symons et al., 2021), dimensional salience (EEG cortical tracking during passive listening; Symons et al., 2021), cue weighting during prosody categorization (Jasmin et al., 2021) and L2 prosody perception and production before and immediately after the training. Additionally, we measured attention and cue weighting after 6 months to assess any long-lasting effects on learners perceptual strategies. After prosody training, participants showed enhanced use of duration during phrase boundary categorization, where duration is the primary cue. However, these effects did not endure. Control training participants showed stronger reliance on pitch over time in categorizing lexical stress, suggesting that in the absence of alternatives, listeners tend to solidify the use of the strategy that works well for them. Our findings provide evidence of strong reliance on pitch among Mandarin learners and difficulties resisting default strategies in L2 speech perception.

---

\*Speaker

†Corresponding author: m.kachlicka@bbk.ac.uk

Nevertheless, we show that perceptual strategies that took a lifetime to develop can be adjusted with as little as three hours of training. These findings open a new direction for designing more targeted language training paradigms.

**Keywords:** perceptual training, L2 prosody, salience, attention, cue weighting

# Vocabulary Learning from Documentaries: Are all documentaries equal?

Kadir Kaderoglu <sup>\*† 1</sup>, Roger Gilabert <sup>1</sup>, Maria Del Mar Suárez <sup>1</sup>

<sup>1</sup> Universitat de Barcelona – Spain

L2 vocabulary studies have frequently used documentaries as their audiovisual input, possibly because they convey information clearly and have visual support. Rodgers (2018) investigated the audiovisual congruence of words and images in documentaries and narrative TV series. He found that images and aural forms of words co-occur significantly more in documentaries than in TV series, which, he argued, might impact vocabulary learning from these genres. Peters (2019) empirically confirmed his hypothesis since she found that words with visual support were three times more likely to be learned compared to those without support. Nonetheless, documentaries are shot, organized, and edited differently. Nichols (2017) proposed six modes, each of which has distinctive features that distinguish one from another. This proposal focuses on three of them – expository, observational, and participatory. Therefore, further research is needed to explore (1) whether some documentary modes offer higher word-image congruence than other documentary modes and (2) how these differences in visual support affect L2 vocabulary learning from different modes. Thus, the following research questions have been formulated:

- Do different documentary modes have the same degree of imagery support?
- In which documentary mode do participants learn more single words from captioned viewing? If there exist differences, how do the input-related factors like (a) visual support, (b) frequency of reoccurrence, (c) lexical coverage, (d) recency, and a learner-related factor – vocabulary size – mediate learning?

To answer the first research question, Rodgers' (2018) 5-second before and after measure, and Ahrabi Fakhri et al.'s (2021) imagery scale will be used to explore the visual support of the TWs across documentary modes. The reason to use the two measures is to explore which one is a better predictor of visual support. To answer the second research question, 100 intermediate-level Turkish learners of English at university will watch approximately 2 hours of each documentary mode – expository, observational, and participatory – in 6 weeks. A within-subjects design will be employed as it reduces individual differences and increases the power to identify possible differences between the documentary modes since all participants will view all in a counterbalanced order. Half of the TWs will be visually supported while the other half will not. Input-related factors like frequency of reoccurrence, lexical coverage, recency effect, and one learner-related factor, vocabulary size, will also be explored. Moreover, a control group will only take the vocabulary tests (form recognition and meaning recall). Additional potential exposure to the TWs, i.e., the coursebook, will be controlled.

---

\*Speaker

†Corresponding author: kkaderoglu@ub.edu

Overall, it is hypothesized that the expository mode will have the highest visual support for words as narration and the images tend to complement one another in this mode. However, this feature is not strictly followed in the other two modes. Consequently, participants might learn more from the expository mode than the other two modes since the additional support of imagery might offer extra assistance to process and retain vocabulary (see the Dual Coding Theory, Paivio, 1990; and the Cognitive Theory of Multimedia Learning, Mayer, 2014 for more on how multimodality helps learning).

**Keywords:** documentary, imagery, audiovisual input, vocabulary learning, multimedia learning



# Pedagogical construction grammar meets multilingual didactics: How to teach the construction (Verb of motion + a/de + noun) in Spanish classes at German grammar schools?

Eleni Kanli \* <sup>1</sup>

<sup>1</sup> University of Tübingen – Germany

The use of prepositions is one of the most difficult grammar topics for Spanish learners. The highly frequent prepositions *a* ‘to’ and *de* ‘of’ are often confused or omitted, especially in motion verb constructions (Campillos Llanos, 2014). This is due to their low salience and polysemous character (Wulff & Ellis, 2018) as well as crosslinguistic influence (e.g. Rothman, 2019; Höder et al., 2021).

How can Spanish teachers support their students in acquiring these prepositions so that they can express themselves idiomatically correctly when describing, for example, their daily routine?

The current state of research and a pilot study conducted at the beginning of the second year of learning (9th grade) at a German grammar school (Gymnasium) provide a hypothesis basis that it could be helpful to draw attention to *image schemata* (*applied cognitive linguistics* (Scheller, 2009)), to form-meaning pairs with varying degrees of schematicity (*pedagogical construction grammar* (Bürgel, 2021)) and to compare functionally equivalent constructions from different languages (*multilingual didactics* (Fäcke & Meißner, 2019)).

In this presentation, I would like to present and discuss the designed teaching materials for the three-hour unit *Planeamos un fin de semana en Madrid* (‘We are planning a weekend in Madrid’), where I combine current research results from the above-mentioned disciplines in an innovative way. These teaching materials are supposed to help students who learn English, French and Spanish successively at German grammar schools (Gymnasium) to acquire the prepositions *a* and *de* in the construction (*motion verb + a/de + noun*) in the first year of learning (8th grade). They will be empirically tested in 2024 and in 2025 in several Spanish classes using a pre/post-test design (Albert & Marx, 2016, 90–91) and *Stimulated Recall* (Gass & Mackey, 2016).

## Bibliography:

Albert, R. & Marx, N. (2016). *Empirisches Arbeiten in Linguistik und Sprachlehrforschung: Anleitung zu quantitativen Studien von der Planungsphase bis zum Forschungsbericht*. Narr Francke Attempto.

Bürgel, C. (2021). Lexiko-grammatische Konstruktionen als Kern der Spracharbeit im Spanischunterricht. *Hispanorama*, 4 (174), 10–17.

---

\*Speaker

Campillos Llanos, L. (2014). Las preposiciones en el habla no nativa de nivel intermedio: análisis de la interlengua basado en corpus. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, (16), 5–27.

Fäcke, C. & Meißner, F.-J. (eds.) (2019). *Handbuch Mehrsprachigkeits- und Mehrkulturalitätsdidaktik*. Narr Francke Attempto.

Gass, S. M. & Mackey, A. (2016). *Stimulated Recall Methodology in Applied Linguistics and L2 Research*. Routledge.

Höder, S., Prentice, J. & Tingsell, S. (2021). Additional language acquisition as emerging multilingualism: A Construction Grammar approach. In H. C. Boas & S. Höder (eds.), *Constructions in Contact 2: Language change, multilingual practices, and additional language acquisition* (p. 309–338). John Benjamins Publishing Company.

Rothman, J. (2019). *Third Language Acquisition and Linguistic Transfer*. Cambridge University Press.

Scheller, J. (2009). *Animationen in der Grammatikvermittlung: Multimedialer Spracherwerb am Beispiel von Wechselpräpositionen*. LIT.

Wulff, S. & Ellis, N. C. (2018). Usage-based approaches to second language acquisition. In D. Miller, F. Bayram, J. Rothman & L. Serratrice (eds.), *Bilingual Cognition and Language* (p. 37–56). John Benjamins Publishing Company.

**Keywords:** prepositions, intransitive motion constructions, pedagogical construction grammar, applied cognitive linguistics, multilingual didactics, usage, based approach, third or additional language acquisition, crosslinguistic influence

# Rating rubrics of EFL writing performance in published research articles: a systematic review

Huimin Ke <sup>\*† 1</sup>, May Wu <sup>1</sup>, Rasmus Steinkrauss <sup>1</sup>, Marjolijn Verspoor <sup>1</sup>

<sup>1</sup> University of Groningen – Netherlands

Using rating rubrics or scales to assess performance or proficiency is a common practice in L2 writing research. The aim of the current study is a systematic review of research articles to investigate (1) what rating rubrics for L2 writing are used in the field of L2 writing development, (2) what rating criteria (i.e. writing constructs) these rubrics have specified, and (3) how these constructs have been delineated. We further explored (4) whether there are any associations between rubrics, constructs, and the proficiency levels of the writers investigated in the studies. For the systematic review, we extracted 378 articles on EFL writing that were published in 38 high-ranking journals between 2000 and 2019. After excluding articles not providing sufficient details for analysis, 66 articles were kept. After tallying which rubrics were used in these 66 articles (RQ1), we mapped which constructs the rubrics used (RQ2). Taking the seven constructs formulated in Polio's (2001) methodological review (Content, Coherence & Discourse Features, Lexical Features, Syntactic Complexity, Linguistic Accuracy, Mechanics, and Fluency) as a starting point and Style as an additional construct, we categorised rubric segments according to these eight constructs, using ATLAS.ti for tagging. After exporting the tagged rubric segments from ATLAS.ti, we use a natural language processing software to explore the topics used to delineate each construct (RQ3). Lastly, we added proficiency level codes (A1 and A2 level as Beginner, B1 and B2 level as Intermediate, and C1 and C2 level as Advanced; based on CEFR) to the rubric segments and explored the connections and co-occurrence patterns among the codes for rubrics, constructs, and proficiency levels (RQ4).

Our findings show that researchers make use of both standardized rubrics (e.g. the Written Assessment Grid from CEFR) and customized ones to fit specific research contexts (Bachman & Adrian, 2022). In our review, standardized writing rubrics from TOEFL and Cambridge English language tests prevailed, but individually designed rubrics still occurred in about 30% of the studies. The three most frequently used constructs in rubrics across all three proficiency levels were Content, Coherence & Discourse Features, and Linguistic Accuracy. There were clear thematic clusters in the way the constructs were delineated. Taking Coherence & Discourse Features as an example, the rubrics allude to *organization*, *content*, *communication*, *support*, *ideas*, *relevance*, *task*, *details*, and *development* to articulate requirements to define this assessment criterion.

No association between rubrics and proficiency levels or between constructs and proficiency levels were found. However, we did see co-occurrence patterns among constructs. Content and Coherence & Discourse Features showed a strong co-occurrence pattern, and Linguistic Accuracy clustered closely with Lexical Features and Syntactic Complexity in how the rubrics defined rating criteria. Style on the other hand linked to most other constructs.

---

\*Speaker

†Corresponding author: h.ke@rug.nl

The findings contribute to a deeper understanding of constructs and criteria used in writing rubrics, and will be relevant for L2 writing researchers and teachers who need to decide on rubrics for their studies and teaching.

**Keywords:** EFL writing, rating rubrics, assessment criteria, writing construct, systematic review

# Redundancy and salience in initial L2 acquisition

Panagiotis Kenanidis \*<sup>1</sup>, Miquel Llompart<sup>1,2</sup>, Sara Fernández Santos<sup>1</sup>,  
Ewa Dabrowska<sup>1,3</sup>

<sup>1</sup> Friedrich-Alexander Universität Erlangen-Nürnberg = University of Erlangen-Nuremberg – Germany

<sup>2</sup> Universitat Pompeu Fabra [Barcelona] – Spain

<sup>3</sup> University of Birmingham [Birmingham] – United Kingdom

Grammatical redundancy is a widespread feature across languages. Although redundant cues can be seen to increase the complexity and processing burden of structures, their presence is also thought to confer several advantages to language users. One of these pertains to the learnability of languages (and constructions) (Monaghan, 2017; Tal & Arnon, 2022). In this study, we report two experiments that examined the effect of redundant linguistic cues at the very initial stages of second language (L2) acquisition in adult learners, as well as the extent to which this effect may be modulated by the level of salience of these cues. In Experiment 1, a group of adult native speakers of English ( $N=28$ ) were exposed to a novel artificial language that contained fixed word order, namely Verb-Object-Subject, while a second group of learners ( $N=30$ ) was presented with an identical language in which the subject and the object of the sentence were additionally marked by a redundant case marker (*-i* and *-o*, respectively) (**1a-b**). Participants were initially trained on the vocabulary of the novel language and subsequently completed a sentence comprehension test that was divided into four identical blocks. In this test, two videos, each showing two aliens performing an action, were presented, accompanied by a sentence corresponding to one of them. For trials assessing vocabulary learning, the videos differed in either one of the aliens or the action performed, while for those testing grammatical comprehension, the two videos differed in that the agent/patient roles were reversed. Although both groups managed to quickly learn the novel structure, the non-redundant group achieved significantly better grammar learning outcomes than the redundant group. To test whether this hindering effect of redundancy can be mitigated by increasing the salience of the redundant marker, in Experiment 2, following Tal & Arnon (2022), we exposed a new group of learners ( $N=30$ ) to a language that was exactly the same as in the non-redundant group but contained the highly salient redundant marker *-pazz* on objects (**1c**). The procedure was identical to Experiment 1. Once again, results failed to show a beneficial effect of redundancy, with participants demonstrating greater grammar learning effects compared to the redundant group of Experiment 1, but still smaller, though non-significantly, than the non-redundant group. In sum, the findings of this study highlight the hindering effect of redundancy in adult language acquisition, which can, however, vary depending on the salience of the redundant cues. Our findings also speak to the ongoing debate about the function of redundancy in languages, its impact on language learnability and processing (e.g., Linguistic Niche Hypothesis; Lopyan & Dale, 2010), as well as the factors that modulate its effect, which, as shown here, can include the learners' biases from their prior first language experience (English vs. Hebrew in Tal & Arnon, 2022) and the stage of L2 development that they are at.

---

\*Speaker

**Keywords:** redundancy, salience, artificial language learning, second language acquisition, grammatical comprehension

# Testing a phonological awareness test: A validated reduced scale predicts L1 and L2 receptive lexical and grammatical skills

Kristin Kersten <sup>\*† 1</sup>, Ann-Christin Bruhn <sup>\*</sup>

<sup>1</sup>, Lynn Jedamski<sup>‡ 2</sup>

<sup>1</sup> University of Mannheim – Germany

<sup>2</sup> University of Hildesheim – Germany

Phonological awareness (PA) is described as a metacognitive ability to recognize, analyze and manipulate the sound structure of spoken language (Mayer, 2008). It has been found in previous research to be highly predictive for L1 acquisition, in particular with regard to literacy skills (Schründer-Lenzen, 2013). For L2 proficiency, however, its predictive power is much less investigated and more controversial, with a stronger focus on L2 lexicon (Farnia & Geva, 2011; Hu, 2008) than on L2 grammar (Hopp et al, 2019).

A particular challenge is the operationalization of PA in language tests. On the one hand, different measures render comparability between studies difficult. On the other, specific types of measurements often seem to include aspects of literacy skills and do not always seem compatible with current phonological theory. In Germany, the PA test *Basiskompetenzen für Les-Rechtschreibleistungen* (BAKO (Basic competences for reading and spelling skills), Stock et al., 2017) is among the most widely used instruments. In our own linguistic analysis of the BAKO, however, we found only 17 out of 59 items to be linguistically valid with regard to German phonological and phonotactic patterns (Jedamski, 2023). The other items exhibited problems with transcriptions of the German phoneme inventory, use of graphemes rather than phoneme symbols, and tense vs. lax vowel quality. Other aspects included missing differentiations between strong and reduced syllables, an incorrect handling of schwa-elision, final devoicing, and vocalization of in syllable codas. Based on this linguistic analysis, we then created a BAKO short-scale comprising these 17 items. The present study uses this short-scale to compare the effects of PA on L1 and L2 acquisition, focusing on the following question:

*Does PA, operationalized with this short-scale, predict the lexical and grammatical receptive skills of L1 German and L2 English learners in fourth grade?*

The sample consisted of  $N = 113$  4th-graders (56 female) from German mainstream and bilingual primary schools. Data analysis of receptive skills was carried out using standardized instruments, i.e., the (L1-)WWT (Glück, 2011) and (L2-)BPVS-III (Dunn et al., 2009) for German and English lexical reception, respectively; and the (L1-)TROG-D (Fox, 2013) and the (L2-)ELIAS Grammar Test-II (Kersten et al., 2012) for German and English grammar reception.

Results showed that the internal consistency of the short-scale of 17 items was acceptable (Cronbach's alpha = .655). Results of four separate regression analyses revealed significant predictions

---

\*Speaker

†Corresponding author: kristin.kersten@uni-mannheim.de

‡Corresponding author: jedamskil@uni-hildesheim.de

of the PA short-scale on all linguistic L1 and L2 measures (WWT:  $\beta = .208$ ,  $p = .027$ ; BPVS-III:  $\beta = .304$ ,  $p = .001$ ; TROG-D:  $\beta = .393$ ,  $p < .001$ ; ELIAS-II:  $\beta = .346$ ,  $p < .001$ ). It is notable that, contrary to previous research (but with the exception of Hopp et al.'s study), predictive effect sizes are stronger for grammar comprehension than for vocabulary for both L1 and L2 skills. Results will be discussed with regard to the role of PA for L1 and L2 acquisition and to the significance of methodological precision in language testing.

Jedamski, L. (2023). *Assessing phonological awareness: Predicting L2 competence using a full and reduced BAKO 1-4 (Basic competences for reading and spelling skills)*. Unpublished Manuscript. Hildesheim University.

**Keywords:** phonological awareness, test validity, operationalization, young learners, L1 receptive skills, L2 receptive skills



# A Longitudinal Investigation into Second Language Writing Development: Examining Complexity, Accuracy, Fluency, and their Interactions

Barkaoui Khaled \* <sup>1</sup>

<sup>1</sup> York University – Canada

Despite calls for more research on the nature and causes of second language (L2) writing development, previous studies have mostly focused on task effects on L2 writing performance at a single point in time. These studies typically explore L2 writing performance in terms of linguistic complexity, accuracy, and fluency (referred to as CAF) (Johnson, 2017; Michel, 2017). These three dimensions play an important role in writing and can provide important insights about L2 performance and development.

From a development perspective, as L2 learners progress in their L2 competence, they should demonstrate increased complexity, accuracy, and fluency in their writing. However, L2 learners often encounter challenges in balancing these three dimensions at the same time. Generally, they may focus on one dimension at the expense of the others (Robinson, 2011; Skehan, 2016). Additionally, the three dimensions do not often progress in tandem because they "are inter-related in complex and not necessarily linear ways" (Michel, 2017, p. 53; Housen & Kuiken, 2009; Larsen-Freeman, 2009). Therefore, to understand L2 writing development, it is crucial to examine the degree to which the three CAF dimensions are in(ter)dependent in L2 performance and development and the factors influencing their manifestations, inter-relationships, and development in L2 writing (Housen & Kuiken, 2009).

To address these questions, we adopted a complexity theory perspective and a longitudinal approach to examine changes over time in multiple CAF measures, their inter-relationships, and their relationships to writing quality in the texts of 402 English language learners from different first language (L1) backgrounds and across five different levels of English language proficiency (ELP, B2 to A1 on the CEFR). Each learner wrote three essays (before, during and after a period of English language study), with 2-3 months between essays, resulting in a corpus of 1,206 essays. Each essay was computer-analysed in terms of six CAF indices: global syntactic complexity, noun-phrase complexity, lexical variation and sophistication, the ratio of lexical and grammatical errors per 100 words, and fluency (word count). Subsequently, multilevel modeling was used to statistically examine the rate of change in CAF measures over time and how the rate of changes in CAF measures varied with learner ELP, L1, and length of English study. Additionally, correlation analyses examined whether the inter-relationships between CAF measures and their relationships to essay holistic scores varied across time points and ELP levels.

The findings revealed that the essays of students with higher English language proficiency generally exhibited significantly higher fluency, accuracy, lexical complexity, and some aspects of syntactic complexity. Second, fluency and accuracy improved significantly, while syntactic complexity declined; lexical complexity did not exhibit significant changes over time. Third, the relationships between the six CAF measures varied across proficiency levels and time points. Fi-

---

\*Speaker

nally, fluency, accuracy, and certain indices of syntactic and lexical complexity were significantly associated with essay holistic scores. The findings, along with their implications for theory, instruction, assessment, and research on L2 writing development, will be discussed.

**Keywords:** Second Language Writing, longitudinal, Complexity, Accuracy, Fluency, L2 proficiency, L1

# A Closer Look at the Acquisition of the Present Progressive by Japanese-speaking Learners of English

Takayuki Kimura <sup>\*† 1</sup>, Runa Kohama<sup>‡ 2</sup>, Hikari Otsuka<sup>§ 2</sup>, Ryusho Osawa<sup>¶ 2</sup>, Urara Shimoyama<sup>|| 2</sup>, Seyoung Jeong<sup>\*\* 2</sup>

<sup>1</sup> Utsunomiya University – Japan

<sup>2</sup> Utsunomiya University – Japan

This study investigates the acquisition of the present progressive (*-ing*) by Japanese-speaking learners of English (JLEs), delving into the effect of semantics of verbs and the present progressive.

In English, the present progressive morphology *-ing*, typically affixed to activity/accomplishment verbs, denotes a process in progress (e.g., *The boy is painting a wall.*). It may also be attached to achievement verbs if the verb possesses a (+process) feature (e.g., *The train is arriving.*). In contrast, the present progressive form in Japanese, *-teiru*, has dual meanings: progressive and resultant state. While *V-teiru* is understood as a progressive event when the verb belongs to the activity/accomplishment class, it is construed as a resultant state when the verb belongs to the achievement class.

Building on these facts, Gabriele (2009) conducted an investigation into the interpretation of English *-ing* by JLEs through a picture-based truth-value judgment task (TVJT). The results revealed that JLEs at elementary and intermediate levels often misinterpreted V(ACHIEVEMENT)-*ing* as indicative of a resultant state, possibly influenced by their L1. However, Gabriele identified substantial variability among verbs, indicating a potential semantic effect on the use of *-ing* by L2ers.

In light of this background, our study aims to scrutinize the impact of semantic properties of verbs on the acquisition of the present progressive *-ing* by JLEs. To address this research question, we conducted two experiments: an acceptability judgment task (AJT) in Japanese and a truth-value judgment task (TVJT) in English. In the AJT, 41 native Japanese speakers evaluated the acceptability of phrases containing *kake* "be about to" in Japanese, which exclusively co-occurs with (+process) verbs (Kuno & Takami, 2017), within provided picture-based contexts. Based on the AJT results, we identified verbs compatible with *kake* (i.e., (+process) verbs) (e.g., *freeze, close*), and those incompatible (i.e., (–process) verbs) (e.g., *fall, burn*). Subsequently, the TVJT involved 70 native English speakers and 50 JLEs at elementary and intermediate proficiency levels, who judged the acceptability of sentences with *-ing* in provided picture-based contexts. The TVJT included i) control types with accomplishment verbs with

---

\*Speaker

†Corresponding author: tkmr32@gmail.com

‡Corresponding author: lun.beach@gmail.com

§Corresponding author: o27hikari@gmail.com

¶Corresponding author: ryu.telecustom@gmail.com

||Corresponding author: aomusi5450@gmail.com

\*\*Corresponding author: wjdehrtk@naver.com

*-ing* in progressive and resultative contexts, ii) achievement verbs ((+process)) with *-ing* in both contexts, iii) achievement verbs ((-process)) with *-ing* in those contexts, and distractors. The results from the TVJT revealed that, through the manipulation of verb semantics, even JLEs at non-proficiency levels accurately interpreted V(ACHIEVEMENT)(-process) *-ing* as indicative of a progressive rather than a resultant state ( $p < 0.05$ , logistic-regression analyses). This implies that the acquisition and use of the present progressive form are significantly influenced by verb semantics. In the presentation, we will discuss further implications of the results obtained in relation to the Aspect Hypothesis (e.g., Andersen & Shirai, 1996).

**Keywords:** present progressive, aspect, L1 transfer, semantics

# L3 Processing of Verbal Aspect in English by Russian-German Bilingual Children: Evidence from Eye-Tracking

Nadine Kolb <sup>1</sup>, Serge Minor <sup>\*† 2</sup>, Natalia Mitrofanova <sup>\* ‡ 2</sup>, Marit  
Westergaard <sup>2,3</sup>

<sup>1</sup> University of Stavanger – Norway

<sup>2</sup> Uit The Arctic University of Norway – Norway

<sup>3</sup> NTNU – Norway

A growing body of experimental studies has suggested that grammatical representations from both previously acquired languages are likely to influence the development of a third language (see Westergaard et al., 2023 for a review). Two recent studies, Jensen et al. (2023) and Kolb et al. (2022), investigated the acquisition of L3 English by bilingual Norwegian-Russian and German-Russian adolescents in comparison to age- and proficiency-matched monolinguals (Norwegian, German, and Russian). Both studies found that the L3 learners differed significantly from both L2 control groups, suggesting combined CLI from both previously acquired languages. The current study continues this line of research adding a processing dimension to L3 research.

We use Visual World eye-tracking to investigate processing of English aspectual forms by bilingual German-Russian children in comparison to age- and proficiency-matched monolinguals. The aspectual systems are dissimilar in the three languages involved in the study. In Russian, there is a strict grammatical opposition between perfective and imperfective verbs with speakers uniformly associating perfective forms with completed events, and imperfective forms with ongoing events (see Minor et al., 2023). On the other hand, German lacks verbal aspect and doesn't grammatically encode the contrast between completed and ongoing events in the past. Finally, English employs a specialized progressive form to refer to ongoing events, while the so-called simple past forms are equally compatible with ongoing and completed event interpretations. Ermolina (2022) adapted the eye-tracking paradigm developed in Minor et al. (2023) to investigate the processing of L2 English Past Progressive (PP) and Simple Past (SP) forms by speakers of L1 Russian and L1 Norwegian (note that Norwegian is similar to German in that it doesn't encode grammatical aspect). The results revealed that L1 Russians interpreted English SP forms as denoting completed events, while L1 Norwegians favored ongoing event interpretations. Both groups showed a similar preference for ongoing events when they heard PP forms. We adapted the paradigm from Minor et al. (2023) and tested whether bilingual Russian-German adolescents would be influenced by aspectual systems from both of their languages when processing L3 English. Data collection with 50 German-Russian bilingual adolescents (8-13 y.o., tested in Berlin) has been completed. The results indicate that when the learners heard PP forms, they preferred looking at pictures of ongoing events - more than when they were listening to SP forms. Data collection with L1 German controls (matched in age and English proficiency with the bilinguals) is ongoing and will be fully completed in February 2024 – with the results ready to be presented at the time of the conference.

---

\*Speaker

†Corresponding author: sergey.minor@uit.no

‡Corresponding author: natalia.mitrofanova@uit.no

**Westergaard et al. (2023)**. Full Transfer Potential in L3/Ln acquisition: Crosslinguistic influence as a property-by-property process. *CUP*. **Jensen et al. (2023)**. Crosslinguistic influence in L3 acquisition across linguistic modules, *IJM*. **Kolb et al. (2022)**. Crosslinguistic influence in child L3 English: An empirical study on Russian-German heritage bilinguals. *IJB*. **Minor et al. (2023)**. Aspect processing across languages: A visual world eye-tracking study. *Front.inLangSci*. **Ermolina, I. (2022)** Acquisition of Verbal Aspect in L2 English by advanced learners with L1 Russian and L1 Norwegian. *UiT*.

**Keywords:** grammatical aspect, third language acquisition, processing, English, eye, tracking

# Processing of computer-mediated feedback: An eye-tracking study

Nektaria Kourtali \* <sup>1</sup>, Ziwei Guo <sup>1</sup>

<sup>1</sup> University of Liverpool – United Kingdom

Drawing on perspectives from cognitive interactionism (Long, 2007), the benefits of interactional feedback in SLA are widely accepted. A type of feedback which has received attention by a considerable number of studies is reformulations of learners' errors (i.e., recasts). In the CALL environment, it could be argued that when recasts are provided during written synchronous computer-mediated communication (SCMC), they might benefit learners because they remain permanently on the screen and learners have more time to process them. Numerous studies have explored potential benefits of written SCMC recasts (e.g., Arroyo & Yilmaz, 2018; Kourtali & Borges, 2023; Sauro, 2009); however, only a handful of studies (e.g., Smith, 2012) have explored real-time processing of recasts during SCMC interaction.

To fill this gap, the aim of the current study is to delve into attentional processes of 28 university students who are L2 users of English (IELTS 6–6.5) when reading written SCMC recasts supplied during an online task. The study examines (1) the extent to which the linguistic focus of recasts (i.e., lexical, morphological and syntactic) influences learners' levels of attention, and (2) whether there are links between attentional processes and foreign language aptitude. The researchers used eye-tracking methodology to capture participants' eye movements and Weblink software to identify recasts (i.e., interest areas) and generate measures which show levels of attention, including number and length of fixations, duration of first fixation, and number of times learners returned to an interest area whereby correction was located. For triangulation, we also used stimulated recall (SR) interviews so as to gain information about learners' levels of noticing of target language features when reading recasts. Regarding foreign language aptitude, we used the LLAMA test designed by Meara and Rogers (2019).

After normalizing the recasts taking into account the number of letters, we used the Friedman non-parametric test and Bonferroni adjustment to examine the extent to which the linguistic focus of recasts affects attention to target features. The study showed that learners produced more and longer fixations when reading lexical recasts as opposed to morphological and syntactic ones. SR interviews support this finding as participants reported significantly higher levels of noticing when receiving lexical recasts as compared to morphosyntactic recasts. Regarding the role of aptitude, Spearman correlations showed a positive correlation between noticing of syntactic errors demonstrated by SR comments and language analytic ability measured by LLAMA F. Furthermore, a negative correlation was found between participants' phonetic coding measured by LLAMA D and number of fixations, number of times learners return to an interest area, and duration of first fixation when reading morphological recasts. Although replication of this finding is needed, this could tentatively demonstrate that learners with a low implicit aptitude score are more likely to engage in language analysis when grammatical features are corrected; however, more research is needed regarding tests of implicit and explicit aptitude. When discussing the findings of this project, the researcher will make suggestions for future directions in the area of SCMC research and pedagogical implications for language practitioners.

---

\*Speaker

**Keywords:** corrective feedback, computer, mediated communication, eye, tracking, language aptitude



# The construction of lexical networks in SLA: longitudinal data of word association and verbal fluency tasks in L1 and L2 of Syrian learners of French

Barbara Köpke \* <sup>1</sup>, Mireille Copin <sup>2</sup>, Ines Saddour <sup>2</sup>

<sup>1</sup> Laboratoire de NeuroPsychoLinguistique – Université Toulouse - Jean Jaurès – France

<sup>2</sup> Laboratoire de NeuroPsychoLinguistique – Université Toulouse - Jean Jaurès – France

Lexical acquisition goes well beyond vocabulary learning: learners need to build a semantic network to use their lexicon efficiently (Meara, 1996). This idea is not new: Richard (1976) and Nation (e.g., 2001) also proposed that the ability to associate words with each other is one of the facets of vocabulary learning. Meara has developed this idea further, focusing on the organization of the lexicon via connections between lexical items (e.g., Meara, 1996; 2009). Two tasks are particularly interesting for exploring semantic networks because they rely heavily on highly automated processes in lexical access: the free association task (whereby one word must be produced in response to another, e.g., *vacances-plage*) and the verbal fluency task (provide as many words as possible from a certain semantic category, e.g., fruits). While the former has been used in some work on vocabulary learning (see review in Fitzpatrick & Thwaites, 2020), the verbal fluency task, common to explore the semantic networks of pathological subjects, is only rarely used with L2 learners.

The present study is part of the CLASS project involving a longitudinal investigation of the linguistic and sociocultural development of Syrian refugees in France. We will present data of 29 Syrian learners of French (age 18-68 years) who spent from 1 month -10 years in France with variable proficiency. Such a broad spectrum of learners allows us to document the connections in the L2 lexicon both longitudinally and cross-sectionally taking into account L2 level at data-collection, age at arrival in France, length of immigration.

All participants were tested twice, at T1 and approximately 12 months later (T2). All tasks were performed first in French (L2) and then in Syrian Arabic (L1). For the word association task, they saw a word written on a card and read by the experimenter and had to produce the first word that came to their mind. The task involved 20 stimuli, 20 concrete and 10 abstract nouns that are generally familiar even for beginners. For the verbal fluency task, two semantic categories known to be particularly productive were used: animals and fruit.

For the word association task, responses were compared between L1 and L2 and to norms established with L1 French speakers (Ferrand & Alario, 1998; Ferrand, 2001). Preliminary analyses of T1 data indicated that performance was more variable in L2 compared to L1 and showed some strategies for compensation of weaker semantic links. For T2 we expect that these strategies should diminish, and so should the variability of the responses.

In the verbal fluency task we analysed the total number of words produced, as well as the

---

\*Speaker

number of clusters (i.e., sub-categories) and switching between categories reflecting more efficient semantic processing (Troyer, Moscovitch & Winocur, 1997). Preliminary results for T1 showed that performance was overall weaker for L2 than for L1. We expect that participants will manage this task more efficiently at T2 with larger clusters emerging. Overall, the study allows us to document the emergence of semantic networks in L2 learners in two tasks varying in difficulty.

**Keywords:** SLA, psycholinguistics, lexical access, language processing, semantic networks

# Acquisition of the mass-count distinction by Mandarin, and Cantonese L2 learners of English

Yılmaz Köylü \* <sup>1</sup>

<sup>1</sup> the Hong Kong University of Science and Technology – Hong Kong SAR China

According to Lima (2018), mass-count distinction is elusive as it is not exactly a straightforward process to determine whether languages make a morphosyntactic distinction in their grammar to represent nouns that are ontologically mass or count. Some languages, such as English, have grammaticized the mass-count distinction. In English, what distinguishes a count noun from a mass noun is its ability to combine directly with numerals (e.g., *four pancakes* versus *\*four honeys*), a feature that Chierchia (1998) refers to as the signature property, to be pluralized (e.g., *wolfs* versus *\*rains*), to be used with an indefinite article (e.g., *a cup* versus *\*a water*), and to be modified by *many* (e.g., *many theories* versus *\*many information/s*). A count noun is also differentiated from a mass noun by its inability to appear bare in the singular form (e.g., *\*car* versus *snow*), and its inability to be modified by the quantifier *much* (e.g., *\*much novel* versus *much pain*). Mass nouns in English are characterized by the opposite features. The current research is an extension and partial replication of Choi et al. (2018), aiming to contribute to the existing literature by investigating the effect of first language (L1) transfer and the semantic universal atomicity on the acquisition of the mass-count distinction by second language (L2) learners of English with Mandarin and Cantonese L1 backgrounds. The study explored whether Mandarin, and Cantonese L2 learners of English were on a par with native speakers of English in pluralizing canonically mass and count nouns that differed with respect to properties such as atomicity, abstractness and concreteness.

The participants were recruited from a major research university in Hong Kong. There were 54 L1 Mandarin, 17 L1 Cantonese L2 learners of English, as well as 15 native English speakers (NSs). The L1 Mandarin L2 learners of English were all postgraduate students studying mostly in engineering degrees. The L1 Cantonese L2 learners of English were all undergraduate students, most of whom were studying business administration. The NSs were all faculty members teaching language courses in the same institution.

The participants completed a Fill in the Gap Task (FGT). There were 6 categories of nouns (i.e., count concrete, count abstract, mass atomic concrete, mass atomic abstract, mass non-atomic concrete, mass non-atomic abstract) and 4 nouns for each category that was tested twice. A one-way multivariate analysis of variance (MANOVA) was conducted to see whether there was a statistically significant difference between NSs, and the L1 Mandarin and L1 Cantonese L2 learners of English in terms of their pluralization patterns in the six experimental categories. The results indicated that while L1 transfer did not affect L2 learners' acquisition of the mass-count distinction, the semantic universal atomicity (boundedness) led them to incorrectly pluralize mass/atomic nouns significantly more than mass/non-atomic nouns both for concrete and abstract nouns. Abstractness and concreteness only played a role in the domain of atomic mass nouns, leading the L2 learners to pluralize mass/atomic/concrete nouns signifi-

---

\*Speaker

cantly more than mass/atomic/abstract nouns, indicating that there is easier access to ‘atoms’ in mass/atomic/concrete nouns.

**Keywords:** second language acquisition, mass/count distinction, Mandarin, Cantonese

# Language change in Japanese-English bilingual returnee children over the course of five years: evidence from accent rating and acoustic analyses

Tim Laméris <sup>\*† 1</sup>, Maki Kubota <sup>2</sup>, Tanja Kupisch <sup>3</sup>, Jennifer Cabrelli <sup>4</sup>,  
Neal Snape <sup>5</sup>, Jason Rothman <sup>2</sup>

<sup>1</sup> Leiden University – Netherlands

<sup>2</sup> University of Tromsø – Norway

<sup>3</sup> Universität Konstanz – Germany

<sup>4</sup> University of Illinois [Chicago] – United States

<sup>5</sup> Gunma Prefectural Women’s University – Japan

Only a handful of studies have investigated global foreign accent (GFA) in bilingual children, and little is known on how it develops over time. Here, we examined the longitudinal development of GFA in bilingual *returnees*: children of immigrant families who spend their formative developmental years (school age) in an L2 majority language (ML) context but return to their L1 environment as older children or teenagers. During their stay abroad, they are exposed to the ML of the host country and typically acquire this as an early L2. At the same time, their L1 becomes a heritage language (HL) which they are only exposed to in the home environment. Upon return to their ‘home country’, this linguistic environment reverses: their L1 HL once again becomes the majority language, whereas the former L2 ML becomes a minority language. This yields two potential linguistic consequences: *heritage language reversal* and *L2 attrition*. We examined whether Japanese-English returnees exhibit signs of such HL reversal and L2 attrition in their speech over the course of five years.

We recorded 17 returnee children a few weeks after return to Japan (T1); one year after (T2); and five years after return (T3). Mean age at return was 10.02 ( $sd = 1.71$ ). Mean age of onset (AoO) to L2 English was 5.15 ( $sd = 2.59$ ) and mean exposure to L2 English (relative to L1 Japanese) whilst abroad was 0.48 ( $sd = 0.14$ ). Recordings were elicited narratives of a picture book from which we created 10-second samples. These samples, in addition to 17 ‘baseline’ samples of Japanese and English monolingual children, were used in two online accent-rating tasks, in which native speakers of American English and Japanese (each  $n = 45$ , and familiar with child speech) rated the degree of GFA of the samples on a 9-point Likert scale. If raters indicated a ‘8’ or ‘9’ (‘very strong foreign accent’), they were additionally asked to indicate what features contributed to their perception of a GFA.

A Bayesian model investigated the effect of language, time, and two experiential factors of interest (AoO to L2 English and exposure to L2 English) on accent rating. The model suggested that GFA in L1 Japanese decreased from T1 to T2 and continued to do so from T2 to T3. By contrast, GFA in L2 English increased from T2 to T3. The model also suggested that individuals with a later AoO to L2 English had stronger GFA in English but weaker GFA in Japanese. Individuals with more exposure to L2 English had weaker GFA in English but stronger GFA in Japanese.

---

\*Speaker

†Corresponding author: t.j.lameris@hum.leidenuniv.nl

We also found that English raters indicated that both segmental and suprasegmental features contributed to a perceived strong GFA, whereas Japanese raters primarily attributed suprasegmental features (intonation and rhythm) to contribute to GFA rating . An acoustic analysis into the longitudinal changes in vowel formants, stop VOTs, rhythmic and pitch-related measures, and how these may correlate with GFA language-dependently is currently underway.

**Keywords:** global accent, returnee bilingualism, japanese, attrition, individual difference

# Is productive lexical knowledge the same as lexical use? The case of advanced L1 Swedish EFL learners

Batia Laufer <sup>\*†</sup> <sup>1</sup>, Per Snoder <sup>\*</sup>

2

<sup>1</sup> University of Haifa – Dept. of English Language and Literature, University of Haifa, Haifa 31905, Israel

<sup>2</sup> Stockholm University – Sweden

A long-established distinction in vocabulary research is between receptive knowledge associated with listening and reading and productive knowledge associated with speaking and writing (Webb & Nation, 2017). However, the distinction between productive knowledge and use is less researched. Laufer (2005) regards the two as different constructs since retrieving a word when prompted by a test is not the same as using it in speech or writing because, in addition to knowledge, word choice is determined by confidence, lexical preferences, speed of access, and avoidance strategies. There is only a handful of studies exploring the relationship between productive knowledge and use of single words (e.g., Danelund & Henriksen, 2015; Levitzky-Aviad & Laufer, 2013) and only one we are aware of (Mizrahi, 2016) comparing the productive knowledge and use of collocations.

Our study examines productive knowledge and use of single words and collocations of 60 high-school advanced EFL learners, L1 speakers of Swedish, by comparing them to 52 native English speakers. We asked whether advanced learners performed similarly to native speakers on two measures of knowledge (productive vocabulary size of single words and collocations) and on four measures of use in essays (the number of infrequent word tokens and types, lexical variation as the proportion of different words, and the number of correct collocations).

Productive vocabulary size and collocation knowledge were measured by self-constructed and validated frequency-based tests of form recall of single words and collocations. The number of infrequent word tokens, types and lexical variation in essays were measured by analyzing participants' essays (300 words each) using the software Vocabulary Profile (Cobb, n.d.). The number of correct collocations was determined by MI score in COCA and cross-checked with a learner dictionary of collocations.

We counted the number of learners who reached a native-like level on each measure, defined as the mean of native speakers' score -1 standard deviation on each of the tests. With regard to vocabulary use in the essays, 30% of learners were native-like on three measures of lexical richness: the number of infrequent tokens, types, and variation. 65% of learners reached a native-like level on two measures: infrequent tokens and types, and 85% on lexical variation. In the use of collocations, close to zero learners reached native-like levels. As for vocabulary knowledge, the number of learners who scored within the native-speaker range on knowledge

---

\*Speaker

†Corresponding author: batialau@research.haifa.ac.il

tests was negligible, both for single words and collocations. These results seem to run counter to previous claims that vocabulary use develops more slowly than knowledge. We will attempt to resolve this apparent contradiction by relating our results and the results of other studies to different developmental routes of productive knowledge and use.

**Keywords:** lexical competence, SLA



# The more proficient the learners, the less sophisticated their L2 vocabulary? Reliability and validity issues in frequency-based lexical measures

Elen Le Foll <sup>\*†</sup> <sup>1</sup>, Raffaella Bottini <sup>\*</sup>

2

<sup>1</sup> Universität zu Köln = University of Cologne – Germany

<sup>2</sup> Lancaster University – United Kingdom

In SLA research, lexical sophistication measures are frequently used to analyse developmental patterns in learner language. Studies have mainly focused on written production in L2 English, analysing the relationship with L2 proficiency. Studies on L2 writing generally show a negative correlation between mean-frequency scores of content words and learners' proficiency (e.g., Kim et al., 2018; Monteiro et al., 2020), indicating that advanced learners produce more infrequent content words than learners at lower proficiency levels. In L2 speech (e.g., Crossley & McNamara, 2013), however, some studies report the opposite trend: the more proficient the learners, the more frequent (therefore less "sophisticated") their vocabulary. In this study, we investigate methodological factors that may explain these seemingly contradictory developmental patterns.

We measured lexical sophistication in the Spoken Monologues and the Written Essays of the International Corpus Network of Asian Learners of English (ICNALE). This corpus contains language produced by L1 and L2 English users (CEFR levels A2 to B2) in controlled settings. We used mean-frequency metrics to compute the average frequency value of each text by extracting a frequency score for each of its words using a reference corpus: the lower the mean-frequency value of a text, the more infrequent its vocabulary. For each learner text, we computed six lexical scores using six reference corpora representing different language modes and registers. We then fit mixed-effects regression models to investigate how lexical scores vary depending on 1) users' proficiency levels, 2) task topic, 3) mode and register of the reference corpus.

We found that for both the spoken monologues and the written essays, mean-frequency scores of content words increase with proficiency. The scores are also affected by task topic. However, the mode and register of the reference corpus describe by far the most variance in CW scores. The direction of the observed proficiency effect would typically be interpreted as more proficient learners using less sophisticated vocabulary. However, we believe that mean-frequency metrics measure register typicality or appropriateness rather than sophistication. For example, proficient learners' speech contains content words that are more frequent in the spoken reference corpora, therefore indicating their ability to produce vocabulary which is typical of the target mode and register.

---

\*Speaker

†Corresponding author: elefoll@uni-koeln.de

These findings have implications regarding the concept and operationalisation of lexical sophistication. They raise questions on the reliability and validity of such measures, concerns for the replicability and comparability of SLA studies that rely on them, and have practical implications for language testing.

### References

Crossley, S. & McNamara, D. (2013). Applications of Text Analysis Tools for Spoken Response Grading. *Language Learning & Technology*, 17(2), 171–192.

Kim, M., Crossley, S., & Kyle, K. (2018). Lexical Sophistication as a Multidimensional Phenomenon: Relations to Second Language Lexical Proficiency, Development, and Writing Quality. *The Modern Language Journal*, 102(1), 120–141.

Monteiro, K., Crossley, S. & Kyle, K. (2020). In Search of New Benchmarks: Using L2 Lexical Frequency and Contextual Diversity Indices to Assess Second Language Writing. *Applied Linguistics*, 41(2), 280–300.

**Keywords:** lexical complexity, lexical sophistication, L2 proficiency, register appropriateness, learner corpora, vocabulary acquisition, language testing, corpus linguistics

# Word engagement and incidental vocabulary learning: the impacts of task coomplexity and task condition

Meiqiong Liang <sup>\*† 1</sup>, Yuanhua Xie <sup>2</sup>

<sup>1</sup> Guangdong University of Foreign Studies – China

<sup>2</sup> Guangdong University of Foreign Studies – China

Two main approaches to vocabulary learning have been extensively investigated: intentional and incidental (Ellis, 2015; Webb, 2019). Notably, the Involvement Load Hypothesis (ILH) postulates that retention of new vocabulary in L2 depends on the degree of involvement load (IL) (Laufer & Hulstijn, 2001). To specifically draw on the influence of task complexity on language learning, Robinson (2001, 2007, 2011, 2022) proposed the Cognition Hypothesis (CH), arguing that cognitively more complex tasks prompt "acquisition-rich interaction (Ellis et al., 2020, p.41)". Relevant research on vocabulary learning has manipulated task complexity along the resource-dispersing dimension (Jung, 2016; Peters, 2006; Wu et al., 2013), of which two studies examined intentional learning of lexical items (Peters, 2006; Wu et al., 2013). To date, no research has been done on the impact of task complexity along the resource-directing dimension on incidental vocabulary learning. Although some adverse or no impacts were reported (Nuevo, 2006; Révész et al., 2014; Torres, 2018), the majority of pertinent studies have found beneficial influences of task complexity on interaction and language learning (Baralt, 2010, 2013; Gilabert, 2007; Gilabert et al., 2009; Kim, 2009; Kim & Taguchi, 2016; Kim & Tracy-Ventura, 2011; Révész, 2009; Révész et al., 2011; Solon et al., 2016). These inconclusive results might be attributed to differences between intentional and incidental learning.

Additionally, the interactional impacts of task complexity and other factors have been underlined (Ellis et al., 2020) but underexplored. Robinson and Gilabert (2007) argued that more interactive complex tasks elicit greater attention to input and uptake of linguistic forms. A few studies have verified this prediction (Ghaderi et al., 2022; Kang & Lee, 2019; Michel, 2011, 2013; Michel et al., 2007, 2012). Besides, some research has examined the impact of task complexity on both interaction and language learning (Baralt, 2010; Kim, 2009; Nuevo, 2006; Solon et al., 2016), but only Baralt (2010) explored their association. Also, several studies on vocabulary learning (Borer, 2007; Newton, 2013; Niu & Helms-Park, 2014) implied that when learners deployed more word engagement during task performance, they attained better lexical learning. However, it remained unknown how word engagement might be influenced by task complexity and task condition.

Therefore, this study examined the combined impacts of task complexity ( $\pm$  few elements) and task condition ( $\pm$  few participants) on Chinese EFL learners' word engagement and incidental vocabulary learning. Forty-four participants were assigned into 2 groups (i.e., individual and collaborative) to complete 2 decision-making tasks of varying task complexity. Nine minutes were allocated for silent reading and 20 minutes for reasoning independently or collaboratively. Learners' oral production was analyzed by the frequency of depth of word engagement. The results indicated that task complexity positively influenced word engagement and incidental lexical learning. Paired performers outperformed the solitary ones on aspects of vocabulary learning.

---

\*Speaker

†Corresponding author: 20230120016@gdufs.edu.cn

A marginal interactional effect between task complexity and task condition was detected in the immediate vocabulary learning, and positive correlations were discovered between word engagement and incidental vocabulary learning. These findings provided novel insights into incidental vocabulary learning and proposed some pedagogical implications.

**Keywords:** task complexity, task condition, incidental vocabulary learning, word engagement

# Visual cues and L2 speaker intelligibility: How do religious and non-religious face coverings affect listeners' understanding of L2 French?

Rachael Lindberg <sup>\*† 1</sup>, Pavel Trofimovich <sup>1</sup>, Sara Kennedy <sup>1</sup>, Oguzhan Tekin <sup>1</sup>

<sup>1</sup> Concordia University [Montreal] – Canada

Listeners' understanding of second language (L2) speech can be affected by their attitudes toward certain speaker characteristics, such as the speaker's ethnicity or language background (Babel & Mellesmoen, 2019; Lindemann, 2005), but it is presently unclear how visual cues indicating a speaker's religious affiliation might impact listener understanding of L2 speech. In Québec, the context of our research, provincial law currently prohibits certain public-sector employees from wearing religious head/face coverings (Bill 96, 2019). Our primary goal was therefore to examine the extent to which French listeners' understanding of L2 French speech is affected by a speaker's wearing of religious (i.e., niqab, hijab) versus nonreligious (i.e., medical mask) face coverings. Our secondary goal was to determine whether listeners' understanding of L2 speech in these visual conditions is associated with their attitudes toward immigrants, cultural values, and religious practices in Québec.

We first recorded four first language (L1) Arabic female speakers (20–28 years old), all intermediate learners of L2 French, reading 40 French sentences (5–8 words each) from a phonetically balanced standardized instrument for French Canadians (Legacé et al., 2010). During audio recording, all speakers wore a 3-ply medical mask. We then recruited 104 L1 Québec French speakers (aged 20–68), all born and raised in Québec, to evaluate these recordings, presented to them under four visual conditions (10 sentences per condition) with a static image depicting a young Arabic-looking woman: (a) not wearing any head covering, (b) wearing a 3-ply medical mask, (c) wearing a hijab (headscarf), and (d) wearing a niqab. The listeners first transcribed the sentence they heard, then rated how comprehensible (i.e., easy or difficult to understand) each utterance was on a 100-point sliding scale. They also filled out questionnaires targeting their attitudes toward immigrants and the importance of immigrants adopting Québec values and speaking French well, as well as their attitudes toward Québec secularism and the wearing of religious symbols in the public domain.

An intelligibility measure was derived from the accuracy of the transcribed sentences (correct words/total words), where mean scores were computed across all utterances, yielding four scores per listener (one per visual condition). Mean comprehensibility ratings were similarly computed across all utterances per listener, separately in each visual condition. A repeated-measures ANOVA revealed that the utterances heard in the medical mask condition were significantly more intelligible than the utterances heard in the other three conditions ( $ps < .001$ , Cohen  $ds = 0.38$ – $0.63$ ) and significantly more comprehensible than those heard in the hijab ( $p = .048$ ,  $d = 0.27$ ) and niqab ( $p < .001$ ,  $d = 0.45$ ) conditions. In addition, Pearson correlations showed that the listeners who believed more strongly that it was okay for immigrants not to speak French

---

\*Speaker

†Corresponding author: rachael.lindberg@concordia.ca

well while living in Québec showed weaker intelligibility scores for utterances in the headscarf condition ( $r = -.25$ ), although this relationship was weak (Plonsky & Oswald, 2014). Implications regarding the role of visual cues to a speaker's religious identity in listeners' understanding of L2 speech will be discussed.

**Keywords:** L2 French, reverse linguistic stereotyping, comprehensibility, intelligibility, Québec

# Investigating cross-linguistic syntactic priming effects in comprehension and production among Chinese learners of English

Christina Liu <sup>\*†</sup> <sup>1</sup>

<sup>1</sup> University of York – United Kingdom

One of the long-standing issues that psycholinguists and bilingualism researchers are interested in lies in how the native and second languages interact with each other at the cognitive levels, and how are these two languages stored in bilingual learners' brains during comprehension and production. The question of whether the syntactic representations are shared or separated in bilingual speakers' minds remains to be debated. 'Syntactic priming', refers to the phenomenon where language users are more inclined to reuse the same or similar structures that they have recently encountered, rather than using the alternative sentence structures in their subsequent language comprehension and production (Bock, 1986; Pickering & Branigan, 1998). As a well-accepted technique, it has long been an effective method to explore the nature of syntactic representations in bilingual studies.

The priming effect has been discovered in both within-language and cross-linguistic settings. It has been observed across various structures and languages, in comprehension and production, and in naturalistic and controlled experimental settings (Chen, Jia, Wang, Dunlap & Shin, 2013; Chen, Jiang & Zhou, 2017; Hartsuiker & Bernolet, 2017; Hwang, Shin & Hartsuiker, 2018). However, research comparing the evidence of cross-linguistic syntactic priming in comprehension and production is still underdeveloped. The current study investigates and compares cross-linguistic syntactic priming effects (from L1 to L2 and from L2 to L1) during the comprehension and production of dative alternations (prepositional objective, double objective) and transitive structures (passive and active sentences) among 100 Chinese learners of English. Meanwhile, whether cross-linguistic syntactic priming is affected by L2 proficiency and lexical boost effects will also be investigated.

The syntactic priming effect in comprehension will be tested by self-paced reading tasks, and the priming effect in production will be investigated by picture description tasks. Multiple linear regression models might be used to analyse the quantitative data and test the priming effect. It is predicted that syntactic priming effect will occur in both comprehension and production. L2 proficiency and lexical boost effect might modulate priming effect, meaning that, priming effect might be more obvious for advanced L2 Chinese learners of English, and might be stronger with lexical boost between prime and target sentences.

**Keywords:** cross, linguistic syntactic priming, L2 proficiency, comprehension and production, lexical

---

\*Speaker

†Corresponding author: xxl504@york.ac.uk

boost effect



# Design, construction and data processing of a longitudinal multimodal interlanguage corpus: An ongoing study with L1 Italian and L2 Chinese

Siyuan Liu <sup>\*† 1</sup>, Adriano Ferraresi<sup>‡ 2</sup>, Maja Milicevic<sup>§ 2</sup>

<sup>1</sup> University Of Bologna (UNIBO) – Italy

<sup>2</sup> University Of Bologna – Italy

Chinese interlanguage corpora have greatly promoted the related research on Teaching Chinese as a Foreign Language (TCFL). In recent years, more and more scholars have joined in the construction of Chinese interlanguage corpora, gradually building a base for teaching and research (Cui & Zhang, 2011; Lee et al, 2018; Li, 2021; Iurato, 2022). There are currently 17 interlanguage corpora of L2 Chinese, but there are still some shortcomings with respect to teaching applications.

This study focuses on a novel corpus, MICICL (a Multimodal Interlanguage Corpus for Italian Chinese Learners), which is an ongoing corpus project. MICICL incorporates data in three modalities: picture, audio, and text. Additionally, as a longitudinal corpus, MICICL tracks each participant, providing insights into their learning progress.

In this workshop paper I outline the design principles, data collection processes, and data processing and analysis approaches of MICICL:

1. Design principles: Addressing corpus framework issues, this study explores how to design a corpus that better aligns with TCFL, integrates well with Data-Driven Learning, and maximizes its value in language education/acquisition. In terms of content, it will address questions that are pivotal in the early research and design phases, such as how to design tasks for participants with different proficiency levels and create tasks that best reflect the learners' language mastery and their learning progress.
2. Data collection process: Data collection primarily occurs in a university setting, employing both online and offline methods. This section discusses methodological aspects of data collection, including how to control variables that may affect the quality of the linguistic data produced to the maximum extent and guide students to express their true proficiency.
3. Data processing and analysis approaches: we employ distinct processing methods for different modalities. The challenges in designing transcription/annotation methods for different modalities, especially dealing with a large number of incorrectly written Chinese characters produced by learners, are addressed. The processing of incorrectly written characters, which cannot be directly typed or stored/displayed in universal computer characters, poses a significant challenge. Determining error annotation standards for various modalities, defining granularity

---

\*Speaker

†Corresponding author: siyuan.liu4@unibo.it

‡Corresponding author: adriano.ferraresi@unibo.it

§Corresponding author: maja.milicevic2@unibo.it

and versatility of error annotation, and striking a balance between redundant and insufficient annotation are additional challenges faced by this research.

**Keywords:** L2 Chinese, multimodal interlanguage corpus, corpus design and construction, Chinese character processing

# Reassessing the gap between sound and word recognition in late L2 learners through highly symmetrical experimental tasks

Miquel Llompart \*<sup>1,2</sup>, Celia Gorba<sup>1,3</sup>, Pilar Prieto<sup>1,4</sup>

<sup>1</sup> Universitat Pompeu Fabra – Spain

<sup>2</sup> Friedrich-Alexander University Erlangen/Nuremberg – Germany

<sup>3</sup> Universitat Autònoma de Barcelona – Spain

<sup>4</sup> Institució Catalana de Recerca i Estudis Avançats – Spain

Learning a second language (L2) involves acquiring sounds that are not part of the native inventory. Especially problematic are cases in which two contrastive sounds in the L2 are perceptually mapped onto the same native language (L1) category, like with English /i/ and // for L1-Spanish speakers (1). Crucially, in order for learners to be able to use such phonological contrasts in communication, they not only need to learn to identify the sounds as two separate phonetic categories, but also to reliably assign these categories to the sets of L2 words that contain them.

A recurrent finding has been that the connection between sound identification and word recognition when the latter depends on these sounds is rather weak. Studies have shown large gaps in learner performances in favor of sound identification (2, 3), and correlations between individual values are rather inconsistent (2, 4). Nonetheless, in previous studies, there were substantial differences between the tasks and the sets of stimuli employed, which could have led to an artificial increase in outcome divergence. Considering recent calls for deeper methodological reflection in L2 speech research (5, 6), the present study revisits the relationship between sound and word recognition while exerting stricter experimental control.

Seventy-nine L1-Catalan/Spanish intermediate-to-advanced English learners took part in a perceptual vowel identification task and a word-picture matching task (WPM) in which the stimuli were the same set of English monosyllabic minimal pairs with the vowels // and /æ/ (easy to distinguish (7)), and /i/ and // (difficult to distinguish (1)). In the identification task, participants were presented with one word at a time and had to select the vowel they heard. In the WPM task, a word was presented first and then an image appeared on the screen, which could match the word (e.g., *heat*-HEAT), correspond to the other member of the pair (e.g., *heat*-HIT) or depict a minimally-different word with another vowel (e.g., *heat*-HAT). Participants had to respond whether the word and the picture matched. Mixed-effects modeling revealed that learners were substantially more accurate in the identification of /i/ and // (75% correct,  $SD = 20$ ) than in recognizing mismatches in WPM when these sounds were involved (53%,  $SD = 31$ ). No difference between the tasks was observed for //-/æ/ (identification: 90%,  $SD = 11$ ; WPM: 86%,  $SD = 15$ ). In addition, while identification predicted mismatch detection in WPM for /i/-//, the correlation between the two was only moderate in size ( $r(77) = .38$ ,  $p < .001$ ), which is in agreement with previous work (4, 8).

This study replicates prior findings making use of a tightly controlled experimental design, providing thus additional support to the claims that the ability of using challenging L2 contrasts to recognize spoken words lags behind that of identifying the sounds in question in phonetic

---

\*Speaker

tasks (2, 9), and that accuracy in the latter is not a guarantee for improvements in the former (9, 10). This emphasizes the need for more individual differences research assessing the role of other potential predictors of word recognition ability.

**Keywords:** speech perception, L2 speech, word recognition, L2 phonology, processing level, late L2 learners, vowels

# Multilingual turn in EAL education: To what extent do EAL teachers practice multiliteracy instruction?

Eliane Lorenz <sup>\*† 1</sup>, Anna Krulatz <sup>\* ‡ 2</sup>, Eivind N. Torgersen <sup>\* § 2</sup>

<sup>1</sup> Justus-Liebig-Universität Gießen = Justus Liebig University – Germany

<sup>2</sup> Norwegian University of Science and Technology – Norway

As English as an additional language (EAL) classrooms worldwide become increasingly multilingual, and as researchers and educators transition from monolingual to multilingual pedagogies, the traditional notion of literacy as the ability to read and write in the language of schooling is expanding to include a range of skills multilingual learners need to interact with the world around them (Flores & Aneja, 2017). Consequently, there is a need to implement (multi)literacy practices that allow all learners to not only draw upon but also further develop literacy skills in all languages in their linguistic repertoires (Christison & Murray, 2020; Hornberger, 2013; Martin-Jones & Jones, 2000). However, few teacher education programs include an explicit focus on the development of multiliteracy skills ((Lypka & DeFelice, 2020), and research on literacy practices in educational settings that serve multilingual learners are scarce. The present study set out to address this gap. The study was guided by two research questions (RQs): RQ1: Do primary school EAL teachers employ multiliteracy practices and how frequently are these employed? RQ2: Which types of multiliteracy practices are employed by the primary EAL teachers? Twenty-four EAL classes taught by six teachers working at a highly linguistically diverse, primary school in Norway were observed. An observation protocol with specific features of multiliteracy practices scored on a scale from 0 to 2 (0 – not observed; 1 – observed once; 2 – observed multiple times) was used to record observation scores and qualitative observation notes. The findings suggest that literacy practices that draw on learners’ multilingual resources were implemented to a minimal degree ( $M=0.21$ ). The teachers encouraged active use of literacy skills in different languages to some extent ( $M=0.38$ ), but a multiliteracy environment was rarely supported ( $M=0.19$ ), and there was minimal encouragement for students and parents to engage in literacy practices in home languages ( $M=0.06$ ). Overall, few types of multiliteracy practices were observed, with main literacy support provided with references to Norwegian and almost no references to students’ other languages. Implications for teacher education and professional development are discussed.

## References

Christison, M. A., & Murray, D. E. (2020). An overview of multilingual learners’ needs for the 21st century. In G. Neokleous, A. Krulatz, & R. Farrelly (Eds.), *Handbook of research on cultivating literacy in diverse and multilingual classrooms* (pp. 1-21). IGI Global.

Flores, N., & Aneja, G. (2017). "Why needs hiding?" Translingual (re)orientations in TESOL teacher education. *Research in the Teaching of English*, 51(4), 441–463.

---

\*Speaker

†Corresponding author: eliane.lorenz@anglistik.uni-giessen.de

‡Corresponding author: anna.m.krulatz@ntnu.no

§Corresponding author: eivind.torgersen@ntnu.no

Hornberger, N. (2013). Bilingual literacy. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Blackwell. Retrieved from <https://onlinelibrary.wiley.com/doi/epdf/10.1002/9781405198431.wbeal>

Lypka, A. E., & De Felice, D. (2020). Telecollaborative storytelling: Reframing English language learners' and pre-service teachers' identity, multimodal literacy, and intercultural competency. In G. Neokleous, A. Krulatz, & R. Farrelly (Eds.), *Handbook of research on cultivating literacy in diverse and multilingual classrooms* (pp. 146-165). IGI Global.

Martin-Jones, M., & Jones, K. (Eds.). (2000). *Multilingual literacies: Reading and writing different worlds*. John Benjamins.

**Keywords:** classroom observation, diverse learners, literacy practices, multilingual classrooms, multiliteracy, young EAL learners

# Predicting upcoming speech based on semantic and/or gender cues among L1 and L2 users of German: a visual-world eye-tracking experiment

Pernelle Lorette \* <sup>1</sup>

<sup>1</sup> University of Mannheim – Germany

Actively anticipating upcoming information is an important mechanism underlying language comprehension. However, prediction abilities appear to be more restricted in L2. van Bergen & Flecken (2017) found that only those L2ers whose L1 lexically encodes the semantic contrast *put.LIE* versus *put.STAND* used this semantic contrast as a cue to predict in their L2. Moreover, Hopp (2015) found that L2ers used semantic cues to form prediction but not morphosyntactic ones (i.e., case marking). Crucially, case marking was absent from his participants' L1 while the semantic cues were transferable from their L1. It is thus unclear whether L2ers' more readily rely on semantics or on possible transfers from the L1 to enable prediction. This study investigates whether semantic cues which cannot be transferred from the L1 still lead to stronger predictive behaviour compared to morphosyntactic ones which are available in the L1.

In a visual-world eye-tracking experiment, 21 German L1ers and 19 L1-French/L2-German bilinguals heard 96 German SVO sentences, while seeing a visual display containing four images: one image depicting the object of the sentence and three distractors / competitors (see Figure 1). Depending on the condition (manipulated within-subject, 24 items per condition), the sentence contained either no semantic or gender cue to enable prediction (no-prediction condition), or the object of the sentence could be predicted based on the semantics of the verb (semantic condition), the gender marking of the article (gender condition), or the combination of the semantics of the verb with the gender marking of the article (semantic+gender condition). In the semantic condition, half of the items involved the position verbs *stellen*(put.LIE) or *legen*(put.STAND) (this semantic contrast is not encoded in French), while the other half involved "non-position" verbs, which have a translation equivalent in French (e.g., *essen*(eat)).

Mixed-effects regression models on the proportion of anticipatory looks to the target object showed that both L1ers and L2ers used the semantics of non-position verbs to form predictions. Gender marking and the semantics of position verbs were only used by L1ers and led to delayed and lower proportions of predictive looks to the target compared to the semantics of non-position verbs. Overall, the results confirm a more restricted use of cues to predict in L2. The data do not provide evidence that L2ers use morphosyntactic cues that are present in the L1 more so than semantic cues that are not transferable from the L1. In fact, these two types of cues were not used by our L2ers. They were only able to predict based on semantic cues that are transferable from their L1. Finally, among L1ers, the results highlight weaker predictions based on morphosyntax compared to semantics. This may be accounted for by the good-enough approach to language comprehension (Ferreira et al., 2022) and opens new questions for both L1 and L2 processing research.

---

\*Speaker

**Keywords:** language prediction, eye, tracking, visual world paradigm, bilinguals, semantic cues, morphosyntactic cues



# Monolingual vs. multilingual university settings: an investigation of students' socio-economic, language and motivational profiles

Laetitia Louis <sup>\*† 1</sup>, Pauline Degrave <sup>1</sup>, Philippe Hiligsmann <sup>1</sup>

<sup>1</sup> UCLouvain – Belgium

The internationalisation of higher education has been gaining momentum in recent decades. Universities increasingly offer study programmes (or parts of them) in languages other than the university's primary language of instruction. Usually, that language is English which is referred to as English Medium Instruction (EMI) (see Macaro (2018) for an overview), but there are also programmes available in other foreign languages, such as Dutch in French-speaking Belgium. These programmes are referred to as Dutch Medium Instruction (DMI). Belgium is an interesting area of research for DMI, given the importance of multilingualism, which reflects the country's linguistic diversity and the fact that Dutch is one of its official languages.

This contribution focuses on the socio-economic and language profile of students opting for a monolingual degree in French versus a multilingual degree incorporating English and/or Dutch in their curriculum. Moreover, we compare the students' motivations for enrolling in the monolingual French degree versus a multilingual EMI and/or DMI programme. We also examine the students' regrets and hesitations regarding their chosen language track.

To this end, an online questionnaire was administered to a total of 710 undergraduate law, political sciences, and economics and management students at the French-speaking Université Saint-Louis (Brussels). The survey consisted in multiple choice questions about students' sociodemographic characteristics, language education background, and current university education.

The results indicate that multilingual higher education attracts students with a higher socio-economic status. Moreover, students exhibit personal, instrumental and integrative motivations when selecting a university degree. When comparing the motivations of DMI students in contrast to EMI students, the findings reveal that both groups primarily prioritise career prospects and linguistic factors. However, DMI students placed greater emphasis on career prospects as their primary motivational factor. Another major finding was that local students are less attracted to a degree with Dutch as a minority language in comparison to a degree with EMI. Multinomial and binary logistic regressions were carried out to identify the factors influencing the students' choice of language stream. It was found that students with higher self-reported proficiency in English or Dutch, those who received CLIL secondary education in the target language, were engaged in extracurricular activities in the target language, or had more English or Dutch classes in secondary school, tended to opt for the multilingual university programme over the monolingual one. Furthermore, encouraging findings indicated that nearly all students

---

\*Speaker

†Corresponding author: laetitia.louis@uclouvain.be

who attended a Dutch-speaking secondary school are willing to pursue their university studies in multilingual programmes integrating Dutch.

Taken together, these results suggest that multilingual higher education is closely linked with access inequity, and strengthen the idea that language support prior to entry and throughout EMI and DMI university programmes is needed. In addition, the findings on students' motivational experiences offer pertinent suggestions for universities offering multilingual education and for institutions considering its introduction.

#### References:

Macaro, E. (2018). *English Medium Instruction*. Oxford University Press.

**Keywords:** EMI, DMI, Belgium, higher education, motivation, language proficiency, multilingualism

# Textual enhancement to help learners acquire L2 features: A Meta-analysis

Liting Luo <sup>\*† 1</sup>, Frank Boers <sup>\* ‡ 1</sup>

<sup>1</sup> University of Western Ontario – Canada

Given the role of attention in second language acquisition (Doughty & Williams, 1998; Schmidt, 2001), many researchers have explored ways of making learners take notice of language features that they are likely to overlook. One rather straightforward way of directing learners' attention is to make certain language features visually stand out in texts, for example, through underlining, **bolding**, *italicization*, CAPITALIZATION, highlighting, and using different font sizes (Sharwood-Smith, 1993). According to a meta-analytic review published 15 years ago (Lee & Huang, 2008), the effect of visual (or typographic) enhancement relative to reading-non-enhanced texts, tends to be small. However, this was based on a relatively small collection of primary studies (N = 16), and they all concerned grammar features. Numerous studies have been published since and, interestingly, a fair number of those more recent studies have examined the effect of enhancement on the acquisition of lexical items, especially collocations (e.g., Boers et al., 2017; Vu & Peters, 2022). Therefore, this field needs a new meta-analysis to address the question of the overall effectiveness of the technique and to additionally address the question of whether the effectiveness is likely to depend on the nature of the language feature to which it is applied.

In this meta-analysis, we examined 37 effect sizes from 26 studies (N = 2460). Twelve of them focused on the acquisition of grammar features, and the rest are focused on lexical features. These 26 studies were selected from an initial set 50 studies based on the following inclusion criteria: (a) studies should be between-group experimental or quasi-experimental studies; (b) they should focus on either vocabulary or grammar learning; (c) they should include measures of the participants' prior knowledge of the targeted vocabulary items or grammatical features (or use pseudowords); (d) they should focus on the effect of visual enhancement as an independent variable; (e) they should be situated in the realm of incidental language acquisition rather than language-focused exercises; (f) they should provide a clear description of the research procedure ensuring replicability; (g) they needed to include effect sizes or the necessary descriptive statistics to calculate effect sizes; (h) they needed to be written in English.

Results showed a small to medium effect for the effectiveness of this technique on grammar learning,  $d = 0.329$ , 95% CI (1.86, 5.09),  $p = 0.01$ , while the results for lexical-item acquisition showed a medium effect,  $d = 0.6$ , 95% CI (1.20, 2.43),  $p < 0.01$ . However, the difference fell short of statistical significance ( $p = 0.078$ ). The presentation will also report the analyses regarding other moderator variables, such as input type (written texts vs. captioned videos), test type, and the participants' proficiency level.

**Keywords:** Meta, analysis, Textual Enhancement, Lexical Acquisition, Grammar Learning

---

\*Speaker

†Corresponding author: lluo95@uwo.ca

‡Corresponding author: fboers@uwo.ca

# L2 Learners' listening efforts and cognitive processes in the listen-to-summarize task: a mixed-methods study

Yuhan Luo \* <sup>1</sup>, Dagmar Divjak<sup>†</sup> <sup>1</sup>, Petar Milin<sup>‡</sup> <sup>1</sup>

<sup>1</sup> University of Birmingham [Birmingham] – United Kingdom

Integrated listen-to-write language tasks have found widespread utilization in the assessment of second language (L2) proficiency, primarily due to their capacity to simulate real-world language-use scenarios (Cumming et al., 2004). These tasks mimic key dimensions of language usage such as, for example, everyday and academic communication (Fauzi, 2019; Hosogoshi & Takahashi, 2015). Most research into integrated language tasks in L2 focused on cognitive processes during read-to-write (Michel et al., 2020), and limited research has explored cognitive processes during the listen-to-write task (Rukthong & Brunfaut, 2020). Even fewer studies have focused on recording moment-to-moment indicators of cognitive effort, such as pupillometry, during the listening component of integrated language tasks (Borghini & Hazan, 2018).

In our study, we applied the listening cognitive processes framework developed by Field (2013) for English L2 integrated listen-to-write tasks on a sample of 80 Chinese students with IELTS score of 6.0 or above. We employed a mixed-methods approach, collecting and analysing (a) temporal patterns of cognitive effort as captured by pupillometry, (b) patterns of listening cognitive processes at two levels – lower-level (input decoding, lexical research, parsing) and higher-level (meaning construction, discourse construction) from the stimulated recall task. The listen-to-write task materials were adapted from the Pearson Test of English (PTE), requiring participants to listen to 1-1.5-minute mini-lectures and provide written summaries after listening. The cognitive effort during listening was indicated in terms of pupil size, measured with an Eyelink Portable Duo eye tracker. After finishing the tasks, participants engaged in stimulated recall interviews to elucidate their cognitive processes while listening. Task performance was assessed by using PTE marking criteria, to establish connections between listening efforts, listening cognitive processes, and task performance.

Preliminary data analysis showed a significant pupil size peak at the start of listening and further peaks and troughs of diminishing magnitude, towards the end of the task. This suggests that participants' most intense cognitive engagement occurred at the onset of the task. Subsequent smaller peaks in the middle and towards the end suggest continued task engagement and language processing through listening. During stimulated recall, lower-level listening cognitive processes were more frequently reported than high-level processes. Measures of cognitive efforts and task performance are positively correlated: the larger the average pupil size, the better the performance ( $\rho=0.47$ ,  $p < 0.01$ ). In addition, students who more frequently reported higher-level cognitive processes during stimulated recall are likely to achieve better results in listen-to-write task performance.

The novel mixed-method approach gives detail-rich data from both uncontrolled (pupillome-

---

\*Speaker

†Corresponding author: d.divjak@bham.ac.uk

‡Corresponding author: p.milin@bham.ac.uk

try) and controlled (stimulated recall) behavioural responses. This is a step-change towards an appropriately multi-faceted data collection for multi-faceted integrative language tasks, which not only expands the methodological toolkit for SLA researchers but also provides a more fine-grained understanding of how learners engage with, and process L2 in real-time. This type of data enables much-needed insights into how L2 educators and learners can focus on task engagement and specific cognitive processes to enhance L2 integrated language skills.

**Keywords:** L2 integrated language tasks, pupillometry, cognitive processes, stimulated recall

# Ecological Validity and Diversity in Heritage Bilingualism Research: Examining The Role of Objective, Subjective, and Language Experience Measures.

Alicia Luque <sup>\*† 1</sup>, Bernard Issa <sup>2</sup>, Bryan Koronkiewicz <sup>3</sup>, Mandy Faretta-Stutenberg <sup>4</sup>, Harriet Bowden <sup>2</sup>, Brechje Van Osch <sup>5</sup>

<sup>1</sup> Universidad Antonio de Nebrija – Spain

<sup>2</sup> The University of Tennessee [Knoxville] – United States

<sup>3</sup> University of Alabama [Tuscaloosa] – United States

<sup>4</sup> Northern Illinois University – United States

<sup>5</sup> The Arctic University of Norway [Tromsø, Norway] – Norway

The multidimensional nature of bilingualism necessitates ecologically valid research methods that adequately capture its dynamism and diversity (e.g., De Bruin, 2019; Gullifer et al., 2021; López et al., 2021). In light of this, it is critical for researchers to examine the validity and reliability of proficiency measures to serve as proxies to reflect the wide-ranging outcomes and experiences associated with the continuum of bilingualism. This is particularly pertinent when working with minoritized and racialized communities, such as heritage bilinguals (HLBs), as misrepresentations of bilingualism’s inherent diversity can lead to the propagation of misinformation and perpetuate prescriptive and dominant hegemonic narratives about what it means to be bilingual (e.g., Flores & Rosa, 2015; Ortega, 2020; Bayram et al., 2021). In this context, the present study aims to (a) evaluate the reliability and consistency of commonly used proficiency measures to assess and characterize variability in language outcomes among Spanish-English HLBs and (b) better understand how diversity in language experience influences these outcomes. In particular, we focus on examining (1) the internal reliability and validity of each measure, (2) the interrelationships between measures, and (3) the connections between proficiency measures and language experience factors. Language background, experience, and proficiency data were collected from 43 Spanish-English HLBs in the US. Participants completed an extensive bilingual language experience questionnaire, adapted from both the Bilingual Language Profile (Birdsong et al., 2012) and the questionnaire for Spanish-English bilingual speakers from the National Heritage Language Resource Center (Montrul, 2012), which included subjective self-reports of proficiency in all four language skills (speaking, reading, listening, writing). Objective measures included a lexical decision task (Izura et al., 2014) designed to assess vocabulary knowledge and overall proficiency and a 50-item, multiple-choice written test adapted from a portion of the *Diploma de Español como Lengua Extranjera*, or DELE (Montrul & Slabakova, 2003), designed to tap into grammar and vocabulary knowledge. Preliminary data analyses revealed that the objective and subjective measures are positively related to one another, with the two objective measures having the highest (positive) correlation ( $r(84) = .58$ ). Additional ongoing analyses include (a) the examination of internal reliability of each measure, (b) a correlational analysis that should reveal the strength and direction of associations between these measures and different aspects of the participants’ bilingual experience and, (c) a cluster analysis to understand how

---

\*Speaker

†Corresponding author: aluque@nebrija.es

HLBs with different language experiences pattern across the different objective and subjective proficiency measures, among others. The results of this study have the potential to illuminate the extent to which different proficiency measures and language experience factors can reliably assess and characterize the dynamism and diversity in heritage bilingualism outcomes, enhancing the ecological validity of bilingualism research.

**Keywords:** heritage bilingualism, ecological validity, assessing language proficiency, bilingual experience factors, diversity in bilingualism

# Initial acquisition of L2 Arabic verbal inflection: input processing and first productions

Hedi Majdoub \*<sup>1</sup>, Marzena Watorek \*

<sup>2</sup>, Rebekah Rast <sup>3</sup>, Pascale Trévisiol <sup>4</sup>

<sup>1</sup> UMR-SFL, CNRS – Université Paris 8 Vincennes-Saint-Denis – France

<sup>2</sup> UMR-SFL, CNRS – Université Paris 8 Vincennes-Saint-Denis – France

<sup>3</sup> UMR-SFL, CNRS – The American University of Paris – France

<sup>4</sup> DILTEC - Didactique des langues, des textes et des cultures - EA 2288 – Université Sorbonne Nouvelle - Paris 3, Université Sorbonne Nouvelle - Paris 3 : EA2288 – France

The learnability of L2 inflectional systems is known to be particularly challenging for L2 learners (Bardovi-Harlig 1992), and although numerous studies have investigated underlying reasons for this difficulty (Gor & Chernigovskaya 2005), missing from the research is a clear description of the learners' linguistic environment (input) and their performance relative to this input. Such a description is necessary for understanding L2 input processing and word formation processes in the initial stages of L2 acquisition.

Expanding "first exposure" research (Carroll 2013; Dimroth et al. 2013), this contribution discusses how absolute beginners retrieve words and inflections from novel input and generalize to new forms that are absent from the input. The study reports data collected from French native speakers ( $n=11$ ), unfamiliar with Arabic, who attended a course in Modern Standard Arabic for a total of 6.5 hours. The classroom input was carefully controlled, transcribed, and coded for a precise description of linguistic properties in the input.

A challenge for learners of Arabic is the complex system of subject-verb agreement. In SVO sentences with imperfect aspect, affixes on the verb mark number and gender in agreement with third person subjects. In the singular, the distinction between masculine and feminine is found on the verb prefix (e.g., masculine *ya-* vs feminine *ta-*) with no difference on the suffix. In the plural, the reverse is true: no difference is found on the prefix, but a distinction is found on the suffix (only masculine forms end in a long vowel).

This paradigm was tested in two tasks: aural grammatically judgment and oral question-answer. Both were administered after 4.5 hours of instruction (T1) and after 6.5 hours (T2). During the judgment task, learners heard sentences with number/gender verb markers, balanced at 50% correct/incorrect and were instructed to indicate whether the sentence was correct or not. In the production task, learners saw pictures eliciting number/gender verb markers, such as *al bintu tarsumu* (fem/sg) (the girl is drawing) and were asked in French *Qu'est-ce qui se passe?* ("What is happening?"). Crucially, to observe the categories of frequent vs absent in the input, not all target/elicited items had been introduced in the lessons.

Data collected from both tasks were analyzed quantitatively for accuracy relative to frequency and affix type. The production data were also analyzed qualitatively to see how learners integrated new forms into their interlanguage systems. While results revealed an overall increase

---

\*Speaker



of accuracy from T1 to T2 on both tasks, a frequency effect was found; only performance on present items increased. Concerning affix type, masc/sg forms showed no increase, whereas the accuracy of other forms increased from T1 to T2. An effect of task was also found; scores for plural items were higher in judgment than in production at T1, becoming nearly equal at T2. Finally, at T2 in production, learners began overgeneralizing inflectional endings onto verbs absent from the input. These findings shed light on how learners extract inflectional markings from typologically distant L2 input and use them productively in L2 speech.

**Keywords:** input processing, L2 production, grammaticality judgment, inflectional morphology, L2 word formation, first exposure, Arabic

# Linguistic concepts and L2 research into L2 Classroom: a case of L2 Greek Clitics

Anastassia Mangana \*<sup>1</sup>, Kook-Hee Gil \*

2

<sup>1</sup> University of Athens – Greece

<sup>2</sup> University of Sheffield – United Kingdom

**Background** One of the newly emerging research strands within Generative SLA explores implications of linguistic generalisations and L2 research findings for L2 pedagogy (Melinda et al. 2013, 2018; Gil et al. 2019, White 2022). The current paper takes a step further in taking research into classroom: we take an L2 target property that is a persistent challenge to L2 Greek learners, i.e, clitic pronouns, identify what linguistic concepts underscore the use of variant clitic forms, design accessible and engaging lines of grammatical rules and activities and measure its effect. We implement these steps within the context of an intervention study.

**L2 Clitics and SLA** Clitics as functional morphemes in Romance languages have been shown to be persistently problematic in SLA due to their multifaceted properties, including their position (Bruhn de Garavito 2013), dependency on Discourse (Valenzuela and McCormack 2013) and variant case forms (Santoro 2007).

**Greek Clitics** While Greek shares similar properties, it also has an added complexity due to case alternation both in monotransitive and ditransitive structures. The direct and indirect object can take two cases: accusative and genitive, rendering two case forms in clitics. They are highly idiosyncratic and associated with specific verb types (Anagnostopoulou and Sevdali 2020), which makes L2 acquisition a difficult task. Current textbook instructions, however, appear to overlook this complexity and present one-to-one mapping between syntactic function and case marking. The instruction we propose specifically targets to help learners notice (i) how direct objects of monotransitive verbs are case-marked, specific to each verb type and (ii) how genitive and accusative clitic forms are teased apart in the case of ditransitive verbs and frequent collocations. The underlying motivation in this instruction design is to provide more input for clitics which lie at the lexicon and syntax interface.

**The Study** Our study follows a pretest-treatment-posttest procedure. The participants come from different L1 backgrounds and at lower intermediate level in proficiency (B1). They were randomly assigned to the control group (n=15) and the experimental group (n= 20). A multiple choice task was used in the pretest and posttest where participants had to choose a clitic form under different verb types. As for the treatment, the experimental group had to identify different verb types linked to different clitic forms, while the control group continued with traditional textbook instruction.

**Results/Discussion** Firstly, the pretest results confirmed that both learner groups had difficulty with case-marking in clitics that we suggested above. Further, the two groups were similar in their performance. The posttest results, however, show that the experimental group was

---

\*Speaker

target-like by 70% in identifying correct clitic forms specific to each verb type, while the control group was target-like only by 56%. This suggests the positive effect of the chosen treatment in facilitating the link between case-marking in clitics and specific verb types. These findings further support our claim that the traditional distinction between vocabulary (verbs) and grammar (clitics) is misleading, as it overlooks the complexity in the underlying relationship between lexicon and syntax.

**Keywords:** SLA, Greek, clitics, language pedagogy, case, marking, intervention study

# Can classroom training facilitate predictive processing and oral production of case in L2 German?

Heather Marsden <sup>\*† 1</sup>, Emma Marsden <sup>1</sup>, Holger Hopp <sup>\*</sup>

<sup>2</sup>, Carrie N. Jackson <sup>3</sup>, Thomas Jochum-Critchley <sup>1</sup>, Vivienne Rogers <sup>\*</sup>

4

<sup>1</sup> University of York – York, United Kingdom  
<sup>2</sup> Technische Universität Braunschweig – Germany  
<sup>3</sup> Pennsylvania State University – United States  
<sup>4</sup> Swansea University – United Kingdom

Research on L2 processing of case and gender morphology has found only limited evidence of morphology-based prediction among L2 learners (e.g., Hopp, 2015; Lew-Williams & Fernald, 2010). However, Hopp (2016) found that, after intensive training on German noun gender, upper-intermediate-level L2 German speakers used gender predictively in a visual-world eye-tracking experiment. Building on this, we investigate whether classroom training based on Processing Instruction (PI, VanPatten & Cadierno, 1993) can facilitate predictive use of case morphology in L2 German learners. Activities in PI train learners to attend to morphosyntactic cues, targeting automatization of form-to-meaning processing. We additionally investigate whether comprehension-based training also yields gains in oral production. The structures investigated were SVO (1) and OVS (2) sentences, where determiner case for masculine nouns identifies the noun's role ("der": subject; "den": object).

(1) Der Kellner ruft den Koch.

the.NOM waiter calls the.ACC cook

"The waiter calls the cook."

(2) Den Kellner ruft der Koch.

the.ACC waiter calls the.NOM cook

"The cook calls the waiter."

Participants were L2 German learners from B1/B2-level university classes in England: 18 undertook the intervention and 10 formed a comparison group without training. Both groups

---

\*Speaker

†Corresponding author: heather.marsden@york.ac.uk

completed testing before the intervention (Time 1), then 1 week (Time 2) and 8 weeks (Time 3) after it. Testing included a visual-world eyetracking comprehension task and an oral production task. In the comprehension task, participants heard sentences such as 1 or 2, while viewing two pictures: a waiter calling a cook; a cook calling a waiter. Participants selected the matching picture. The production task aimed to elicit OVS sentences following a discourse context favouring this word order. Between Times 1 and 2, the intervention group received two 45-minute training sessions including an explanation of German subject/object case followed by exercises with SVO and OVS sentences, where correct answers depended on processing of determiner case morphology.

**Results:** At Time 1, both groups had comparably low rates of production and comprehension of OVS, and no evidence of predictive looks towards the correct picture in OVS sentences before the final noun onset. At Times 2 and 3, the intervention group showed significant gains in both production and comprehension tasks, and earlier looks to the target picture in OVS. The comparison group also made gains in production, and small gains by Time 3 in comprehension, but showed no changes in eye movements. Mixed effects models confirm significantly higher gains by the intervention than the comparison group for production and comprehension. Cluster mass permutation modelling of the eyetracking data confirms earlier looks to the target picture in the OVS condition by the intervention group than the comparison group at Times 2 and 3 relative to Time 1, although looks to target continued not to rise until after the final noun onset.

Overall, findings suggest that the training instigated a change in production and processing behaviour, although predictive processing driven solely by case morphology was not fully established. Targeted comprehension practice thus affects early implicit processing.

**Keywords:** L2 processing, morphosyntax, case morphology, German, input processing, grammar teaching, eyetracking, comprehension, oral production

# Playing the catch-up game: How long do (not so) newly immigrated students lag behind?

Nicole Marx \* <sup>1</sup>

<sup>1</sup> University of Cologne – Germany

Since 2015, Germany has witnessed a notable increase in the numbers of newly immigrated schoolchildren, who now account for approximately 12% of the total student population. At the same time, little is known about the academic and linguistic progression of these students. This is due to both the high diversity of the group as a whole and to the differences in educational measures enacted to support them, which range from full integration with no additional language support to full schooling in German as an Additional Language classes for up to two years (Will et al., 2021). Regardless of support received, students are commonly expected to catch up with their non-immigrant peers, i.e., reach grade-level norms in their various academic subjects, after an unspecified period of time. In Germany, it is unclear whether students reach this goal – and, if so, how long they need to do so.

The dearth of research is a central problem for schools and research alike. Of special concern is the development of reading skills in German, since much classroom learning is based on written texts. Earlier studies indicate that the target group needs much longer to converge on mainstream norms than previously thought (i.e., Marx et al., 2021). However, the evidence base remains sparse and unclear, especially following the Covid-19 pandemic.

The present study investigated the reading skills of immigrated students in Grades 5 to 8 ( $n = 76$ ) after a minimum transition period of one year after migration. Since students are expected to catch up to their mainstream peers, students in the same classes who received their full education in Germany were investigated as a comparison group ( $n = 192$ ).

In addition to questionnaires on individual education and language backgrounds, students completed standardized reading tests with measures of reading comprehension and reading fluency in spring 2022. A subgroup was retested annually for two years to investigate learning trajectories. Results of the multifactorial analyses of variance show a main effect for group, with immigrant students performing between 0.7 and 1.24 standard deviations below mainstream students on all measures for all data points, and little evidence that immigrant students were beginning to close the gap to their mainstream peers even after a number of years in mainstream education. The results are supported with details on individual differences in reading results and discussed with regard to further research desiderata, but also to considerations of the usefulness of studies carried out with such a diverse learner population.

## References:

Marx, N., Gill, C., & Brosowski, T. (2021). Are migrant students closing the gap? Reading progression in the first years of mainstream education. *Studies in Second Language Acquisition (SSLA)*, 43(4), 813–837. <https://doi.org/10.1017/S0272263120000583>

---

\*Speaker

Will, G., Homuth, C., Maurice, J. von, & Roßbach, H.-G. (2021). Integration of recently arrived underage refugees: research potential of the study ReGES - Refugees in the German Educational System. *European Sociological Review*, 37(6), 1027–1043. <https://doi.org/10.1093/esr/jcab033>

**Keywords:** immigrant learners, secondary school, reading

# A Corpus-based Spoken Academic Belgian Dutch (SABeD) Word List

Jolien Mathysen \* <sup>1</sup>

<sup>1</sup> KULeuven – Belgium

Research has shown that listening comprehension increases with lexical coverage, or the percentage of known words in a text (Durbahn et al., 2020; van Zeeland & Schmitt, 2013), and that vocabulary knowledge is a strong predictor of academic listening comprehension (Wallace, 2020). For English, studies (Dang, 2022b; Dang & Webb, 2014) have revealed that students need to know 3,000-4,000 word-families (plus proper nouns and marginal words) to reach 95% coverage, or reasonable comprehension, of lectures and 7,000-9,000 word-families (plus proper nouns and marginal words) to reach 98%, or very high comprehension. Lectures were also found to be more (lexically) demanding than labs and tutorials (Dang et al., 2021). In spite of its prevalence, previous studies have demonstrated that (spoken) academic Dutch, and more specifically academic vocabulary, constitutes a stumbling block for second/foreign language (L2) learners of Dutch (in the comprehension of lectures) (Deygers, 2017; Deygers et al., 2017; Deygers et al., 2018). Pre-existing (English) spoken academic word lists have already confirmed their usefulness in developing and validating learning materials, as well as analyzing the lexical demands of academic lectures (Dang, 2022a; 2022b; Dang et al., 2017). However, to date, there was no corpus-based academic word list available for Dutch. Creating a similar list for Belgian Dutch will allow us to analyze the lexical demands of real-life Flemish lectures, as well as develop new corpus-based learning materials. The list can also raise awareness among teachers and students about their academic vocabulary use and struggles (Dang, 2019; Dang et al., 2021; Nation, 2016; Uchihara & Webb, 2022).

The current study aims to develop a spoken academic Belgian Dutch word list as part of the interdisciplinary SABeD (Spoken Academic Belgian Dutch) project. This list is being compiled based on the SABeD corpus. This domain-specific spoken corpus has been created after the example of the BASE-corpus (Thompson & Nesi, 2001), consists of about 200 lectures from Flemish university first bachelor programmes and encompasses four disciplinary domains (Arts and Humanities, Life and Medical Sciences, Physical Sciences, and Social Sciences). It was composed by manually correcting transcriptions from a pre-trained automatic speech recognition (ASR) system (Van Dyck et al. 2021). The list takes into account the parameters of frequency and dispersion (Dang et al., 2017; Szudarski, 2017) and is lemma-based. It is obtained by comparing our raw frequency list to a general Dutch vocabulary word list by Tiberius and Schoonheim (2013) and the use of tools developed by the ‘Leuven Language Institute’ to separate general academic words from proper names and domain-specific words. This presentation will zoom in on the compilation of the SABeD corpus, and illustrate the principles underlying the corpus-based academic word list. It will end with some pedagogical implications.

**Keywords:** SLA, vocabulary, word list

---

\*Speaker



# L2 writers' tripartite engagement with written corrective feedback in different timing conditions. Relationship between engagement with feedback and written accuracy

Sophie McBride \* <sup>1</sup>, Florentina Nicolás-Conesa <sup>2</sup>, Lourdes Cerezo† <sup>1</sup>

<sup>1</sup> Dpto. de Filología Inglesa, Universidad de Murcia – Spain

<sup>2</sup> Facultad de Educación, Universidad de Murcia – Spain

Research on written corrective feedback (WCF) has received abundant attention in the last quarter century, mostly confirming its beneficial effects on L2 written accuracy (see Kang & Hang, 2015). Recent studies have focused on analyzing how L2 learners engage with feedback to shed light on language development (see Roca de Larios & Coyle, 2022). Learner engagement with feedback has been defined as "how learners respond to the feedback they receive" (Ellis, 2010, p.342) and characterized as a multidimensional construct, in as much as it can be examined from three perspectives: cognitive, behavioral, and affective (Han & Hyland, 2015). To analyze L2 writers' engagement with feedback, researchers have used qualitative instruments in case studies, including retrospective verbal reports, student-teacher conferences, semi-structured interviews, or questionnaires. Despite providing rich qualitative data, these instruments involve non-concurrent procedures, which means that they provide information on the result of feedback processing and not on the actual processes whereby learners understand the feedback and engage with it (see Coyle et al., 2023). In contrast, using concurrent procedures such as screencasts, researchers can visualize learners' engagement with WCF in a less disruptive way, "often running invisibly on a computer as students write" (Séror, 2013, p.3). There is also scant research that investigates different feedback timing conditions. It is considered that delayed synchronous feedback does not interrupt the flow of thought compared to immediate synchronous feedback (Harmer, 2007). However, most research has just focused on the beneficial effects, in terms of written improvements, of immediate synchronous feedback or asynchronous feedback without comparing how learners engage with WCF in different timing conditions and the relationship between learners' engagement with WCF and L2 written accuracy. This study investigates learners' tripartite engagement with WCF via screencast technology (SCT) in three feedback timing conditions and the relationship between such engagement and L2 written accuracy. The participants were 20 advanced EFL writers (C1-C2 level) divided into three homogeneous groups according to the results obtained in a placement test and accuracy rates (i.e., percentages of errors) in an initial writing task. The participants completed two writing tasks that were revised after receiving and processing WCF. For each task, participants were given 50 minutes to write and process unfocused indirect WCF: (i) immediate synchronous (i.e., while composing); (ii) delayed synchronous (10 minutes after composing), and (iii) asynchronous (i.e., 4 days after composing) WCF (see Aubrey & Shintani, 2022). The participants in the three conditions were video recorded while writing, processing feedback, and revising their texts using SCT. Our re-

---

\*Speaker

†Corresponding author: lourdesc@um.es

sults indicate the existence of similarities in learners' cognitive and affective engagement with feedback in all three conditions. There were also some similarities in the behavioral engagement of the delayed feedback group and the asynchronous feedback group, but not with the synchronous feedback group. Differences were also found in the decrease of errors among the three feedback timing conditions. The potential advantages of learners' engagement with the three feedback timing conditions will be analyzed and discussed from a writing-to-learn-language perspective (e.g., Manchón, 2011).

**Keywords:** written corrective feedback, error correction, synchronous feedback, asynchronous feedback, feedback timing, engagement with feedback, L2 writing, screencast technology, written accuracy

# Acquisition and development of grammatical gender in L3 Spanish: an eye-tracking study on heritage speakers of Polish in the UK

Carla Menares \* <sup>1</sup>

<sup>1</sup> King's College London – United Kingdom

Most of the evidence on the intralinguistic and extralinguistic factors mediating morphosyntactic transfer in third language (L3) acquisition comes from offline methodologies (Rothman et al., 2019), although there are some exceptions using online techniques (e.g., Abbas et al., 2021; Pereira Soares & Rothman, 2021). Similarly, most of the evidence is derived from cross-sectional studies, so there is scarce data on L3 grammatical development (Puig-Mayenco et al., 2020; Cabrelli & Iverson, 2023). The current longitudinal eye-tracking project aims to fill these gaps by examining the acquisition and development of grammatical gender in L3 Spanish among heritage speakers (HSs) of Polish in the UK.

The heterogeneity among heritage bilinguals will provide an opportunity to investigate how variations in bilingual experience, including factors such as family language dynamics, age of acquisition, heritage language exposure, and participation within the heritage-speaking community (Kupisch & Rothman, 2018; Rodina et al., 2020; Polinsky & Scontras, 2020), influence both transfer selection and developmental sequencing. In addition, this study will delve into intralinguistic factors, namely holistic structural similarity between Polish, English, and Spanish, and the characteristics of grammatical gender in these languages to test current models and theoretical accounts concerning morphosyntactic transfer in L3 (Rothman, 2011, 2015; Westergaard et al., 2017; Slabakova, 2017).

To do so, this eye-tracking study will use the visual world paradigm to monitor participants' real-time eye movements as they process gender assignment and gender agreement cues in L3 Spanish. Lexical and syntactic overlap will be manipulated as well. Facilitative effect of Polish could be observed in the processing of nouns that share the same gender assignment cue (e.g., feminine nouns ending in *-a*) or are assigned to the same gender class in both Polish and Spanish. Conversely, processing costs are anticipated for opaque nouns (e.g., feminine nouns ending in *-e*) and gender agreement cues since gender is marked syntactically on different words in these languages. Participants in the study will include a group of Polish-English bilinguals learning L3 Spanish in the first year of secondary school (aged 11-12) and an L1 English group of learners of Spanish as an L2 to isolate the effect of Polish (Westergaard et al., 2023). Potential changes across development stages will be examined by comparing two testing times.

The findings could shed light on the potential challenges faced by HSs acquiring grammatical gender in an L3 since the evidence suggests it is a domain that poses challenges for both L2 learners and HSs as well (Montrul et al., 2008; Fuchs et al., 2021; Fuchs, 2022), contributing to the discussion on heritage bilingualism. This project aims to collect real-time data about language processing that is expected to provide insights into the source of morphosyntactic transfer, the role of intra and extralinguistic factors, and how they restructure according to the dynamicity of L3 development (Cabrelli & Iverson, 2023). Ultimately, the longitudinal nature

---

\*Speaker

of the study will allow us to go beyond the initial stages of acquisition to start modeling L3 grammatical development.

**Keywords:** third language acquisition, transfer, grammatical gender, heritage bilingualism, eye tracking

# Informal learning of English: Finnish pupils' experiences

Lea Meriläinen \* <sup>1</sup>

<sup>1</sup> University of Eastern Finland – Finland

Learning English informally in out-of-school settings is a growing phenomenon in many countries as a result of globalization and technological development (Reinders et al. 2022). Informal learning has been found to have an impact, among others, on learners' vocabulary development and spoken communication skills (Sundqvist 2009, Sylvén & Sundqvist 2012, Schurz & Sundqvist 2022). Its effects in Finland have been studied little, apart from national evaluations, which have shown pupils' self-reported use of English in their free time to be an important factor influencing their learning outcomes at primary school level (Härmälä et al. 2019, Härmälä & Marjanen 2022).

This study examines Finnish 9th graders' (pupils aged 15-16) experiences with informal learning through the analysis of written tasks, where the pupils were asked to describe their English learning and use outside school. The data for the study was collected as a part of a national assessment of learning outcomes in English in basic education by the Finnish Education Evaluation Centre in the year 2020 (Härmälä & Marjanen 2022). The data consists of 4026 short written texts, which were compiled into a text corpus (altogether 434 387 words). The analysis of the texts is conducted in two steps. The texts are first analysed with corpus-based keyword analysis, which is a commonly used method in corpus-assisted discourse studies in order to identify textual patterns and recurring themes in a data-driven manner (Scott & Tribble 2006, Partington 2013). After this, the keywords are grouped thematically and their broader textual context is analysed qualitatively (Partington 2013). The study seeks to answer the following broad research questions:

1) Which words describing out-of-school learning of English occur frequently in the pupils' texts and in what kinds of contexts?

2) What kinds of broader themes emerge from the closer contextual analysis of the keywords? The results show that the most prominent theme in the pupils' texts is the English-speaking media landscape that they encounter in their free time. Among the most frequent keywords were *games*, *videos*, *series*, *music*, *online* and *social media*, which describe the pupils' common free time activities. Closer contextual reading reveals that the pupils are not only passively exposed to English through these channels, but they also describe themselves as active producers of language in social media and game chats (cf. Sundqvist & Olin-Scheller 2013). The pupils also attach great social importance to free time activities such as gaming and participating in different online communities. In these contexts, English functions as a tool to participate in global communities and to form social networks that shape the teenagers' identities (Edwards & Seargeant 2020). Learning English in out-of-school contexts emerges as motivating and personally meaningful for the pupils, while learning at school is often described as less effective and less interesting. The findings are in line with the notions of demotivation (Sundqvist & Olin-Scheller 2013) and motivational dissonance (Ushioda 2013) observed in other countries as well, thus shedding light on new kinds of challenges that English language classrooms are facing.

---

\*Speaker

**Keywords:** informal learning, Finnish learners of English, learner experiences, corpus based keyword analysis

# Heritage multilingualism in the UK – the case of mood selection in Heritage Speakers of Spanish

Camila Merlo \* <sup>1</sup>

<sup>1</sup> King's College London – United Kingdom

The aim of this project is to understand the acquisition and development of heritage languages (HL) by investigating the selection of the indicative vs subjunctive mood by UK heritage speakers (HS) of Spanish. Over the last ten years, Spanish speakers have gone from 120,000 to 215,000 in the UK, and it's in the top 6 minority languages (Office for National Statistics, 2021) but, unlike in the US, where this field was developed, research on UK Spanish HS is still in its infancy. Some scholars have started to study HS in the UK (e.g. Corbet & Domínguez, 2020), stressing the need to continue to develop the field in the European context, where the linguistic landscape differs from that of the US. In particular, tri/multilingualism is largely present in European households (De Houwer, 2004) and it is not uncommon to find families in which two HL are spoken.

Of note, indicative vs subjunctive mood selection in Spanish is notoriously difficult to master for HS (Polinsky & Scontras, 2020) and shows variability across ages. Blake (1983) states that the system is completely acquired by monolinguals by age 8/10, a time when HS tend to shift dominance to the majority language of the society (Dracos & Requena, 2023). For this reason, this study will focus on children aged 4-17, a key age group to explore the acquisition of mood selection and one that has not been studied in depth, constituting a gap in the literature (Bayram et al., 2021; Montrul, 2018).

The aims of this project are: (i) to analyse the production and comprehension of indicative vs subjunctive mood in a variety of contexts by HS of Spanish in the UK, across two distinct age groups (a) HL bilinguals (b) HL trilinguals; (ii) to explore the individual factors that contribute to the variability observed in HS production and comprehension, such as quantity and quality of input, contact with the home country, amongst others; (iii) to test the competing approaches to HL acquisition and development.

For this purpose, HS of Spanish aged 4-17 and 18 and over and their parents living in the UK will be recruited. They will be asked to complete background questionnaires (De Cat et al, 2022) to tap into their use and exposure to the languages they speak, age of onset and biographical information. An independent measure of proficiency in both languages (Spanish and English) will also be taken into consideration to calculate their language dominance (Montrul, 2015). Their comprehension and production of mood selection in Spanish in earlier and later acquired syntactic contexts (Blake, 1983.) will be assessed with an elicited production task and a comprehension task. I will also explore the methods of Personal Social Network Analysis (Cuartero et al., 2023) to study the links between participants and the community of speakers and complement the self-reported surveys regarding language exposure to the different languages.

---

\*Speaker

**Keywords:** Heritage speakers, Spanish, UK, subjunctive mood, psycholinguistics



# Collaborative Writing Based on Generative AI Models: Revision and Deliberation Processes in the German as a Foreign Language Classroom

Marije Michel \* <sup>1</sup>, Iryna Menke-Bazhutkina<sup>† 1</sup>, Niklas Abel<sup>‡ 1</sup>, Carola Strobl<sup>§ 2</sup>

<sup>1</sup> University of Groningen – Netherlands

<sup>2</sup> University of Antwerp – Belgium

Over the past years, second language (L2) writing instruction has seen a growing number of tools transforming writing practices, ranging from early on spell checkers and translation tools to most recently, generative artificial intelligence (GAI) like ChatGPT. Current generations of L2 writers, therefore, need to learn to critically assess and leverage these technologies to maximize their learning. Simultaneously, L2 instruction faces the challenge of guiding students through the novel contexts and possibilities afforded by GAI (cf. Sasaki, 2023). In the past decades, also collaborative writing has received growing interest of L2 scholars given its effectiveness as an instructional tool (cf. Dobao, 2012). Deliberating with a peer about how to present ideas and structure information in a coherent text using appropriate linguistic forms has shown to promote noticing and support L2 development (Storch, 2002). With the availability of GAI, yet another player can be added to the interaction between L2 writing peers and the evolving text, that is, an AI-generated model (Su et al., 2023).

In this article, we present an in-depth analysis of how GAI influences collaborative writing, revision, and deliberation processes within eight focal students (aged 18 to 23) majoring in German at a Dutch university. Data were gathered during a three-week classroom-based intervention. In session 1, students wrote a synthesis combining the main arguments of two popular scientific articles into one coherent text. For session 2, participants received two AI-generated syntheses of the same source articles and were tasked to compare and evaluate their own text and the GAI-models before revising their own writing. Next, students collaborated in pairs on a joint computer following the same process of synthesis writing on a new topic in session 3; and in session 4 joint guided evaluation of and comparison with GAI-models and revision of their original collaboratively written text. Finally, in session 5 they wrote again an individual synthesis of the texts received during the initial session. Throughout the study, screen recordings captured the development of evolving text and ongoing revision behaviour of all sessions; during collaborative work an external microphone recorded pair's discussions.

In our previous work (author et al., in press), we scrutinized the data based on frameworks by Diels (2022) and van Steendam (2010) for revisions, revision focus, and revision success. Complementing these earlier analyses, the current paper (i) relates the aforementioned revision behaviour to text quality evaluated for communicative adequacy (Kuiken & Vedder, 2017); and

---

\*Speaker

<sup>†</sup>Corresponding author: i.a.menke-bazhutkina@rug.nl

<sup>‡</sup>Corresponding author: n.abel@rug.nl

<sup>§</sup>Corresponding author: Carola.Strobl@uantwerpen.be

(ii) presents a more in-depth review of pair dialogues in sessions 3 and 4 in order to identify patterns of collaborative writing in interaction with the GAI models. For example, some pairs actively engaged with the models by evaluating content, language, and cohesion, while others seemed to ‘blindly’ copy-paste from the GAI-models without caring for adequacy or cohesion. We discuss our findings in light of models of collaborative writing (Li & Zhu, 2013; Storch, 2009) as well as the writing process-product relationship (van Weijen, 2009) by focusing on what aspects of GAI might be used effectively to foster L2 collaborative writing.

**Keywords:** generative AI, collaborative writing, German

# Unpacking L2 explicit linguistic knowledge and online processing of the English modals *may* and *can*: A comparison of acceptability judgments and self-paced reading

Nadia Mifka-Profozic \*<sup>†</sup> 1,2, David O'reilly \*<sup>‡</sup> 2

<sup>1</sup> University of York [York] (UoY) – Heslington, York, YO10 5DD, United Kingdom

<sup>2</sup> University of York – United Kingdom

This study reports on a project that investigated second language (L2) linguistic knowledge and processing of two frequently used modals in the English language, *can* (most frequent in spoken language) and *may* (most frequent in written language). As a contribution to the ongoing discussion on the nature of L2 knowledge (e.g., DeKeyser, 2008; Ellis et al., 2009; Maie & Godfroid, 2022; Suzuki, 2017), the study examined the online (real-time) and off-line comprehension of three different meanings expressed by these two auxiliaries: ability (*can*), epistemic possibility (*may*), and permission (*can/may*). Modals present significant challenges for L2 English learners, primarily due to (1) the intricacies of their form-meaning mapping, with specific forms (no infinitive, no 3rd person singular, no progressive, etc.) conveying different meanings depending on the context, and (2) their interchangeability in some, but not all situations. In addition, some meanings seem to be easier to acquire than others. Previous research has found that epistemic meanings in modals present particular challenges and develop very slowly in both L2 and child first language (L1) (e.g. Giacalone Ramat, 1992; Papafragou & Ozturk, 2006). However, no L2 study to date has investigated online (real-time) processing of the English modals in comparison with off-line comprehension of these verbs.

To address this research gap, the current study sought to answer research questions (RQs) on the extent to which acceptability judgement ratings (RQ1) and reading times (RQ2) are affected by modal-context mismatches/manipulation. Participants were two groups of university students: 40 first language (L1) English speakers and 41 L1 Croatian speakers majoring in (L2) English (estimated to be CEFR B2 level). All participants completed two tests with *can/may* sentence manipulation for ability, epistemic possibility, and permission contexts: a self-paced reading (SPR) task, tapping online processing and, indirectly, implicit knowledge, and an untimed acceptability judgment task (AJT) enabling respondents to make judgments based on conscious decisions and possibly engaging explicit knowledge. L1/L2 English group performance was compared on each test separately, using mixed-effects regression and Cohen's *d* effect sizes to offer a robust and nuanced understanding of effects (Plonsky & Oswald, 2017), with instrument reliability estimates of the instrumentation's psychometric properties and measurement error (Marsden et al, 2018). The results revealed that the L1 and L2 speakers rated the acceptability of sentences in offline tasks similarly; however, there were differences in performance on self-paced reading. Specifically, the L1 English speakers provided empirical evidence of processing sensitivity to modals denoting ability (*can*) and epistemic possibility (*may*), slowing down after modal-context mismatches (*may* and *can* respectively), but no processing sensitivity to prag-

---

\*Speaker

<sup>†</sup>Corresponding author: [nadia.mifka-profozic@york.ac.uk](mailto:nadia.mifka-profozic@york.ac.uk)

<sup>‡</sup>Corresponding author: [david.oreilly@york.ac.uk](mailto:david.oreilly@york.ac.uk)

matic choices expressing permission (*can/may*). L2 learners demonstrated similar sensitivity to L1 speakers when processing modals expressing ability and permission but showed no sensitivity to verb-context mismatches in epistemic modality. Implications for L2 acquisition of modals and future research are discussed.

**Keywords:** modals, sentence processing, acceptability judgments, implicit/explicit knowledge

# The interaction of linguistic and visual cues for the processing of Case in Russian by Russian-German bilinguals: An eye tracking study

Serge Minor <sup>\*†</sup> <sup>1</sup>, Natalia Mitrofanova <sup>\*</sup>

<sup>1</sup>, Marit Westergaard <sup>2,3</sup>

<sup>1</sup> Uit The Arctic University of Norway – Norway

<sup>2</sup> Uit The Arctic University of Norway – Norway

<sup>3</sup> NTNU – Norway

The development of psycholinguistic methods over the last decades has allowed for a deeper and qualitatively new understanding of language processing as it unfolds in real time. Converging evidence suggests that language processing is highly incremental and involves interconnected sub-systems or ‘modules’ (phonological, syntactic, lexical, etc.) which are activated in parallel to decode and anticipate the incoming linguistic signal (Brodbeck et al. 2022; Hagoort 2008; Huetig 2015). However, less is known about the parallel processing of linguistic and non-linguistic information, especially the interaction between non-linguistic (e.g., visual) and grammatical cues.

In this study, we investigate how subtle changes in visual representations can affect the processing of grammatical case cues in heritage Russian by German-Russian bilingual children (N=50, 8-13 y.o.). The linguistic manipulation followed previous designs (Kamide et al. 2003) contrasting SVO sentences – with NP1 marked with nominative case and NP2 marked with accusative, and OVS sentences – where the case marking was reversed. For example, 1) ‘The hare (acc) will now find the fox (nom)’; 2) ‘The hare (nom) will now find the cabbage (acc)’.

Unlike previous studies, we contrasted three types of visual displays: representing **individual referents** (potential agents/themes of the action; e.g., ‘the fox’ vs ‘the cabbage’), **pairs of referents** (the referent of NP1 + potential agent/theme; e.g., ‘the hare + the fox’ vs ‘the hare + the cabbage’), and **events** (representing interactions between the referent of NP1 and the potential agent/theme; e.g., ‘the fox finding the hare’ vs ‘the hare finding the cabbage’). We found that the participants were significantly more sensitive to the case manipulation when they viewed depictions of **events** compared to the other types of visual display. No significant differences in effect size were observed between the displays representing individual referents vs pairs of referents. This suggests that the participants were able to quickly integrate the thematic role information signalled by grammatical case with the visual cues in the **event representations**, facilitating the identification of the target picture. On the other hand, just co-presenting the pairs of referents together doesn’t lead to facilitation in grammatical case processing (as opposed to presenting them in an event representation with clear visual cues to thematic roles).

---

\*Speaker

†Corresponding author: sergey.minor@uit.no

**Keywords:** eye, tracking, visual context, grammatical case, parallel processing, bilinguals

# Revisiting Agent-Preference in language and perception : An eye-tracking study of causal scene narration by Syrian learners of French

Copin Mireille \* <sup>1,2</sup>, Ines Saddour <sup>1</sup>, Cyrille Granget <sup>1</sup>

<sup>1</sup> Laboratoire de NeuroPsychoLinguistique – Université Toulouse - Jean Jaurès – France

<sup>2</sup> Université Toulouse - Jean Jaurès – Université de Toulouse – France

Our study aims to investigate the influence of reading direction in the perception and representation of actants in causal events' verbalizations produced by L1 and L2 French speakers. As part of a larger project on conceptual restructuring in Syrian Arabic learners of French, we collected eye-tracking and verbalization data in a scene-retelling task with 2 groups, 29 Syrian (SA) adult learners of French (FR) and 22 speakers of French L1.

Previous research on event apprehension has suggested an "Agent-first" preference in retelling events in both L1 and L2, where the Agent ("Doer") of an action is used in first subject position and receives more visual attention than the Patient ("Undergoer") (Isasi-Isasmendi et al. 2023). However, this preference can be modulated by syntactic priming (Sauppe & Flecken 2021). Reading direction has also been shown to influence cognitive preferences in Agents' position, affecting language production (Esaulova 2020). Clearly, more research is needed to test the Agent-first preference based on dynamic stimuli, especially examining oral productions along with visual attention in speakers of languages with different reading directions.

Our study aims to contribute to filling this gap by examining oral verbalizations of collision events as well as visual attention, through eye-tracking. All the participants had to watch and narrate 16 scenes in which one female character (Doer) hits another (Undergoer) with an instrument. The Doer's position was counter-balanced across stimuli, and mirror versions of the video were presented to half of the participants. Each group did the task in French, L1 and L2, and the learners group re-did it in Syrian Arabic L1 after some distraction tasks.

In narrations, we identified the preferential subject of the collision verb as well as the first introduced character. We compared dwell times and fixations, specifically at the time of the collision (instrument-undergoer contact), for our two Areas of Interest (AOI) : the characters' head and shoulders.

At the utterance level, analyses of verbalizations corroborate the linguistic Agent-first preference for both groups in L1 and L2, independently of the Doer's on-screen position. Both groups overwhelmingly chose the Doer as the subject of the collision verb, which shows no effect of reading direction. However, analyses of all the utterances in the discourse show no preference to mention the Doer's first and suggest that the order of mention "Undergoer > Doer" is more frequent in French L1. Interestingly, this preferential order is influenced by the characters' on-screen positions : Doers on the left of the screen are more likely to be mentioned first. Conversely, in Syrian Arabic L1, first mentions of the Doer are more frequent when shown on the right. This effect is not observed in FRL2, which suggests conceptual restructuring susceptible of lessening the right-to-left preference. Finally, contrarily to what is expected, eye-tracking data suggest that more attention is paid to the Undergoer, both in total dwell times and in first fixations

---

\*Speaker

at the time of the collision. These observations encourage revisiting the notion of "Agent-first Preference", exploring it on the discourse level and based on dynamic stimuli.

**Keywords:** Eye tracking, French L2 acquisition, Syrian Arabic L1, Reading direction, Agent



# Learning to use single words and formulaic sequences creatively in second language writing: The impact of explicit instruction

Rebecca Moden \* <sup>1</sup>

<sup>1</sup> Institute of Education – United Kingdom

Developing second language learners' productive knowledge of vocabulary is essential for effective communication. In English for Academic Purposes (EAP) contexts, a sophisticated range of single words (SW) and formulaic sequences (FS) need to be mastered to produce complex texts. However, there are few pedagogical intervention studies focusing on productive mastery. It also remains unclear whether SW and FS can be learned equally well from the same instructional method, with a limited number of studies showing mixed results (Alali & Schmitt, 2012; Laufer & Girsai, 2008).

To address these gaps, the present study assessed whether explicit instruction boosted productive knowledge of SW and FS to two levels: 1) successful use in creative sentence-writing, 2) autonomous use in coursework essays. Due to their popularity in EAP coursebooks, two types of contexts were chosen as the instructional method. First, cohesive contexts involving sentences developing a single theme, and unconnected sentences describing unrelated events. A further aim was therefore to compare the gains from these two methods. Based on Schema theory, it was predicted that activating learners' schemata on one topic in the cohesive condition could produce superior learning (Chilton & Ehri, 2015). A comparison of the productive gains of SW and FS was also made.

Sixty-five postgraduate students learned 34 SW and 30 FS in one of three conditions (cohesive, unconnected, control) in six hourly lessons. Both groups encountered the vocabulary in sentences with the meanings provided, then the cohesive group ( $n = 20$ ) practised the vocabulary in sentences on one theme, while the unconnected group ( $n = 20$ ) practised the items in unrelated sentences. A control group ( $n = 25$ ) completed the pre-tests and post-tests. The tests measured the ability to write original sentences using the items. Forty coursework essays provided by 20 learners were analysed. A holistic three-point rating scale was developed to measure vocabulary use in the tests and essays. Five criteria were included to evaluate learners' sentences: level of appropriacy, detail and creativity; and lexical and grammatical choices.

A mixed-effects model showed that both experimental groups outperformed the control group in creative sentence-writing. The odds of successfully using the SW and FS were approximately five times and nineteen times greater than the control group for the unconnected group and the cohesive group respectively. There was a significant advantage of the cohesive condition over the unconnected condition, but with a small-sized effect. SW and FS were learned at similar rates. However, the preliminary results showed that the experimental groups used significantly more SW than FS in their coursework essays. Pedagogical implications will be discussed.

## References

---

\*Speaker

Alali, F., & Schmitt, N. (2012). Teaching formulaic sequences: The same or different from teaching single words? *TESOL Journal* 3(2), 153-180.

Chilton, W. C., & Ehri. L. C. (2015). Vocabulary learning: Sentence contexts linked by events in scenarios facilitate third graders' memory for verb meanings. *Reading Research Quarterly* 50(4), 439-458.

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: a case for contrastive analysis and translation. *Applied Linguistics* 29(4), 1-23.

**Keywords:** Vocabulary, Formulaic language, Teaching formulaic sequences, Contextual vocabulary learning, Second language vocabulary acquisition, Second language vocabulary instruction.

# Diagnosing L2 pronunciation challenges for Japanese medical students: top-down and bottom-up approaches

Ingrid Mora-Plaza <sup>\*†</sup> <sup>1</sup>, Yui Suzukida <sup>\*</sup>

2

<sup>1</sup> Universitat de Barcelona – Spain

<sup>2</sup> Juntendo University – Japan

In the field of second language (L2) pronunciation research, one crucial question addressed by scholars is how to identify effective teaching methods that enhance L2 learners' speech comprehensibility for successful communication (Derwing & Munro, 2015). While most phonetic training studies (Rato, 2014) and pronunciation instruction interventions (Gordon & Darcy, 2019) have targeted difficult L2 pronunciation features as a function of L1-L2 phonetic similarity (Cebrian, 2021) or relative functional load (Suzukida & Saito, 2021), fewer studies have adopted a diagnosis approach consisting of identifying problematic L2 phonological forms from learners' spontaneous speech (Isbell, 2020; Levis & Echelberger, 2022). This diagnosis may be particularly important in unique contexts, such as aviation and medicine, where it is crucial for learners to become precise communicators and to accurately pronounce words and expressions (e.g., medical terminology, diagnoses, procedures, and drug names) because the safety and well-being of individuals are at stake (Baker & Murphy, 2011; Moghadam et al., 2016). The present study is an attempt at diagnosing L2 pronunciation challenges from research-based predictions and teacher/learner-informed perspectives (top-down approach) and specific and global assessments of L2 learners' speech during interaction (bottom-up approach).

Twenty-five Japanese intermediate learners of English who were enrolled in a medical English course participated in the study. As part of their course assessment, learners engaged in patient-doctor conversations that simulated real cases of history taking. The participants served as doctors who were to complete history taking from patients (i.e., identify the chief complaint, details of the symptoms, social and family background, and past medical history), and highly proficient L2 interlocutors served as simulated patients who responded to the inquiries from doctors. On the one hand, qualitative data about learners' pronunciation difficulties was collected from two experienced teachers and participants. On the other hand, fine-grained phonological analyses of pronunciation errors were conducted by two experts in English phonology, and global measures of L2 pronunciation proficiency (i.e. comprehensibility and accentedness) were taken immediately after the patient-doctor conversations by the L2 interlocutors.

Preliminary results from specific assessments of L2 pronunciation errors suggested that segmental errors (e.g., drug /dr/ as /dlæ/) were more common than suprasegmental errors and they mainly originated from L1-L2 phonetic differences (e.g., English /r/ assimilated to Japanese /l/) and as a result of high functional load (e.g., English /æ/-// contrast). These errors were detrimental to the accurate production of medical-specific phrases and lexical items, and resulted in communication breakdowns, thus affecting overall L2 comprehensibility and accentedness.

---

\*Speaker

†Corresponding author: imoraplaza@ub.edu

This tendency was also reflected in the questionnaire responses by experienced teachers and L2 learners. The observations from this triangulated diagnosis (i.e., learners-teachers-experts) will be reflected in the design of a communicative task-based intervention (see Mora-Plaza, 2023) that will focus on phonetic forms that may negatively affect L2 learners' comprehensibility, with the future aim of implementing task-based pronunciation teaching as a methodology to enhance L2 pronunciation development in medical courses in Japan.

**Keywords:** L2 pronunciation learning, pronunciation diagnosis, English as a foreign language, task, based pronunciation teaching, L2 interaction

# Factors affecting learning of pragmatic formulaic sequences from audio-visual input

Natalia Moskvina \* <sup>1</sup>, Roger Gilabert <sup>1</sup>

<sup>1</sup> Universitat de Barcelona – Spain

Formulaic sequences are an important part of vocabulary that greatly contributes to overall proficiency (e.g., Boers et al., 2006; Crossley et al., 2015). However, the input language learners receive is not always authentic in terms of formulaic patterns (e.g., Northbrook & Conklin, 2018). Rich authentic input is especially important for pragmatic formulaic sequences since they rely heavily on the context. Recent research on audio-visual input has shown its potential for reflecting accurate distribution of formulaic sequences and promoting their incidental learning (Lin, 2014; Majuddin et al., 2021), although a number of factors associated with learner and learning conditions seem to have an impact on the process (Puimège & Peters, 2020). The present study seeks to further explore formulaic learning from this type of input by extending the research focus to pragmatic formulaic sequences and introducing previously unexamined learner- and context-related factors that might give us further insight in the learning of formulaic language from video watching, as well as by investigating the role of the factors that have been already discussed in relation to other formulaic types.

The data was collected from 114 Catalan-Spanish learners of English, assigned to one of the 4 genre groups: sitcom, TED talks, drama, and action. Each group was exposed to approximately 30 minutes of captioned videos of the corresponding genre. Learning of the same 6 target formulaic sequences was measured through the pre- and immediate post-tests. Participants' vocabulary size, statistical learning ability, and metalinguistic awareness of formulaicity were evaluated as well.

Results show that metalinguistic awareness and prior knowledge of the target items are significant predictors of learning outcomes, with positive and negative effects correspondingly. Interaction between genre and learner-related factors was observed as well, suggesting that not all genres are equally beneficial for all types of learners. These findings contribute to the discussion on the potential of audio-visual input to promote learning of formulaic sequences and the most efficient ways of using this type of context.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language teaching research*, 10(3), 245-261.

Crossley, S. A., Salsbury, T., & Mcnamara, D. S. (2015). Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*, 36(5), 570-590.

Lin, P. M. (2014). Investigating the validity of internet television as a resource for acquiring L2 formulaic sequences. *System*, 42, 164-176.

Majuddin, E., Siyanova-Chanturia, A., & Boers, F. (2021). Incidental acquisition of multiword

---

\*Speaker

expressions through audiovisual materials: The role of repetition and typographic enhancement. *Studies in Second Language Acquisition*, 43(5), 985-1008.

Northbrook, J., & Conklin, K. (2018). "What are you talking about?" An analysis of lexical bundles in Japanese junior high school textbooks. *International Journal of Corpus Linguistics*, 23(3), 311-334.

Puimège, E., & Peters, E. (2020). Learning formulaic sequences through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, 42(3), 525-549.

**Keywords:** formulaic language, audio, visual input, incidental learning, genres, metalinguistic awareness, statistical learning ability

# How do phonetic reductions affect the perception of French learners of English?

Marine Mouquet \* <sup>1</sup>

<sup>1</sup> Laboratoire Parole et Langage - UMR 7309 – Aix Marseille Université, Centre National de la Recherche Scientifique – France

In everyday speech, words often exhibit considerable variability, leading to the omission of segments. This phenomenon, commonly referred to as *phonetic reduction*, can be exemplified by instances such as *summary* becoming (smri) or even, *yesterday* becoming (je). The present study focuses on the former reduction that is, the deletion of a vowel, also known as *syncope*. Native speakers effortlessly comprehend and produce these reductions, frequently remaining unaware of such variations. However, recent studies have revealed that non-native listeners, including L1 Dutch learners of French (Brand & Ernestus, 2018) and L1 Dutch learners of English (Mulder et al., 2015), face greater difficulties in processing reductions in L2 speech compared to native speakers due to their limited exposure to such phenomena (Morano et al., 2023). This study seeks to contribute to this growing body of knowledge by exploring the interaction between the acoustic input (surface form) and the abstract linguistic representation (underlying form) within an L2 context, with a focus on L1 French learners of English. In doing so, we investigated words with possible phonetic variants reflected by syncope (e.g., for *camera* (kæmr) vs (kæmr)). Various factors including frequency, orthography, phonotactics, and cognates guided our stimuli selection process. Previous studies reported high-frequency words to be more reduced than less frequent words in conversational speech (e.g., Ranbom & Connine, 2007; Brand & Ernestus, 2018). However, existing corpora do not provide frequency for each reduced variant. To address this and to ensure that our selected stimuli accurately reflect native British English speech, we developed a pre-task involving subjective frequency ratings (following Brand & Ernestus) for each variant (full vs. reduced) by native British English listeners. In this task, they had to listen to both variants and identify which form they think is most frequent. These results enabled us to establish a useful basis to get better insights into L2 English acquisition. Currently, we are developing an ABX discrimination task with full and reduced forms, which will reveal whether French learners of English with different proficiency levels benefit from subtle phonetic information when processing reduced speech. We aim to determine how words stored in the mental lexicon influence L2 learners' abilities to distinguish between phonetic variants. Results will assist in providing a clearer insight into L2 learners' difficulties in acquiring and processing English variants and will contribute to L2 phonology models.

## References

Brand, S., & Ernestus, M. (2018). Listeners' processing of a given reduced word pronunciation variant directly reflects their exposure to this variant: Evidence from native listeners and learners of French. *Quarterly Journal of Experimental Psychology*, 1240-1259.

Morano, L., Ten Bosch, L., & Ernestus, M. (2023). Second language learners acquire reduced word forms just like they acquire full forms: From exposure. *Linguistic Approaches to Bilingualism*.

---

\*Speaker

- Mulder, K., Brekelmans, G., & Ernestus, M. T. C. (2015). The processing of schwa reduced cognates and noncognates in non-native listeners of English. *Proc. of ICPHS 2015*.
- Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, 273-298.

**Keywords:** phonetic variation, phonetic reductions, L2 learners, psycholinguistics, language perception



# Use and effect of the Baby Sign on language and communicative acquisition in childcare centers

Lucie Métral \* <sup>1,2</sup>

<sup>1</sup> Laboratoire Dynamique du Langage -UMR 5596 CNRS Université Lumière Lyon 2 – CNRS :  
UMR5596, Université Lumière - Lyon II – France

<sup>2</sup> Forellis – Université de Poitiers, 1 Rue Georges Bonnet, TSA 51106, 86073 Poitiers, France Cedex 9. –  
France

A significant number of hearing pre-school children are exposed to Baby Sign (BS) at home or in childcare settings. BS involves speaking to children with no hearing difficulties, systematically combining sign language with speech to support their language development. Although the majority of users praise BS, but the benefits of its use on children’s comprehension and production and on adult-child interactions have yet to be scientifically demonstrated (Fitzpatrick, 2014; Johnston, 2014). The study, presented here, is an interventional study aimed at describing the use of BS in 4 childcare centers by around 50 children aged between 2 months and 3 years and their trained educators, and at measuring its effect on communicative development. The methodology used consisted of filming and observing signs made by fewer formalised sessions of nursery rhymes and exchanges, at +2 weeks, +3 months, +6 months and +9 months after BS training. We will propose qualitative analyses based on transcriptions and annotations of the videos using ELAN software (MPI, 2023). Preliminary results show that, in the case of adults, signs are produced as early as two weeks after training, particularly during nursery rhymes. Moreover, the signs were systematically associated with speech. Over time, the signs used by the professionals became more diversified and were used in contexts other than nursery rhymes. In quantitative terms, the results vary greatly from one childcare centre to another, from one situation to another and even from one adult to another. We hypothesise that the children are influenced by adult practices, i.e. they sign mainly with the adults who sign, sign mainly during nursery rhyme sessions before extending the use of signs to other activities, and sign simultaneously with or in anticipation of the adults’ signs. However, it seems that the children’s signs are often produced without speech and that they tend to be more precise with repeated use. Based on these initial results, we hope to gain a better understanding of the role of BS in language development in children in ‘atypical’ bilingual situations.

**Keywords:** Baby Sign, multimodality, language development, nursery rhymes

---

\*Speaker

# Supporting heritage languages outside the family: a cross-sectional study on early German-Spanish bilingualism in Germany and Spain

Natascha Müller <sup>\*† 1</sup>, Laia Arnaus Gil <sup>\*</sup>

<sup>2</sup>, Johanna Stahnke <sup>\*</sup>

2

<sup>1</sup> Bergische Universität Wuppertal – Germany

<sup>2</sup> Bergische Universität Wuppertal – Germany

Previous research on first language (L1) acquisition in multilingual children raised in countries with one majority language (ML) has highlighted the role of child-external factors for successful acquisition and active use of the heritage language(s) (HL(s)). Especially, in early child trilingualism, 20% of daily input in one of the child’s languages (Quay 2008) is enough for the children’s success in reaching proficiency levels in the HL(s) comparable to monolingual peers (Chevalier 2015 for an overview). Other factors are family language policy (FLP) and variety of contacts in the HL(s) (Arnaus Gil et al. 2021). Grammatical development can be accelerated in bilingual and trilingual children if compared with monolingual children (Arnaus Gil, Stahnke & Müller 2021). Acceleration can also be observed in children who have a so-called ‘weak’ language, the source of which may be less exposure (Arnaus Gil & Müller 2020). Still open is the role of the educational institution in fostering active bi- and trilingualism. We have analyzed cross-sectional data of 66 simultaneous German-Spanish children (mean age 4;10) being raised in Germany (N=27) or Spain (N=39) and compared them to five Spanish and six German monolingual children from different longitudinal corpora of the CHILDES data collection (MacWhinney 2000). The study consists of spontaneous speech recordings (duration: 15 minutes) in each of the bilinguals’ L1s, from which the Mean Length of Utterance in words (MLUw) was calculated and taken as a proficiency measure. 46 of the 66 children were raised according to the OPOL-strategy (One Person One Language). We distributed the children’s data into four age groups. Spanish monolinguals and Spanish ML bilinguals show a similar pattern, while the MLUw values for Spanish HL children oscillate between 2.0-2.5 from age 3;4. The difference between the bilingual groups is statistically significant ( $t(6) = 4.15, p = 0.006$ ). On the contrary, the two bilingual groups with German as an ML and an HL, respectively, do neither diverge from each other nor from the monolingual group. One possible explaining factor could be the FLP. This analysis shows that the OPOL-strategy strengthens MLUw growth and children generally achieve slightly higher MLUw values for both languages. These differences seem to disappear at older ages and they are not statistically significant across the groups. The high MLUw values achieved in HL German and the significantly lower MLUw values in HL Spanish can be explained by the fact that educational institutions supported multilingualism

---

\*Speaker

†Corresponding author: nmuller@uni-wuppertal.de

to different degrees: while in Spain bilinguals went to a German school and the school's lingua franca was exclusively German, children in Germany attended a bilingual Spanish-German institution, where both languages were used. Acknowledging the fact that grammatical development can be accelerated in multilingual children with less exposure in one of their languages, our results indicate that an institutional setting which guarantees a (sufficiently) high variety of contacts can compensate for disadvantages caused by other circumstances. In this respect, multilingual L1 development does not differ from adult second language acquisition, as outlined by Clahsen, Meisel & Pienemann (1983), among others.

**Keywords:** Heritage Language, multilingual children, cross, sectional, spontaneous production, MLU, Spanish, German, OPOL, variety of contacts

# Could incrementally learned expectations lead to the gradual emergence of the mismatch negativity (MMN) response?

Jessie S. Nixon <sup>\*† 1</sup>, Jacolien Van Rij <sup>2</sup>

<sup>1</sup> Carl Von Ossietzky Universität Oldenburg = Carl von Ossietzky University of Oldenburg – Germany

<sup>2</sup> Bernoulli Institute for Mathematics for Mathematics, Computer Science and Artificial Intelligence Groningen, – Netherlands

The mismatch negativity (MMN) component is one of the most widely researched event-related potentials (ERPs) and is commonly used in second-language acquisition research to measure learning of non-native speech sounds (e.g. Du et al., 2023; Højlund et al., 2022, Shafer et al., 2021). But what underlying mechanisms produce the MMN response is a hotly debated topic. One commonly held view, proposed by Naatanen et al. (1978), is that the MMN reflects detection of a mismatch from a "memory trace". This view posits that the MMN is elicited only in response to mismatch detection; if no mismatch is detected, no MMN is elicited. Alternative proposals have suggested that the MMN is a gradient response that corresponds to the *degree of expectation* of a stimulus (e.g. Jaaskelainen et al., 2004; May & Tiitinen, 2010). In this study, we examine evidence to help to distinguish between these two accounts.

We recorded EEG from 16 native German-speaking participants with no prior knowledge of tonal languages as they passively listened to stimulus sequences sampled from a bimodal distribution of Cantonese mid-level and high-level lexical tones. Participants heard sequences of four same stimuli (*standards*) followed by one different stimulus (*deviant*). There were 8 blocks of 210 stimuli (42 sets of 5 trials = 210 per block). Preprocessing was carried out to remove noise and movement artefacts (1183 trial-channel combinations were excluded, 4.4%). Ocular artefacts were excluded using Independent Component Analysis on unaveraged data.

We carried out two analyses: 1) comparison of ERP amplitude between standard and deviant stimuli in the MMN window, 150-200ms post-stimulus (MMN response); and 2) trial-by-trial response amplitude in standard and deviant stimuli over the experimental block. Results showed a significant MMN response (analysis 1): consistent with previous studies, there was a significant difference between standard and deviant stimuli in the MMN window. Most importantly, there was also a significant *change* in response amplitude in the standard stimuli over the course of the block (analysis 2). At the beginning of block, standard and deviant stimuli elicited essentially the same response; however, with multiple repetitions of the standard stimuli, amplitude decreased over the block. This decrease in amplitude did not occur for the deviant. This pattern of results suggests that the MMN gradually emerges due to incremental changes in response to the standard stimuli.

In summary, the present results provide new evidence regarding the interpretation of the MMN. The finding that standard and deviant stimuli initially elicit the same ERP response, combined with the gradual change in response to the standard stimuli, suggests that the MMN emerges as a relative difference in response amplitude due to a relative difference in *expectation* of the different stimulus types. These results are compatible with recent findings suggesting learning of

---

\*Speaker

†Corresponding author: jessie.nixon@uni-tuebingen.de

acoustic speech cues occurs through prediction (i.e. expectation) and feedback from prediction error (Nixon, 2020; Nixon & Tomaschek, 2021). The MMN may reflect aspects of the underlying neural processes.

**Keywords:** auditory evoked response, ERPs, implicit learning, prediction error

# Turn-Taking in Second Language Acquisition: How do L2 learners make use of prosody in turn-end prediction?

Emilia Nottbeck \* <sup>1</sup>

<sup>1</sup> University of Paderborn – Germany

In this paper I present a research project which focusses on the question in how far L2 learners make use of prosodic cues in predicting the end of a speaker's turn in conversation. Current research findings in turn-taking suggest that speakers make predictions about the end of an upcoming turn (Levinson, 2016). In doing that, speakers draw on different cues, e.g. prosody, semantics, syntax, etc. in order to time and launch their own utterance (e.g. Corps et al. 2018). Regarding the role of prosody, research results are still inconclusive (de Ruiter et al., 2006; Bögels and Torreira, 2015).

My research project targets the role of prosodic cues and seeks to ascertain whether L2 learners of English on different proficiency levels employ prosodic cues in the same way as native speakers. In order to investigate this issue, I partly replicated Bögels & Torreira's (2015) study and conducted an online button-press experiment. Participants were a) 30 beginning and b) 30 advanced learners of English and c) 27 native speakers. The proficiency levels of the learners were determined with communicative tasks categorizing the learners according to the stages of acquisition as outlined in Processability Theory (PT) (Pienemann 1998; 2005). A distinction was made between learners that had acquired all stages of acquisition (stage 6) and those that were at lower levels (3 & 4). In the reaction time experiment, participants were presented with two types of questions that differ in regard to their intonational contour: a) Wh- questions (e.g. *Where do you live?*) and b) declarative questions (e.g. *Your favorite music is rap?*). The questions were presented in four conditions: original short, original long, manipulated short and manipulated long. Short versions (e.g. *Where do you live?*) differed from long versions through the addition of *at the moment* at the end of the question (e.g. *Where do you live at the moment?*). The manipulations consisted of truncated and cross-spliced versions of these original stimuli (e.g. manipulated short: *Where do you live at the moment?*; manipulated long: *Where do you live +at the moment?*).

The results indicate a significant difference for all groups between the two original and the manipulated versions. Responses in all groups were delayed in both types of questions for the manipulated short versions compared to the original short versions and speeded for the manipulated long versions compared to the original long versions. Regarding learner group differences, advanced learners anticipated turn ends in all long (and possibly original short) versions of the stimuli, making their response latencies comparable to native speakers. Beginning learners on the other hand, only anticipated turn-ends in the manipulated long versions of wh-questions. These findings highlight the role of prosody for turn-end prediction in learners at different stages of acquisition and provide a starting point for the investigation as to how learners develop a native-like use of prosodic cues in their language acquisition process.

Bibliography:

---

\*Speaker

Bögels, S., & Torreira, F. (2015). Listeners use intonational phrase boundaries to project turn ends in spoken interaction. *J. Phonetics*, 52, 46-57.

**Keywords:** turn, taking, prediction, reaction time experiment

# Processing tense/aspect agreement violations on-line: A replication of Roberts and Liszka (2013) with L1 Mandarin and L1 Croatian learners of L2 English

David O'reilly <sup>1</sup>, Jelena O'reilly <sup>\*† 1</sup>, Leah Roberts <sup>1</sup>

<sup>1</sup> University of York [York, UK] – United Kingdom

Temporality (the locating of a situation in time) is a ubiquitous concept, but can be expressed morphosyntactically, lexically, and/or pragmatically depending on language (Klein, 2009). In English, temporal relations are expressed via grammaticalised tense-aspect, while languages such as Mandarin do not require overt time marking (Klein, 2009), discrepancies which can cause second language (L2) English learner significant difficulties (Hawkins and Liszka, 2003). Roberts and Liszka (2013) were among the first to investigate both implicit (self-paced reading/SPR) and explicit (acceptability judgement task/AJT) knowledge of L2 English tense-aspect. The results from 20 L1 French and 20 L1 German advanced L2 English learners showed that while both L2 groups had explicit knowledge of past simple and present perfect tenses, in online comprehension the L1 French group was significantly sensitive to mismatches for both tenses, while the L1 German group showed no sensitivity to either tense, findings attributed to L1 transfer.

To test the generalisability of L1 effects on tense-aspect acquisition, the current study conducts a close replication of Roberts and Liszka's (2013) seminal experiment. The replication (preregistered on the OSF) uses 33 L1 Mandarin L2 English learners (aspect but no tense grammaticalised in the L1) and 33 L1 Croatian L2 English learners (tense and aspect grammaticalised in the L1, akin to the original study's L1 French group) of equivalent proficiency. Importantly, neither Mandarin nor Croatian have a dedicated present perfect tense like English. Assuming the L1 influences postulated in the original study, we hypothesise:

- *H1*: the L1 Croatian L2 English learners will have online sensitivity to mismatches with the past simple tense (existing in Croatian) but not the present perfect (not existing in Croatian);
- *H2*: the L1 Mandarin L2 English learners will not have online sensitivity to violations of either tense (neither are encoded in their L1).
- *H3*: Both L2 groups will demonstrate explicit knowledge of both tenses.

Data are currently being collected and, as a departure from the original study's ANOVAs, will be modelled using mixed-effects regression with SPR time and AJT rating as outcomes; group

---

\*Speaker

†Corresponding author: jelena.oreilly@york.ac.uk



and tense as interactive, fixed-effect predictors; and by-participant and -item random slopes for within-measures and intercepts for between-measures (Barr, 2013) to avoid the need to average out observations for each participant, accounting for random subject/item variance in a single analysis.

## References

Barr, D. J. (2013) 'Random effects structure for testing interactions in linear mixed-effects models', *Frontiers in Psychology*. Frontiers Media SA, 4(1).

Hawkins, R. and Liszka, S. A. (2003) 'Locating the source of defective past tense marking in advanced L2 English speakers', in van Hout, R. et al. (eds) *The Lexicon-Syntax Interface in Second Language Acquisition*. John Benjamins Publishing Company, pp. 21–44.

Klein, W. (2009) 'How time is encoded', *The expression of time*, 3(1), pp. 1–43.

Roberts, L. and Liszka, S. A. (2013) 'Processing tense/aspect-agreement violations on-line in the second language: A self-paced reading study with French and German L2 learners of English', *Second Language Research*, 29(4), pp. 413–439.

**Keywords:** tense, aspect, self, paced reading, replication

# How acquirable are English articles for L2 learners? Evidence from online processing and production by L1 Mandarin and L1 Croatian speakers of L2 English

Jelena O'reilly <sup>\*† 1</sup>, Leah Roberts <sup>1</sup>

<sup>1</sup> University of York – United Kingdom

English articles (a(n)/the) are one of the hardest features for second language (L2) learners to acquire, but what role does the first language (L1) play in the L2 article acquisition process? The Morphological Congruency Hypothesis (e.g., Jiang et al., 2011) would predict that where articles are a unique-to-L2 feature they will not be fully acquirable, while the Competition Model (e.g., Van Hell & Tokowicz, 2010) would propose full article acquisition as L1 transfer would be neither a help nor hinderance in this respect. However, evidence for (or against) these predictions about L2 English article acquirability is limited to: (1) metalinguistic and production data (i.e., online processing is less known) and (2) L1 Mandarin L2 English speakers, predominantly. The latter is particularly problematic since Mandarin is suggested to be on a path of grammaticalizing certain markers of (in)definiteness (e.g., Chen, 2004; Jenks, 2018), and so is arguably not strictly fully article-lacking.

To address these gaps, the present study tested 24 L1 Mandarin L2 English speakers (not fully article-lacking), 22 L1 Croatian L2 English speakers (fully article-lacking), and 24 L1 English controls. All L2 English speakers were highly proficient (Oxford Quick Placement Test scores: L1 Mandarin  $M = 49$ ,  $SD = 1.41$ ; L1 Croatian  $M = 53$ ,  $SD = 2.19$ ) and late bilinguals (English learned through classroom exposure, no English-speaking country residence before 18). To probe both explicit and implicit knowledge of the English article system (both omission: *I saw cat.*, and substitution: *I saw a/the cat.*), participants were tested on acceptability judgement (AJ), self-paced reading (SPR), and oral production tasks. For the SPR and AJ tasks, we developed 48 scenarios (24 indefinite and 24 definite) with the critical item consisting of singular noun in anaphoric direct object position. For the oral production task, the participants were asked to watch a cartoon and retell the story as it unfolded. Mixed-effects models analysis showed, first, that L1 Mandarin L2 English speakers outperformed the L1 Croatian L2 English speakers on SPR and oral production tasks, with patterns similar to the L1 English group, suggesting that Mandarin may be on a path of grammaticalizing some of its markers of (in)definites. Second, we observed an asymmetry between the processing and production of the indefinite and definite articles in all groups in that there was a greater processing/production cost for the indefinite article. As such, the study's results better support the Morphological Congruency Hypothesis than the Competition Model, with evidence of detrimental effects of the L1 on L2 English article acquisition. Future research implications are discussed.

## References

Chen, P. (2004). Identifiability and definiteness in Chinese. *Linguistics*, 42(6), 1129–1184.

---

\*Speaker

†Corresponding author: jelena.oreilly@york.ac.uk

Jenks, P. (2018). Articulated definiteness without articles. *Linguistic Inquiry*, 49(3), 501–536.

Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning*, 61(3), 940–967.

Van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language learning: syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research*, 26(1), 43–74.

**Keywords:** syntax, processing, self, paced reading, oral production, articles

# Learning a Second Language at First Exposure: The Role of Audiovisual Input

Paulina Olender <sup>\*† 1</sup>, Imma Miralpeix <sup>\* ‡ 1</sup>

<sup>1</sup> Universitat de Barcelona – Spain

A vast number of First Exposure (FE) studies have proved that novice learners are able to extract linguistic knowledge from minimal input (e.g. Rast, 2008; Ellis & Sagarra, 2010). This research has mostly been conducted in classroom or in laboratory settings, and there are very few studies that have explored audiovisual materials as input source. For instance, Gullberg et al.'s (2012) L1 Dutch participants learned new words from watching a short weather forecast in Chinese. Miralpeix et al. (2023) also reported vocabulary gains in L1 Spanish learners after watching a subtitled TV advert in Polish and showed that learning was significantly correlated with aptitude scores. As positive results have been found for learners when firstly exposed languages completely unknown, which were also typologically distant from their first languages (L1s), possibly longer videos in languages typologically similar to students' L1s could be a good way to start learning a second language (L2) from scratch. However, no attempts have been made to investigate whether and to what extent this could be an effective practice, even if audiovisual input (AV) is nowadays very easily accessible and our exposure to it increasingly common (Vanderplank, 2020).

The present study assesses what lexical knowledge adult novice learners can acquire when exposed to a 20-minute TV series episode (either with captions or without) in a language typologically close to their L1. In this case, 60 Spanish/Catalan L1 speakers watched the episode in Italian (a language they had not previous knowledge of) and were tested on word form recall, meaning recognition and recall of twenty target words, controlling for cognates. A content comprehension task was also included. Furthermore, a control group (N=16) that did not watch the episode took the vocabulary tests as well. Input and learner factors were considered, as both have been shown to have an effect on learning in previous research on AV input (e.g., Peters & Webb, 2018): the possible influence of word frequency and word relevance was analyzed and participants were tested on aptitude through LLAMA B, D, and E (Meara, 2005), as well as on working memory (WM) through a reading span test. Significant differences were found between control and experimental conditions on vocabulary learning, proving the positive effects of video viewing at FE, but not between captions and non-captions conditions. High comprehension scores were also achieved in the experimental groups, evidencing that the plot was followed irrespective of the presence of captions in the episode. Input factors were found to have a more important role on performance than learner factors: language aptitude had a small effect on comprehension and WM moderately affected form recall. However, frequency of word occurrence was determinant for word meaning learning. The study shows that, independently of the presence of captions, videos in a typologically similar language to the L1 provide input that allows for comprehension, while promoting significant vocabulary learning at FE in adult novice learners.

---

\*Speaker

†Corresponding author: polender@ub.edu

‡Corresponding author: imiralpeix@ub.edu

**Keywords:** first exposure, vocabulary acquisition, audiovisual input, frequency, relevance, working memory, language learning aptitude, learner factors, input factors

# The effect of regressive transfer from L3 Scandinavian on word order in L2 English

Anna Olszewska \*† <sup>1</sup>

<sup>1</sup> Adam Mickiewicz University in Poznań – Poland

While present research on TLA focuses on the influence of L1 and/or L2 on L3, less attention has been given to the reversely occurring effect, in which the subsequently acquired L3 impacts the previous L1 and L2. Currently, three factors possibly modulating regressive transfer have been proposed: age of acquisition (Cabrelli Amaro 2017), shared representation between L1 and L3 (Puig-Mayenco 2023) and typological similarity (Brown-Bousfield & Chang 2023). Research on L3-induced regressive effects remains however largely exploratory (Cabrelli 2023). Thus, this PhD project aims to address: Does the L3 exert influence over the previously acquired languages, most notably L2? If so, does the robustness of the effect change in relation to L3 proficiency? The study targets both syntactic and discourse-dependent phenomena that differentiate Mainland Scandinavian (MSc; L3) from English (L2) and Polish (L1), i.e., topicalization, V2 and adverb placement. While all three languages share the mere possibility of fronting different elements, they vastly differ in frequency with which it occurs. In MSc, non-subject-initial sentences constitute ca. 30-40% of all declaratives (Westergaard et al. 2023), while in English and Polish SVO is the most prevalent. Importantly, Scandinavian topicalization is linked to V2 and obligatory inversion, that English and Polish do not have. The languages also dissimilarly place adverbs in main clauses, allowing V-Adv in L3, but Adv-V in L1/L2.

To address the above-mentioned questions, a pseudo-longitudinal study will be conducted among L1 Polish L2 English (C1 level) university students, who acquire L3 Norwegian ( $n = 40$ ), Swedish ( $n = 40$ ) or Danish ( $n = 40$ ) from beginning to proficient (A2-C1) levels. In order to isolate the L3 influence, two control groups will be recruited: one of L1 Polish L2 English bilinguals ( $n = 20$ ) and one of L1 Polish monolinguals ( $n = 20$ ). All participants will be tested by means of an offline acceptability judgement task (AJT), measured on a 1-5 scale, and an online self-paced reading task (SPRT). The participants will be presented with sentences that are grammatical (*Nina lays the keys on the table*), grammatical yet pragmatically odd (*#The keys Nina lays on the table*), and ungrammatical (*\*The keys lays Nina on the table*). The last two will reflect the Scandinavian-like word order. To account for the possibly changing L1, a control AJT in Polish will be employed. Linguistic background and proficiency will be assessed through LHQ3 and DIALANG.

To date, both information structure and variable V2/V3 have been shown to be vulnerable domains for L2 learners. Thus, the study hypothesizes that regressive influence from L3 to L2 will become more robust with raising L3 proficiency and will (a) influence the way the L3 speakers organise information in their L2 and (b) induce V2 and V-Adv in L2 English grammar. The study will thus consider the majorly overlooked negative effect of L3 and serve as a further verification of the recent proposals.

Questions:

---

\*Speaker

†Corresponding author: a.olszewska@amu.edu.pl

1. Are there any methodological shortcomings?
2. How can potential differences between AJT and SPRT data be interpreted, when the regressive direction of influence is considered?

**Keywords:** L3 acquisition, regressive transfer, word order

# Frequency and Mutual Information effects on processing of lexical bundles: evidence from self-paced reading and priming tasks

Christopher Ott \* <sup>1</sup>, Patrick Rebuschat \*

<sup>1</sup>, Raffaella Bottini \*

1

<sup>1</sup> Lancaster University – United Kingdom

Lexical bundles (e.g., *at the same time, is one of the*) constitute an interesting subset of formulaic sequences that span phrasal boundaries and are often semantically and structurally incomplete. While previous psycholinguistic research on lexical bundles has shown a processing advantage for English L1 speakers (e.g., Tremblay et al., 2011), evidence for L2 speakers remains inconclusive. A possible underlying factor contributing to the mixed results is that psycholinguistic studies investigating lexical bundle processing have often used stimuli from pre-existing corpus studies (e.g., Jeong & Jiang, 2019), giving researchers little control over measures such as frequency and collocability. In this dissertation, we use a multimethod approach (Rebuschat et al., 2017), combining psycholinguistic and corpus methods, to investigate lexical bundle processing in non-native learners of English. At present, we are concluding a corpus study (Study 1 of the dissertation), which analyzes lexical bundles in the British Academic Written English (BAWE) corpus to be used as stimuli for a subsequent psycholinguistic study. In this presentation, we will focus on the latter study (Study 2 of dissertation). The aim of the study is to determine whether lexical bundles confer a processing advantage to L2 speakers and to investigate the effects of frequency and Mutual Information (MI) score measures on processing efficiency in English L1 and L2 speakers.

We plan to recruit approximately 40 native speakers of Mandarin and 40 native speakers of English. Participants will complete two tasks. In the first task (self-paced reading), participants will read two types of sentences: sentences with lexical bundles embedded and sentences with non-lexical bundle controls embedded. Participants will read portions of a sentence in sequential order. Processing advantages will be determined by the speed at which participants read the lexical bundle portion of the sentences in comparison to control/non-lexical bundle portions. In the second task (forward priming), participants will be shown the first three words of a lexical bundle (prime) and, after a short pause, the final word by itself (target). Participants will read the final word into a microphone when displayed. The response latency will determine any processing advantage lexical bundles may have. In both tasks, a multiple-regression model will determine effects for frequency and MI score. Based on previous research (Ellis et al., 2008), we predict that lexical bundles with a higher MI score will determine processing speeds for native English participants, while lexical bundles with higher frequency will determine processing speeds for non-native participants.

---

\*Speaker



Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 42(3), 375-396.

Jeong, H., & Jiang, N. (2019). Representation and processing of lexical bundles: Evidence from word monitoring. *System*, 80, 188-198.

Rebuschat, P., Meurers, D., & McEnery, T. (2017). Language Learning Research at the Intersection of Experimental, Computational, and Corpus-Based Approaches. *Language learning*, 67(S1), 6-13. <https://doi.org/10.1111/lang.12243>

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language learning*, 61(2), 569-613.

**Keywords:** psycholinguistics, processing, formulaic sequences, learner corpora

# WHEN and FOR WHOM does extracurricular English exposure work? Introducing new methods to capture the temporal specificity of L2 development and individual differences in young learners

Simone Pfenninger \* <sup>1,2</sup>, Mason Wirtz \* † <sup>3</sup>

<sup>1</sup> University of Zürich – Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

<sup>2</sup> University of Zurich – Switzerland

<sup>3</sup> University of Salzburg – Austria

Today’s multilingual and technology-supported culture is reshaping when, why, and how second languages (L2) are learned (Douglas Fir Group 2016), leading to diverse learning experiences and developmental processes (Peng et al. 2022). In particular, outside-school contact with English significantly influences outcomes for L2 English learners, expanding even very young individuals’ social networks beyond school boundaries (de Graff 2015). While prior research mainly examined extracurricular activities beneficial for L2 outcomes at specific points in time, this study takes a temporal-relational perspective, exploring the quantitative and qualitative relationship between extracurricular English exposure and L2 development over time. We report on findings from a study conducted in a bilingual (pre)primary school in Switzerland investigating the complex ways that L2 English development emerges over time in 45 children who received German/English bilingual instruction from age 5-12 in relation to emerging extracurricular exposure to English. The following RQs are addressed:

- Can we identify L2 supra-developmental patterns that transcend individual L2 developmental trajectories?
- To what extent can periods of rapid L2 development in these supra-developmental patterns quantitatively and qualitatively be explained by individual differences in extracurricular L2 use?

Data collection occurred 4 times annually, resulting in 32 data collection points for each participant and task over 8 years. At each measurement, participants (a) wrote an L2 English narrative; (b) completed an oral re-telling task; (c) filled in a questionnaire on their extracurricular English and a language awareness questionnaire with open-ended questions; and (d) participated in a 10-minute semi-structured interview to collect information about their use of English in daily activities, emotional experiences, and peer and family support.

---

\*Speaker

†Corresponding author: mason.wirtz@plus.ac.at

We proceeded as follows in the analysis: (1) Generalized additive mixed modeling to quantify individual-level L2 developmental trajectories; (2) time-series clustering methodology to identify L2 ‘supra’ developmental patterns; (3) visual-quantitative analyses to identify periods of rapid L2 development; and (4) qualitative content analysis. The analysis phases (3) and (4) were geared towards *explaining* trajectorial differences, homing in specifically on factors relating to extracurricular exposure to English.

The results revealed that heightened extracurricular English activities significantly impact rapid L2 development, primarily at the later testing stages (ages 9-12) when children start using English on social media, the Internet, and streaming services. The impact varies depending on the activity type and learner cluster, though. We relate the findings to the so-called “authenticity gap” (Henry 2013) between English inside and outside of school, as clusters of learners who perceived a discrepancy between in- and out-of-school encounters with English also showed rapid L2 development that was characterized by increased English exposure during extracurricular activities (rather than by English instruction inside of school). Our study demonstrates how the shift in students’ encounters with English has serious implications for teaching, e.g. in terms of how and when young learners’ knowledge of English owing to various out-of-school activities should be acknowledged and utilized in the classroom.

**Keywords:** young learners, individual differences, extracurricular input, bilingual education, research methodology, CDST

# Modality matters in measuring foreign language aptitude: Empirical evidence from a new aptitude battery Tests of Aptitude for Language Learning (TALL)

Junlan Pan <sup>\*† 1,2</sup>, Emma Marsden <sup>\*</sup>

2

<sup>1</sup> Chongqing University [Chongqing] – China

<sup>2</sup> University of York – United Kingdom

Foreign language aptitude is a componential construct conceptualised to predict and explain second language (L2) learning outcomes (Carroll, 1981). L2 learning, across naturalistic and instructed contexts, occurs in different modalities including auditory and written forms. This entails that aptitude components—like associative memory, language analytic ability, and executive control capacity in working memory— should perhaps be operationalised using stimuli in both aural and written modalities. Thus, aptitude assessments would ideally involve tests in both modalities to reflect different learning contexts. However, most major aptitude batteries only include written tests, or do not specify the modality of test items. The selection of modality in operationalising aptitude goes beyond methodological concerns: it holds theoretical significance in understanding the role aptitude may play in diverse learning contexts.

**Design and methods.** This study addresses the neglected issue of modality and its impact on the outcomes of aptitude tests. A new internet-based aptitude battery—*Tests of Aptitude for Language Learning* (TALL; Pan & Marsden, under review)—was developed. TALL integrates constructs from the theoretical frameworks of the Stages Approach (Skehan, 2016) and the Phonological/Executive Model (Wen, 2016) and its design is informed by previous aptitude batteries. It operationalises aptitude components representing: associative memory; phonetic coding ability; language analytic ability; and working memory (specifically, phonological short-term memory and executive control capacity). These components are measured by five subtests: Vocabulary Learning, Sound Discrimination, Language Analysis, Serial Nonwords Recall, and Complex Span Tasks, respectively.

Critically, TALL has both an aural and a written ‘suite’. To allow us to both examine within-subject effects of modality and to control for order of testing, each suite contains two versions of materials with counterbalanced test items.

165 college-level L1 Chinese-L2 English participants (aged 17 - 19) took two rounds of TALL tests: half taking the oral modality first and half the written modality, with a minimum of 30 days before taking the other modality. TALL, designed to reduce confounds with L2 proficiency, underwent scrutiny that demonstrated high levels of internal validity as a battery for measuring the components theorised to constitute the aptitude construct. Mixed-effects modelling was used

---

\*Speaker

†Corresponding author: junlan.pan@york.ac.uk

to analyse data from the three subtests that were administered in two modalities: Vocabulary Learning, Language Analysis, and Complex Span Tasks.

**Findings.** Participants consistently performed better when tested in the written modality compared to the aural modality. Furthermore, correlations of the scores from the two modalities were low (ranging between 0.16 to 0.32, in Kendall's *tau*). Importantly, the effects of modality varied in their impact on results for learners at different levels of aptitude abilities, suggesting that higher aptitude abilities may mitigate the effects of modality. Together, these results suggested substantial differences in aptitude constructs when measured in different modalities.

The study underscores the importance of considering modality as a moderating variable when examining aptitude in L2 learning, especially for learners with different aptitude levels. Additionally, it highlights that research could fruitfully investigate the role of modality-specific aptitude components in different learning contexts that may foreground the oral modality (naturalistic contexts) or written modality (instructed contexts).

**Keywords:** modality, foreign language aptitude, aptitude measurement

# The dimension of phraseological sophistication: Proposals towards more content validity

Magali Paquot <sup>\*†</sup> <sup>1,2</sup>, Hubert Naets <sup>3</sup>

<sup>1</sup> Centre for English Corpus Linguistics, UCLouvain – Belgium

<sup>2</sup> Fonds National de la Recherche Scientifique (FRS-FNRS) – Belgium

<sup>3</sup> Cental, UCLouvain – Belgium

Since Paquot (2019), several unresolved issues have persisted regarding the operationalization of phraseological sophistication in L2 complexity research. One of the most crucial concerns relates to the extent to which the commonly used measures of phraseological sophistication fully represent the intended construct (content validity). So far, phraseological sophistication has mostly been operationalized as average pointwise mutual information (PMI) scores (a measure of the strength of association between words) on the ground that, when calculated on a large reference corpus, this measure highlights word combinations that are highly exclusive. Because of their exclusivity, word combinations with a high PMI tend to be more topic-specific, have more distinctive meanings, and tend to involve more specialized vocabulary; these combinations are therefore considered ‘sophisticated’. However, it could be argued that association is only just one aspect of phraseological sophistication, which measurement today suffers from construct under-representation. In L2 vocabulary research, for example, indices related to word frequency (e.g. proportion of words that are not on the list of the 2,000 most frequent words), range, and specialized vocabulary (e.g. proportion of academic vocabulary) have been used to evaluate the lexical sophistication of learner language (e.g. Kyle & Crossley, 2015). Paquot (2019) also used a measure based on the *Academic Collocation List* (Ackermann and Chen, 2013) to operationalize phraseological sophistication but follow-up studies have not explored further this conceptualization of phraseological sophistication in terms of vocabulary specificity.

In this study, we drew upon insights from L2 vocabulary research to reexamine the conceptualization and operationalization of phraseological sophistication. We conducted new analyses on the VESPA corpus used in Paquot (2019), using alternative operationalisations of phraseological sophistication that represent distinct conceptualizations of sophistication in terms of frequency and vocabulary specificity. To this end, we developed new frequency lists of English general and academic collocations. Our aim was to investigate whether the consistent patterns noted in prior research are replicated when the construct is defined and measured using different linguistic operationalizations. Results provide robust empirical evidence for incorporating vocabulary specificity into the dimension of phraseological sophistication. In terms of concurrent validity, like the PMI-based measures used in Paquot (2019) and follow-up studies, measures of phraseological sophistication in relation to vocabulary specificity (academic collocations) have demonstrated their ability to distinguish at least between some of the CEFR groups that they should theoretically be able to distinguish (cf. Crossley et al., 2013). They have also exhibited correlations with related association-based measures of phraseological sophistication (convergent validity). By contrast, we were not able to make an argument for the concurrent validity and convergent validity of frequency-based measures of phraseological sophistication. We believe this

---

\*Speaker

†Corresponding author: magali.paquot@uclouvain.be

may at least be explained by the metrics implemented (collocations above an 80% text coverage in VESPA).

The reliability and validity of our methods and measurements should be the focus of more theoretical thinking and empirical evaluation in L2 research (e.g. Norris & Ortega, 2003). This study contributes to current research that seeks to evaluate the validity of the construct of phraseological complexity, and phraseological sophistication more specifically.

**Keywords:** learner corpus, phraseology, phraseological complexity, phraseological sophistication, measurement, validity

# The Cognate Facilitation Effect in L3 lexical processing

Kyle Parrish \* <sup>1</sup>

<sup>1</sup> Goethe-University Frankfurt am Main – Germany

The cognate facilitation effect refers to faster processing of words that are cognates between languages (e.g., words that have high orthographic, semantic, and sometimes phonological overlap). This is thought to occur, in part, because lexical items are represented in an integrated space, rather than being language-specific. Robust evidence has been reported for this effect across a variety of language combinations and a range of tasks in L2 research, which suggests that bilingual speakers have an integrated and non-selective lexicon (Costa et al., 2000). It has also been shown that this effect is additive in multilinguals, meaning that words which are cognates in three languages known by a trilingual are processed even faster than cognates shared between two of three languages (Szubko-Sitarek, 2011). While this suggests that L3 speakers have access to both L1 and L2 lexical representations, it is unclear whether the magnitude of access (measured by the cognate facilitation effect) is equal coming from L1 to L3 and L2 to L3, and how this relative impact of previously known lexical knowledge changes as L3 proficiency increases. It has been shown, for example, that a form-specific priming effect from both L1 and L2 lexical items was higher for low proficiency L3 speakers as compared to their high proficiency peers (González Alonso, 2012).

The present study examines this question using a lexical decision task by L3 Spanish speakers (L1 German- L2 English). Participants' L3 proficiency was measured using the Spanish LexTALE (Izura et al., 2014). During the lexical decision task, participants were presented with a word or a non-word in Spanish, and their task was to decide whether the presented item existed or not. Reaction times of these decisions were measured and compared across conditions. The lexical decision task consisted of 210 total words, in which 105 were words and 105 were non-words. Within the real words, 35 were L1 cognates (German-Spanish), 35 were L2 cognates (English-Spanish) and 35 were non-cognates. The present work considers a difference of 30ms to be the smallest meaningful difference ( $sd = 45$ ), and found that 35 tokens were necessary to achieve a power level of .8. Data collection is ongoing. The data will be analyzed using a Bayesian mixed effects linear regression model, in which log-transformed reaction time will be modeled as a function of condition (L1 cognate, L2 cognate or non-cognate), L3 proficiency, and lexical frequency, with a maximal random effects structure.

The results of this study will have important implications both for models of L3 acquisition and models of lexical access. If differences are shown in the cognate facilitation effect as a function of proficiency, or simply overall, this would challenge the notion of an integrated lexicon. L3 models would benefit from determining whether their predictions, such as language status effects and typological effects, can be extended to the lexicon.

(References removed due to word limit)

---

\*Speaker



**Keywords:** L3 acquisition, multilingualism, L3 lexicon

# The role of starting age for L2 English proficiency in input-rich settings: the case of Flanders and Sweden

Elke Peters <sup>\*†</sup> <sup>1</sup>, Pia Sundqvist <sup>\*</sup>

2

<sup>1</sup> KULeuven – Belgium

<sup>2</sup> University of Oslo – Norway

This presentation discusses the role of starting age of formal English instruction for English listening comprehension and vocabulary knowledge in two input-rich learning settings, Flanders and Sweden. It aims to contribute to the ongoing debate on early foreign/second (L2) instruction. An early start for L2 English instruction is on the rise in Europe, as most countries are now offering L2 classes in primary schools (Nikolov, 2009). Nevertheless, the research evidence for the potential benefits of an early start remains inconclusive. Some studies have found positive effects of an early start (e.g., De Graaff, 2015; Jaekel et al., 2022), while others failed to show any advantage for early starters compared to late starters (e.g., Baumert et al., 2020; Pfenninger & Singleton, 2017). However, most studies have been conducted in countries with relatively little out-of-school contact with L2 English (Muñoz & Cadierno, 2021; Schurz & Sundqvist, 2022). Still, little is known about the effects of early L2 instruction in input-rich learning settings, where learners regularly come into contact with English outside of school, often referred to as extramural English (Sundqvist, 2009).

The research contexts of this exploratory study are Flanders (Dutch-speaking region in Belgium) and Sweden. Unlike Sweden, where children receive their first English lessons in grade 1, 2, or 3 (age 7–9), children in Flanders start with formal English instruction in grade 7 (age 12–13) or 8 (age 13–14). Both regions are settings where L2 English learners have large amounts of extramural English (De Wilde et al., 2020; Peters et al., 2019; Puimège & Peters, 2019; Sundqvist, 2009; Sundqvist & Wikström, 2015).

Our study includes participants aged 11–13 (Younger Group,  $n = 139$ ) and 15–16 (Older Group,  $n = 52$ ) in Sweden and Flanders ( $N = 191$ ). The Flemish learners in the Younger Group had had no English instruction at the time of data collection. Data were collected through a questionnaire tapping into learners' extramural English, a listening comprehension test, the Picture Vocabulary Size Test (PVST, Anthony & Nation, 2017), and the Vocabulary Size Test (VST, Nation & Beglar, 2007). The VST was only administered to the Older Group.

The findings indicated that starting age had a large effect on learners' listening comprehension in the Younger Group with higher scores for the early starters (Sweden) ( $t = 10.019$ ,  $p < .001$ , Cohen's  $d = 1.839$ ), but not in the Older Group. Regarding vocabulary knowledge, there were no differences between the early and late starters in the Younger Group. However, the late starters (Flanders) significantly outperformed the early starters (Sweden) in the Older Group on the PVST and the effect was large ( $t = -3.255$ ,  $p = .002$ , Cohen's  $d = 0.903$ ). They also

---

\*Speaker

†Corresponding author: elke.peters@kuleuven.be

obtained higher scores on the VST, but not significantly so ( $t = -1.164$ ,  $p = .250$ , Cohen's  $d = 0.323$ , medium effect). The presentation concludes with some implications for L2 English instruction.

**Keywords:** starting age, extramural English, vocabulary, listening

# Grammatical gender assignment to unknown words in L2 French by speakers of three-gender languages

Francois Pichette <sup>\*† 1</sup>, Maura Cruz Enríquez <sup>1</sup>, Justyna Leśniewska <sup>2</sup>

<sup>1</sup> Université T luq – Canada

<sup>2</sup> Uniwersytet Jagielloński – Poland

Several studies have explored how gender is assigned to L2 words by L1 learners. However, little progress has been made since Lemh fer et al. (2010) described this process as "poorly understood." Research indicates three factors that affect gender assignment for L2 nouns: the tendency to assign masculine by default (Klassen & Liceras, 2014), transfer effects from the L1 (see Vanhove, 2017) and L2 word form (Seigneuric et al., 2007, Velni , 2020). Apparently, no research has compared the relative impact of those three factors. This study fills this gap by examining gender assignment in L2 French, a language whose gender system is challenging even for native speakers (Ayoun, 2018).

Our participants were 2,103 adult learners of L2 French, representing a wide range of proficiency (levels A1 to C2), whose L1 possesses a gender system with a neuter-feminine-masculine distinction. The L1 of the participants was either German (N= 1,148) or Polish (N=955). The two L1 groups show a comparable sociodemographic composition and breakdown across proficiency levels. In both groups, participants overwhelmingly declared English as their main other L2, a language which is considered non-influential in this study due to its lack of grammatical gender.

Based on grammatical gender in both L1 and L2, and on the gender bias of French word endings, four categories of words were selected for which each gender assignment can be explained by only one of the three factors. The test items were 24 French nouns presented sequentially online, in isolation, in the form of an image accompanied by a recording of the word's pronunciation. The items were extracted from the *Lexique3* database (New et al., 2004); selection criteria included word length, word frequency, and word ending. Cognates with German and Polish were avoided. Participants had to indicate which gender they thought the word possessed (masculine or feminine) and whether they already knew the word and its gender.

Our analysis centers on the unfamiliar words (as claimed by our participants), assessing whether the influence of the three above-mentioned factors varies with learners' L1s. A logistic regression reveals strong and significant differences between German and Polish speakers in assigning gender to French feminine and masculine words with feminine gender endings, with a coefficient of -0.758 ( $t = -15.0$ ;  $p < 0.0001$ ). Polish participants, despite having a more consistent and transparent L1 gender system (featuring specific word endings for each gender, with very few exceptions), appear to rely significantly less on word endings when assigning gender to unfamiliar words compared to their German counterparts, who, in contrast, L1 exhibit a marked inclination toward relying on word endings, as opposed esp. to masculine by default.

To explain this phenomenon, we speculate that an L1 gender system which relies on less overt cues (as in German) may foster a heightened sensitivity to less salient gender cues in the L2. The

---

\*Speaker

†Corresponding author: francois.pichette@teluq.ca

findings also stress the need for a nuanced understanding of both linguistic factors (typological distance and morphological cues) and extralinguistic factors (L2 instruction and metalinguistic knowledge) in language learning.

**Keywords:** L2 vocabulary acquisition, Gender assignment, French as a second language, word endings, L1 transfer

# Implicit statistical learning and working memory predict the outcomes of digital L2 practice in adolescents.

Diana Pili-Moss <sup>\*† 1</sup>, Katharina Wendebourg <sup>2</sup>, Torben Schmidt <sup>1</sup>,  
Detmar Meurers <sup>2</sup>

<sup>1</sup> Leuphana University of Lüneburg – Germany

<sup>2</sup> Eberhard Karls Universität Tübingen = Eberhard Karls University of Tuebingen – Germany

To foster a learner-centred approach to L2 learning, it is crucial that cognitive, memory-related individual differences are investigated in ecologically-valid contexts across the lifespan. The significant role of long-term memory, working memory and implicit statistical learning ability in computer-mediated artificial language learning has been widely reported in laboratory-based studies in the last twenty years (e.g., Misyak & Christiansen, 2012; Morgan-Short et al., 2014; Pili-Moss, 2021; Siegelman et al., 2017). Similar research for natural languages in educational environments has been scarce in comparison.

In the present study, 77 German EFL students (13-14 year-olds, three intact classes matched for type and amount of classroom instruction) used a digital platform (The FeedBook; Meurers et al., 2019) to individually practise questions in the present and past simple during the course of two weeks. The digital exercises were fully integrated with the syllabus and designed to provide a pre-communicative phase in preparation for a classroom-based written communicative target task performed at the end of the practice period.

Platform-administered matching pretest and posttest (30 items each), as well as the collaborative written target task, were used to respectively assess grammar accuracy gains and accurate use of questions in classroom-based communicative interaction. Measures of implicit statistical learning ability, explicit learning ability and working memory were collected online in laboratory conditions by respectively administering a visual statistical learning task (Siegelman et al., 2017), a matching explicit sequence learning task (Kenanidis et al., in press) and a visual Backward Span task.

The results of mixed-effect models elucidated the relationships between cognitive predictors and outcome measures (language gains at posttest and accurate question use in the written task). Although explicit L2 learning strategies are known to play an increasingly important role in older children and adolescents, explicit sequence learning was not found to significantly predict L2 development. Instead, the analysis revealed that L2 outcomes in both measures were independently predicted by a significant positive interaction between implicit statistical learning ability and working memory, thus highlighting a key role for these cognitive abilities in L2 digitally-based practice in early adolescence.

## References

Meurers, D., Kuthy, K. D., Nuxoll, F., Rudzewitz, B., & Ziai, R. (2019). Scaling up inter-

---

\*Speaker

†Corresponding author: [diana.pili-moss@leuphana.de](mailto:diana.pili-moss@leuphana.de)

vention studies to investigate real-life foreign language learning in school. *Annual Review of Applied Linguistics*, 39, 161–188. <https://doi.org/10.1017/S0267190519000126>

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, 62, 302–331. doi:10.1111/j.1467-9922.2010.00626.x

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, P. C. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, 17, 56–72. <https://doi.org/10.1017/S1366728912000715>

Pili-Moss, D. (2021). Cognitive Predictors of Child Second Language Comprehension and Syntactic Learning. *Language Learning*, 71(3), 907-945. <https://doi.org/10.1111/lang.12454>

Siegelman, N., Bogaerts, L., & Frost, R. (2017b). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavioral Research Methods*, 49, 418–432.

**Keywords:** implicit statistical learning, explicit learning, working memory, computer, mediated language learning, adolescents, environmentally, valid instructed contexts

# The acquisition of object clitics pronouns in child L3 French

Mihaela Pirvulescu <sup>\*† 1</sup>, Virginia Hill <sup>\*</sup>

2

<sup>1</sup> University of Toronto Mississauga – Canada

<sup>2</sup> University of New Brunswick – Canada

This paper focuses on the acquisition of object clitic pronouns in French as a third language (L3) by 31 children (ages 8-17) who grow up in Toronto and are exposed to Romanian since birth at home (HR - heritage Romanian). These children use English outside their home (their dominant language) and are learning French (L3) in a school immersion system. While Romanian (HR) and French (L3) are clitic languages, English has strong pronouns. The context of L3 is interesting for the investigation of the acquisition of clitics as there are two possible languages that may influence it and several competing models attempt to predict the source of influence. Objective. Our goal is to identify the role of transfer in the context of L3 acquisition. The knowledge of the pronominal domain is investigated in all languages of the participants. Background. Previous research shows that there can be influence from either (or both) background language over the L3 depending on factors such as language typology (Rothman et al. 2019; Westergaard et al. 2017), language proficiency (Tremblay, 2006) or cognitive associations (Bardel & Sanchez 2017 for L2; Hermas 2015 for L1). Questions. How are pronominal clitics learned in French L3, is there divergent production and interpretation of clitic forms? If yes, does transfer play a role? If so, what is the source of transfer? **Methodology.** Both production (Clitic Elicitation task) and comprehension (Picture Choice Task) of pronominal object clitics were measured, in L3 and HR. The production of pronouns was also measured in English. We used a Working Memory Test, a language questionnaire, receptive and production measures for overall language abilities. Results. In HR, clitic pronouns are correctly located. In English, strong pronouns are also correctly located in the post-verbal field. In L3 however, pronouns are produced in clitic positions in clitic form, or post-verbal in what appears to be strong pronoun forms (similar results discussed in Prévost 2009). Interestingly, in our data, the post-verbal forms are used in nominative. In terms of clitic comprehension children are performing in a similar manner in L3 and HR. Among the predictor variables, age of acquisition and the amount of French in the current year are significant factors for better performance ( $p < .001$ ,  $r_s = -.642$ , and  $p = .020$ ,  $r_s = .480$ , respectively). Analysis. Data shows that the parameter setting for object clitic pronouns is well set in the previous languages, i.e., positive for HR and negative for English. In French L3 the object clitic parameter is also well set, since object clitics are correctly located. However, there is transfer of use of post-verbal pronouns. Our results show that both background languages transfer to the L3. Specifically, HR, as a clitic language, seems to positively transfer this aspect to French, as all object clitic forms are well placed in L3 (see similar results in Grüter & Grago, 2012, Spanish L1). There is also negative transfer of case syncretism from HR, in the context where both HR and English have strong forms in object position.

---

\*Speaker

†Corresponding author: ma.pirvulescu@utoronto.ca



**Keywords:** French as L3, multilingual young learners, cross, linguistic influence, pronominal clitics

# Maximizing L2 learning from TV viewing in different conditions

Margarita Popova <sup>\*† 1</sup>, Imma Miralpeix <sup>‡ 2</sup>

<sup>1</sup> Universitat de Barcelona – Spain

<sup>2</sup> University of Barcelona – Spain

Watching videos, films and TV series has become a usual practice to learn a second language (L2) in different learning contexts, especially in autonomous learning, and several studies have proven its effectiveness (Reynolds et al., 2022). However, research has shown that it is hard to learn incidentally from TV viewing (Gesa & Miralpeix, 2023) and that acquiring a word involves numerous encounters with it in different contexts (Rodgers & Webb, 2020). Therefore, it is important to find ways to maximize learning from viewing. Research so far mostly comprises experiments with subtitles (L1 on-screen text) and captions (L2 on-screen text), often ignoring repeated viewing (with few exceptions, e.g., Alm, 2021; Muñoz et al., 2022), or the use of computer-assisted technology to enhance learning from viewing. For instance, only a few studies (e.g., Goulet et al., 2020; Fievez, 2023) have been published on the Chrome extension ‘Language Reactor’, an application providing individually regulated instruments when watching videos, such as simultaneous L1+L2 subtitles, word translations, pausing and speed control, replay function and flexible navigation between subtitles. To our knowledge, no studies have compared the use of this toolbox with other viewing conditions, keeping exposure time constant.

The current mixed-method study investigates whether L2 learning from TV viewing can be maximized with teenagers when using this edutainment application, in comparison to repeated viewing without using this app. An online intervention was conducted to assess comprehension and vocabulary gains from video viewing in L1 Russian (upper)-intermediate adolescents (N=40, mean age 15.5), who learned English as a foreign language. The learners watched the two first episodes of a Netflix series in English in two conditions: (1) Repeated Viewing with captions from the teacher’s screen (RV Group, N=19) and (2) Individual viewing with Language Reactor, with full access to all its tools (LR Group, N=21). After each of the two sessions, both taking 50 minutes, all participants took comprehension and vocabulary meaning recall tests with 20 keywords per episode. Then, students filled in post-viewing questionnaires, addressing the way they approached the task, as well as their perceptions and attitudes towards learning from viewing the episodes.

Results showed that TV viewing in both conditions was beneficial for comprehension and vocabulary learning (mean relative gains in word meaning recall were up to 51% per episode). The absence of significant differences between the groups suggests that, in the short run, the effect of repeated viewing is comparable to that of more individualized watching with LR, provided that watching time is the same. Qualitative analysis of learners’ behaviour while watching the videos showed that the RV Group paid more attention to the content, strongly relied on captions and perceived the first and the second viewing differently. On the contrary, LR participants explicitly focused on linguistic issues and appreciated pausing, replaying some parts and accessing the

---

\*Speaker

†Corresponding author: margaritona.p@gmail.com

‡Corresponding author: imiralpeix@ub.edu

translations the app provides when clicking on words in the subtitles.

**Keywords:** audiovisual input, edutainment, repeated viewing, vocabulary learning

# Processing Gender Stereotypes in L2 learners

Joanna Porkert \* <sup>1</sup>, Anna Siyanova-Chanturia <sup>2</sup>, Merel Keijzer <sup>1</sup>

<sup>1</sup> University of Groningen – Netherlands

<sup>2</sup> Te Herenga Waka - Victoria University Wellington – New Zealand

Research has shown that children being brought up bilingual show less racial bias than their monolingual peers (Singh et al., 2019; 2020). In addition, Singh et al. (2021) found that implicit racial bias was predicted by better cognitive flexibility in bilinguals. It raises the question whether this can also be found in young adults who are either sequential or simultaneous bilinguals, respectively. Moreover, it would be interesting to investigate whether other aspects of social cognition, such as gender stereotyping, can also be predicted by cognitive flexibility. Psycholinguistic EEG research using the event-related brain potential (ERP) technique has shown that gender stereotypes are a relevant cue during language processing, for example, when establishing coherence in pronoun resolution, such as in "Our aerobics instructor gave *himself* a break." (p. 284, Osterhout et al., 1997). Findings have shown that in the case of pronoun resolution, participants tend to perceive gender stereotype violations as grammatical mistake, even though the sentences were grammatical, by showing a P600 effect (typically elicited for morphosyntactic violations). For pronoun resolution in general, individual differences were observed on whether a P600 or Nref (a component typically elicited for ambiguous pronouns) is elicited based on working memory abilities (Nieuwland & van Berkum, 2006). The aim of this project is to investigate gender stereotype violations during pronoun resolution of simultaneous and sequential bilingual young adults with different L1 backgrounds. A secondary aim of this project is to examine individual language processing patterns as past gender stereotype studies found, for example, that participants who score low on sexism tend to show reduced amplitudes in the course of the ERP experiment (Grant et al., 2020). Recent research has also shown that interesting individual ERP patterns may become lost when only looking at ERP effects on a group level (Grey, 2023).

We tested 66 university students who were either Dutch or Spanish L1 speakers, using an English reading task. All of our participants were sequential English L2 learners, and around half of our participants were, in addition, Dutch-Frisian (n = 16), or Spanish-Catalan (n = 14) simultaneous bilinguals. The stimuli sentences were sentences with both personal, and reflexive pronouns as critical words (e.g., "The lifeguard threw *himself* into the water."). The collected background variables consisted of for instance the language and social background, the LexTale English proficiency score, but also scores in ambivalent sexism (a scale that measures hostile and benevolent sexism), in social dominance (the preference of strong hierarchies in society), or attitudes towards gender equality. Performance in working memory was measured with a Corsi blocks task, and cognitive flexibility with a color-shape switch task.

We hypothesize that abilities in working memory will affect how the pronoun violations will be processed, with higher ERP amplitudes for reflexive in comparison to personal pronouns. We expect that attitudes towards sexism and gender equality will lead to higher and more stable ERP amplitudes. Additionally, we think that cognitive flexibility will correlate with gender stereotyping, and that simultaneous bilinguals will show higher cognitive flexibility.

---

\*Speaker

**Keywords:** language processing, EEG/ERP, psycholinguistics, gender stereotyping, bilingualism

# Learning English beyond the classroom: Young EFL learners' grammaticality judgement and writing proficiency prior to formal English instruction

Elieen Prophète <sup>\*†</sup> <sup>1</sup>, Pia Sundqvist <sup>2</sup>, Elke Peters <sup>1</sup>

<sup>1</sup> KULeuven – Belgium

<sup>2</sup> University of Oslo – Norway

In many European (subtitling) countries, young L2 learners have ample opportunity to engage with English beyond the classroom setting. Studies have not only reported learners' extensive amounts of extramural English (EE; Sundqvist, 2009), they have also consistently demonstrated positive relationships between EE and several language proficiency measures, including, for example, vocabulary knowledge (e.g., Hannibal Jensen, 2017; Jóhannsdóttir, 2018; Puimège & Peters, 2019), listening proficiency and reading proficiency (e.g., De Wilde et al., 2020; Lindgren & Muñoz, 2013). These positive associations were found with learners who had already received English lessons (e.g., Hannibal Jensen, 2017) and with those who had not (e.g., De Wilde et al., 2020; Puimège & Peters, 2019).

The present study aims to extend existing knowledge on the link between young learners' EE and language proficiency by focussing on grammaticality judgement and writing, two language proficiency facets that have received little attention. More specifically, we investigate the relationship between different EE activities (e.g., watching TV, using social media, gaming) and learners' grammar knowledge and writing skills. We collected data from 445 children ( $N = 244$  monolingual Dutch;  $N = 176$  multilingual(1)) in the final year of primary school (age 11-12) who had not received formal English instruction. The participants completed an English writing test (European Commission, 2012) and a grammaticality judgement task (Pfenninger, 2014). A questionnaire was used to tap into learners' EE and their language backgrounds.

Our data indicated high levels of EE among participants, especially activities related to music, games, social media and YouTube videos. The results showed that learners had acquired grammar knowledge ( $M = 52.4$ ,  $SD = 20.90$ , max. = 98). There was considerable variance in the writing test scores ( $M = 11.41$ ,  $SD = 3.62$ , max. = 32), as some participants were not able to write in English while others managed to produce a short English text. Multiple regression analyses revealed that EE significantly predicted both grammar knowledge ( $R^2 = .37$ ,  $F(4, 364) = 53.36$ ,  $p < .001$ ) and writing proficiency ( $R^2 = .17$ ,  $F(4, 254) = 13.27$ ,  $p < .001$ ), after controlling for parental education, cognate linguistic distance and multilingual status. The strongest predictors for grammaticality judgement were watching TV without subtitles ( $\beta = .19$ ,  $p < .001$ ), YouTube videos ( $\beta = .27$ ,  $p < .001$ ), social media use ( $\beta = .18$ ,  $p < .001$ ) and speaking and writing in English ( $\beta = .22$ ,  $p < .001$ ). The strongest gains for writing were found from watching YouTube clips ( $\beta = .18$ ,  $p = .017$ ) and social media ( $\beta = .21$ ,  $p = .004$ ). Our findings demonstrate that grammar skills and writing can be acquired outside the classroom, through mere exposure. The results will be interpreted in light of usage-based theories which

---

\*Speaker

†Corresponding author: elien.prophete@kuleuven.be

emphasize the role of input for L2 acquisition.

(1) 25 responses were missing.

**Keywords:** grammar, writing, input, extramural English, young learners

# Eye-tracking evidence for pretesting effects in incidental L2 vocabulary learning

Eva Puimege \* <sup>1</sup>, Takumi Uchihara <sup>2</sup>

<sup>1</sup> KULeuven – Belgium

<sup>2</sup> Tohoku University – Japan

Incidental vocabulary learning is one of the most extensively researched topics in SLA (e.g., Uchihara et al., 2019). A common design to measure the effects of incidental learning activities (e.g., reading, listening, TV viewing), is the pretest-posttest design. Using pretests allows for the control of various aspects of lexical knowledge prior to a learning intervention. However, pretesting may also trigger inadvertent learning effects that could affect findings from intervention studies (Swanborn & de Glosper, 1999). Currently, little is known about the scale or cognitive nature of pretesting effects in L2 vocabulary research.

Cognitive research has found that pretesting information can boost posttest performance for a variety of learning materials and test formats, including paired associates and trivia questions (Chan et al., 2018). One explanation for this is that pretesting alters learners' attention to difficult or test-relevant information (e.g., Soderstrom & Bjork, 2014). These attentional effects, and how they might moderate learning outcomes, have thus far not been empirically tested in incidental L2 vocabulary learning. In line with effects of pre-teaching and previewing (Elgort et al., 2022; Pellicer-Sánchez et al., 2020), vocabulary pretests may alter learners' overt attention to novel words during L2 reading, for example leading to longer fixation times on novel words. Furthermore, the test format might alter learners' attention to semantic or orthographic information. These attentional shifts could affect posttest scores for pretested word knowledge aspects, compared to non-pretested information.

The current study examines the effects of two pretest formats on learners' online processing and learning of new L2 vocabulary from meaning-focused reading. Eighty-five Japanese university students learning English (38 female, Mage = 20.64) took part in an eye-tracking experiment. They were quasi-randomly assigned to a pretested group ( $n = 42$ ) or a non-pretested group ( $n = 43$ ). During the experiment, they read an English story while their eye movements were recorded. The story contained 12 pseudowords, each occurring 3 times in the reading text. Learners in the pretested group completed a form recognition pretest (yes-no format) and a meaning recall pretest (L2-L1 translation). Target words appeared in one of the formats in a counterbalanced fashion.

The preliminary results ( $n = 55$ ) suggest that pretesting affected learners' online processing of the pseudowords. Specifically, the number of fixations ( $B = 0.13$ ,  $p = .06$ ,  $d = 0.21$ ), and odds of rereading ( $B = 0.33$ ,  $p = 0.17$ ,  $d = 0.77$ ) were higher in the pretested group. Within the pretested group, dwell times ( $B = 0.12$ ,  $p = .007$ ,  $d = 0.19$ ) and fixation counts ( $B = 0.11$ ,  $p = .01$ ,  $d = 0.17$ ) were significantly higher for words pretested in the form recognition format. Posttest scores did not differ significantly between the two groups. Within the pretested group, learning gains were higher for words pretested at the meaning recall level ( $B = 0.72$ ,  $p = .009$ ,  $d = 1.13$ ). The results will be discussed in relation to theoretical accounts of pretesting effects and their application in L2 word learning experiments.

---

\*Speaker



**Keywords:** vocabulary, pretesting, incidental learning, eye, tracking

# The effectiveness of bilingual subtitles for foreign language learning and comprehension in beginner learners

Georgia Pujadas <sup>\*† 1</sup>, Carmen Muñoz <sup>1</sup>, Stuart Webb <sup>2</sup>

<sup>1</sup> Universitat de Barcelona – Spain

<sup>2</sup> University of Western Ontario – Canada

Bilingual (or dual) subtitles have been commonly used in multilingual areas for quite some time, but research has only recently started to investigate their effectiveness as a language learning tool. Bilingual subtitles, which display the L1 and L2 subtitles simultaneously, potentially combine the benefits of both types of subtitles by presenting the L2 forms of unknown words together with the L1 translations. This format allows learners to decide where to focus their attention according to their needs, which might be especially useful in classroom contexts with a disparity of proficiency levels. Studies also concur that the simultaneous presentation of the L1 and L2 does not result in cognitive overload (e.g. Wang & Pellicer-Sanchez, 2022). However, to date, the scarce research on bilingual subtitles has reported inconclusive results for comprehension and vocabulary learning. While some studies have found bilingual subtitling more effective for content comprehension (Dizon & Thanyawatpokin, 2021; Wang, 2019), others have reported no significant advantage over other subtitling types (Hao et al., 2021; Lwo & Lin, 2012; Wang & Pellicer-Sanchez, 2023). For vocabulary learning, bilingual subtitles appear to be superior for the acquisition of meaning (Li, 2016) compared to monolingual subtitles, but not for word-form learning (Wang & Pellicer-Sanchez, 2022). Additionally, studies have generally been carried out with adult learners and in contexts where bilingual subtitles are commonly used, but little is known about their value for younger, beginner learners with no experience with bilingual subtitling.

The present on-going study investigates the effectiveness of bilingual subtitles for viewing comprehension and vocabulary learning (compared to other forms of subtitling) with young learners who are unfamiliar with this subtitling mode. Participants are 150 Catalan-Spanish beginner learners of L2 English in Grade 8 (aged 13) from a secondary school in Spain, and the study is embedded as part of a regular English course. Four Grade 8 classes will view four age-appropriate 20-minute documentaries, each one in a different subtitling condition: bilingual subtitles, L1-subtitles, L2-subtitles and no-subtitles. The combination of documentary and subtitling condition will be counterbalanced across classes. Initial proficiency and vocabulary knowledge have been assessed at pre-test time, together with prior knowledge of target words. Vocabulary gains and comprehension will be assessed after each documentary. Results and pedagogical implications will be discussed.

Dizon, G., & Thanyawatpokin, B. (2021). Language learning with Netflix: Exploring the effects of dual subtitles on vocabulary learning and listening comprehension. *Computer Assisted Language Learning Electronic Journal*, 22(3), 52-65.

Liao, S., Kruger, J. L., & Doherty, S. (2020). The impact of monolingual and bilingual subtitles

---

\*Speaker

†Corresponding author: georgia.pujadas@ub.edu

on visual attention, cognitive load, and comprehension. *The Journal of Specialised Translation*, 33, 70-98.

Wang, Y. (2019). Effects of L1/L2 captioned TV programs on students' vocabulary learning and comprehension. *Calico Journal*, 36(3), 204-224.

Wang, A., & Pellicer-Sánchez, A. (2022). Incidental vocabulary learning from bilingual subtitled viewing: An eye-tracking study. *Language Learning*, 72(3), 765-805.

Wang, A., & Pellicer-Sánchez, A. (2023). Examining the effectiveness of bilingual subtitles for comprehension: An eye-tracking study. *Studies in Second Language Acquisition*, 45(4), 882-905.

**Keywords:** bilingual subtitles, viewing comprehension, vocabulary learning, young learners, audio-visual input

# Reading adjacent and non-adjacent collocations: An eye-tracking investigation of "nested" collocations

Manuel Pulido \* <sup>1</sup>, Marijana Macis <sup>2</sup>, Suhad Sonbul \*

3

<sup>1</sup> Pennsylvania State University – United States

<sup>2</sup> Manchester Metropolitan University – United Kingdom

<sup>3</sup> Prince Sultan University [Riyad] – Saudi Arabia

**Background:** Research in the last two decades has shown that collocations (i.e., statistically associated word combinations) shape language processing in important ways. A growing body of evidence shows that words that form a predictable sequence facilitate processing (Durrant & Doherty, 2010; Pulido, 2022; Siyanova-Chanturia et al., 2011, 2017). However, our current understanding of collocational processing is almost entirely restricted to immediately adjacent collocations. There is limited evidence suggesting that non-adjacent collocations facilitate L1 and L2 processing, but to a lesser extent than adjacent associations (Vilkaitė, 2016; Vilkaitė & Schmitt, 2017). However, no previous research has examined how non-adjacent collocations interplay with other (non)collocational intervening material. In the present study, we employed the novel concept of "nested" collocations, in which more than one element (verb, adjective) holds differential associations with the node (noun). To illustrate, in a Verb-Adjective-Noun (V-A-N) phrase, two collocations are "nested" if both V and Adj are associated with the noun ("influence behavior" + "violent behavior" = "influence violent behavior"); alternatively, only the verb (non-adjacent) or adjective (adjacent) might be collocational (see Examples below).

**Present study:** Using eye-tracking, the goal of the present study is to broaden our understanding of non-adjacent collocational effects in language processing, by investigating contexts in which two collocations are nested within a complex V-A-N phrase. By doing this, we go beyond previous research, which limited the scope of experiments to one single collocation, and we consider the simultaneous effects of non-adjacency, the role of intervening material, and the syntactic structure of a complex phrase. Using four conditions (Examples 1-4), we aim to replicate the non-adjacency V-N collocational effect vis-à-vis controls (Examples 1-2); and to compare processing of individual adjacent and non-adjacent collocations (Examples 2-3) relative to full "nested" collocations (Example 4). We specifically ask: (1) How is processing of non-adjacent collocations affected by the collocational status of intervening material in V-A-N phrases? (2) Is the collocational effect cumulative for two collocates (verb and adjective) priming the same noun? The materials included 39 quartets (following the quartet in Examples below), divided into four counter-balanced lists.

**Results:** Eye-tracking data from 46 L1 English speakers was analyzed using mixed-effects linear regression models for first-pass and total duration measures. Our results replicate the basic adjacent effect (Example 3 vis-à-vis Example 1). Importantly, we find an advantage for non-adjacent V-N collocations when intervening adjectives are non-collocational (Example 2 vs.

---

\*Speaker

Example 1). Further, this non-adjacent collocational V-N effect persists with an intervening collocational adjective (Example 4 vs. Example 3) (RQ1). Finally, we find a cumulative effect in "nested collocations" (Example 4), above and beyond the individual effects observed in adjacent and non-adjacent collocations (Examples 2-3) (RQ2). Altogether, the study extends our understanding of collocational effects beyond single collocations. We discuss the findings in relation to theories of L1 collocation processing and potential implications for L2 processing.

#### Examples

1. Control (VerbControl-AdjControl) "recommend vulnerable behavior"
2. Non-adjacent V-N collocation (VerbTarget-AdjControl) "influence vulnerable behavior"
3. Adjacent A-N collocation (VerbControl-AdjTarget) "recommend violent behavior"
4. Nested collocation (VerbTarget-AdjTarget) " influence violent behavior"

**Keywords:** collocations, processing, adjacency, eye, tracking, reading

# A by-product of being bilingual: Changes at the syntax-discourse interface in the L1 of L1 Spanish-L2 English bilinguals

Teresa Quesada \* <sup>1</sup>, Cristóbal Lozano <sup>2</sup>

<sup>1</sup> Universidad de Córdoba = University of Córdoba [Córdoba] – Spain

<sup>2</sup> Universidad de Granada = University of Granada – Spain

A key concern in Second Language Acquisition (SLA) is how the two languages of the bilinguals coexist and interact in a bidirectional manner: L1 > L2 and L2 > L1 being the latter is known as attrition (Schmid, 2010; Schmid & Köpke, 2017). Interestingly, such interaction seems to be more problematic when dealing with phenomena that involve two interfaces (Román & Gómez-Gómez, 2022). More specifically, the integration of syntactic and discursive information with referential expressions (REs) is a problematic area for bilinguals (Sorace, 2011). An ideal testing ground to investigate L1/L2 interaction is the processing of REs like null/overt pronominal subjects and repeated noun phrases (RNPs). In null-subject languages like Spanish, null pronouns are the privileged topic-continuity markers, irrespective of their syntactic configuration (1a), whereas overt pronouns would be pragmatically redundant. In non-null subject languages like English topic-continuity is marked via overt pronouns in all syntactic configurations, though null pronouns are allowed in coordination only (1b) (Quesada & Lozano, 2020). Thus, the question is whether the L1 of L1Spanish-L2English bilinguals would be modified after long exposure to the L2, in which native-like competence of REs is achieved (i.e., whether they would tolerate pragmatically redundant overt pronouns in topic-continuity).

(1) a. *Lucía*<sub>i</sub> interrumpió a *Pedro*<sub>j</sub> durante la reunión  $\emptyset$ <sub>i</sub> y discutió enérgicamente.  $\emptyset$ <sub>i</sub>/#*Ella*<sub>i</sub> no estaba de acuerdo con su idea.

b. *Lucy*<sub>i</sub> interrupted *Peter*<sub>j</sub> during the meeting and  $\emptyset$ <sub>i</sub> argued strongly. *She*<sub>i</sub> did not agree with his idea.

Additionally, previous corpus results show that several factors modulate the RE choice: the number of antecedents (e.g., one (*Lucy*) vs. two antecedents (*Lucy* and *Peter*)), and, crucially, the information status (i.e., topic-continuity contexts are more problematic than topic-shift).

Departing from this, we initially tested 24 advanced L1Spanish-L2English bilinguals in their L1 Spanish and a control group of 24 Spanish monolinguals. We focused on contexts where the two languages behave alike (topic-continuity and coordination) and conducted a self-paced reading task (SPRT) in Spanish to determine whether the processing of REs in their L1 Spanish was affected by their L2 English by manipulating the number of activated antecedents (1 vs. 2) and RE type (null/overt/RNPs).

Using mixed-effect models (Baayen et al., 2008), we found no interaction of number of antecedent by RE type, but we found a main effect of i) number of antecedents (longer reading times (RTs) with 2 than with 1 one antecedent); and ii) of RE, whose pattern is different according to group. Monolinguals show significantly shorter RTs for null than for overt and RNP, which suggests

---

\*Speaker

that null pronouns are the privileged topic-continuity marker in line with previous research. However, bilinguals processing their L1 show significantly shorter RTs for null and overt pronouns than for RNPs indicating that they privileged null and overt pronouns as topic-continuity markers, which goes against the Spanish monolingual norm. After corroborating these results with more participants, they would provide a valuable contribution to the field of SLA by showing that native-like attainment in the L2 entails changes in the L1 (attrition), which reflects the dynamism of being bilingual.

**Keywords:** Referential expressions, syntax, discourse interface, L1 Spanish, L2 English bilinguals, attrition

# Exploring a single coefficient of utterance fluency. Longitudinal analysis of speak production in French L2

Minerva Rojas \* <sup>1</sup>

<sup>1</sup> Université Côte d'Azur – Université Côte d'Azur (UCA) – France

Fluency in L2 is a field of research that continues to inspire the SLA scholars (e.g., Gilyuk et al., 2021 ; Huensch, 2023 ; Lintunen, 2019 ; Tavakoli et al., 2020). Much of the research focuses on the analysis of temporal variables in order to study the development of fluency. Their analysis provides information about the functioning of the cognitive mechanisms underlying speech production, and consequently the state of cognitive fluency (Segalowitz, 2010). One of the challenges of studying fluency is to establish concrete quantitative measures; this has led to an exhaustive list of units of analysis. For example, Segalowitz, French and Guay (2017) list a total of 23 potential fluency markers.

Fluency measures can be categorised into negative and positive measures (Osborne & Rutigliano, 2007). Negative measures, whose excess could indicate weaknesses in the automation of oral production, include the mean length of pauses, the frequency of pauses (*breakdown fluency* based on Skehan, 2003), and the frequency of repetitions and repairs (*repair fluency*). Positive measures (*speed fluency*) can be divided into those relating to the speed of speech (articulation rate, speech rate and phonation rate), and those relating to the length of production (mean length of runs and mean length of utterances) (Hilton, 2009, 2014).

Most studies on utterance fluency report results by unit of analysis, which provides a high degree of detail on the state of fluency in L2. However, this can also make it challenging for the researcher to report and interpret the results, and can also be difficult for the reader. Furthermore, very few publications deal with the calculation of an overall coefficient of fluency (Osborne, 2011), making it possible to combine the results of different measurements under a single coefficient expressed from 0 to 1 (1=fluency values comparable to an experienced speaker).

To explore the calculation of this coefficient in L2 French, we compiled an oral corpus of monologic production from 12 long-term immersion learners of French (30 months) and oral production from 22 proficient speakers of French, including 14 native French speakers and 10 non-native speakers living in France for more than 15 years. All participants performed the same tasks, 4 narrative tasks which, in the case of the learners, were collected in 4 data collections. The transcription, creation and analysis of the corpus were accomplished with EXMARaLDA (Schmidt & Wörner, 2009) and CLAN (MacWhinney, 2000). We analyzed 10 units of analysis to find the characteristic features of fluency in French L2. We then chose 4 units of analysis, 2 negative (hesitation rate and frequency of pauses) and 2 positive (articulation rate and MLR) in order to calculate the single fluency coefficient and trace the development of fluency for each learner. The mean results of the group of experienced French speakers (N=22) constitute the reference points for calculating the coefficient.

In our poster, we will present the methodology and the results of our study in order to show

---

\*Speaker



that the fluency coefficient makes it possible to trace the long-term evolution of oral fluency in FFL.

**Keywords:** learner corpus, fluency, French foreign language, study abroad, methodology

# ”Here comes the sun”. Differential pupillary embodiment of emotion and brightness vocabulary in L1 and L2

Stefanie Radetzky <sup>\*† 1</sup>, Dieter Thoma <sup>\*</sup>

2

<sup>1</sup> University of Mannheim – Germany

<sup>2</sup> University of Mannheim – Schloss EW 266 68163 Mannheim, Germany

Unbalanced bilinguals are less responsive to emotional stimuli in their L2 than L1. For example, their pupil dilates less in response to L2 words of high emotional arousal. While the evidence accrues, the causes of language-dependent emotions remain debated. Most authors argue that an L2 learned later in life, in more academic and less familial contexts has a persistent emotion disadvantage because the disjunct age-of-acquisition of emotions and L2 yields weaker semantic representations. The explanation links to embodiment theory, i.e., the idea that semantic learning mirrors physical perception. For example, when we hear words of brightness, the pupil constricts, irrespective of the physical luminance we see. Alternatives to learning-context explanations argue that L2 emotion semantics are less accessible because L2 words are processed less frequently.

Against this background, we investigated if language-dependent pupillary embodiment differs for emotion and brightness vocabulary. This is theoretically interesting because context-based theories predict *comparably* reduced L2 pupillary reactions to words of emotion and brightness. If the acquisition of emotion concepts and brightness concepts both co-occurred with learning their L1-word labels in childhood, their semantics should be equally more embodied in L1 than L2. In contrast, processing-based theories predict *differentially* reduced L2 pupillary reactions to emotion vs. brightness words. As non-immersive L2ers use abstract L2 emotion words more frequently than concrete L2 brightness words, pupillary reactions in L2 should be even more pronounced for emotion than for brightness, relative to L1.

Methodologically, we replicated and extended a monolingual study on the embodiment of brightness (Mathôt et al., 2017). We used a 2 (pupil embodiment: small — wide) by 2 (domain: brightness — emotion) by 2 (language: L1 German — L2 English) experiment with embodiment and emotion as within-subjects and language as between-subjects factors. Forty-nine unbalanced German-English bilinguals completed a go/no-go task on an eye-tracked screen. They read altogether 180 words presented individually in random order in 3 blocks. The task was to detect altogether 30 animal words among 30 brightness, 30 darkness, 60 neutral, and 30 emotion words. Linear mixed-effects models found a main effect of embodiment, such that the pupil was generally wider for words of emotion and darkness (see Figure 1). This embodiment was weaker in L2 than L1. Consistent with processing-based predictions, pupillary reactions were comparably sensitive to the neutral-emotional contrast and the bright-dark contrast in L1. In L2, however,

---

\*Speaker

†Corresponding author: stefanie.radetzky@students.uni-mannheim.de

pupil reactions were less sensitive to emotions and did not reflect the bright-dark contrast. We discuss implications for bilingual semantic processing as well as methodological challenges.

Figure 1. Pupil reactions indicating embodied emotion and brightness.

**Keywords:** embodiment, emotion, pupillometry, semantics

# Comparing maze task and judgement task data in crosslinguistic influence research: null and overt complementizers in L2 English relative clauses

Youssef Rami \*†<sup>1</sup>, Alaa Al-Maani \* ‡<sup>2</sup>, Heather Marsden \* §<sup>1</sup>

<sup>1</sup> Department of Language and Linguistic Science, University of York, York, United Kingdom – United Kingdom

<sup>2</sup> Department of English Language and Literature, Al al-Bayt University, Jordan – Jordan

Acceptability judgement tasks (AJTs) have been widely used in L2 research, yielding offline evidence of crosslinguistic influence. The Maze task (Forster et al., 2009) measures processing ease or difficulty in real time, but has, to date, been little used in L2 research. The present study investigates whether L1 influence effects found in an AJT are reflected in a Maze task targeting the same property.

The investigation exploits differences between English, Moroccan Arabic (MA) and French in relative clauses. English object relatives allow both null and overt complementizers, with definite and indefinite head nouns, as in (1a–d). However, in MA, the complementizer "lli" must occur with a definite head noun; with an indefinite, the complementizer must be null. French requires an overt complementizer "que" regardless of the definiteness of the head. In short, only (1b–c) would be grammatical in MA, and (1b) and (1d) in French.

(1)

- a. Sarah read the book her father bought. (DEFINITE\_NULL)
- b. Sarah read the book that her father bought. (DEFINITE\_OVERT)
- c. Sarah read a book her father bought. (INDEFINITE\_NULL)
- d. Sarah read a book that her father bought. (INDEFINITE\_OVERT)

Al-Maani (2020) used an AJT to investigate L2 knowledge of the paradigm in (1) by Jordanian Arabic speakers whose L1 behaves as MA. The results showed clear L1 influence: high acceptance of (1b–c); significantly lower acceptance of (1a) and (1d). We replicated Al-Maani's AJT, but also created a Maze version. In the Maze task, participants proceed through the sentence word-by-word, by selecting the word that makes sense from pairs of alternatives. For example, for (1a), after reading "Sarah", the sentence advances by selecting from each pair shown in (2). The time taken to select the correct word is measured.

---

\*Speaker

†Corresponding author: [yousseframi.nrl@gmail.com](mailto:yousseframi.nrl@gmail.com)

‡Corresponding author: [alaa.maani@aabu.edu.jo](mailto:alaa.maani@aabu.edu.jo)

§Corresponding author: [heather.marsden@york.ac.uk](mailto:heather.marsden@york.ac.uk)

2. (read / download) (the / sir) (book / applies) (her / sec) (father / myself) (bought / billion)

The critical word is "her". L1-English speakers should select this word with uniform speed in all conditions, whereas, if affected by L1 influence, L1-MA speakers are predicted to exhibit delay on "her" in (1a) and (1d) while L1-French speakers should exhibit delay in (1a) and (1c).

Participants to date are 44 intermediate-level L2-English speakers and 22 L1-English speakers (AJT: 11 L1-MA, 7 L1-French, 10 L1-English; Maze: 17 L1-MA, 8 L1-French, 12 L1-English). Results show clear between-group differences. On the 1–7 AJT scale, the L1-English ratings are  $> 6$  for all conditions, whereas the L1-MA and L1-French groups rated  $> 6$  only on conditions that would be grammatical in their L1s, and rated  $\leq 4.5$  elsewhere. In the Maze, L1-English speakers had uniform mean response times for the critical word across conditions, whereas L1-MA and L1-French speakers had strikingly longer response times in conditions that would be ungrammatical in their L1s ( $\geq 980$  ms v.  $\leq 858$  ms). Mixed effects models provide statistical support of L1-influenced behaviour.

Our findings show that the Maze task can capture L1-influence effects. Moreover, the combined results provide evidence from a novel crosslinguistic perspective of L1 grammar influencing both implicit processing and offline comprehension.

**Keywords:** crosslinguistic influence, relative complementisers, definiteness, comprehension, processing, maze task, acceptability judgement task, research methods

# Phonetic training meets statistical learning

Patrick Rebuschat \* <sup>1,2</sup>, Yuxin Ge <sup>3,4</sup>, João Dinis Fernandes <sup>3,4</sup>, Anabela Rato <sup>5</sup>, Susana Correia <sup>3</sup>

<sup>1</sup> Lancaster University – Bailrigg, Lancaster. UK LA1 4YW, United Kingdom

<sup>2</sup> University of Tübingen – Germany

<sup>3</sup> Linguistics Research Centre (CLUNL), NOVA University Lisbon – Portugal

<sup>4</sup> Lancaster University – United Kingdom

<sup>5</sup> UNIVERSITY OF TORONTO – Canada

In this paper, we report the results of a study that explored the impact of phonetic training on statistical learning. Recent research (e.g., Tuninetti, Mulak and Escudero, 2020) has shown that adult learners can rapidly acquire novel words of a foreign language by tracking cross-situational statistics, but also that learning is substantially reduced when the target words are phonologically similar, e.g. when they form pseudoword minimal pairs (Ge, Monaghan & Rebuschat, in press). Here, we expand on this research by investigating if perceptual discrimination training on non-native target sounds facilitates cross-situational learning of new words which overlap phonologically. To our knowledge, this is the first experiment to directly combine these research literatures.

Our target items were 24 Portuguese disyllabic pseudowords which were randomly associated with 24 novel and unfamiliar objects from Horst and Hout (2016). The linguistic focus was on four sound contrasts that are phonemic in Portuguese but not in English, the native language of our participants, namely the consonant contrasts /l/-// and /n/-// and the vowel contrasts /e/-// and /o/-//. The 24 pseudowords formed 12 minimal pairs. This manipulation allowed us to determine if and how phonological overlap between the pseudowords affected word learning. The study was preregistered, and the materials, anonymized data and R scripts are available on the OSF platform.

In the study, 68 native speakers of English were randomly assigned to one of the three conditions. All groups were trained on the 24 Portuguese pseudowords by means of a cross-situational word learning task (CSWL). In this task, participants were presented with two objects on the screen and played a single pseudoword (e.g., /palu/). Their task was to decide, as quickly and accurately as possible, which object the pseudoword referred to. No feedback was provided, and the next trial played once the response had been entered. The difference between the groups related to whether or not they received perceptual discrimination training (over two days) on our target contrasts before completing the CSWL task. One group (n = 22) was trained by means of an AX discrimination task, another (n = 24) by means of an oddity discrimination task. A third group (n = 22) received no phonetic training and just completed the CSWL task. Our results confirmed the effectiveness of perceptual training on the discrimination of L2 contrasts (e.g., Sakai & Moorman, 2017), with participants in the AX condition outperforming those in the oddity discrimination condition. The analysis of the CSWL task indicates that adults can learn non-native words from cross-situational statistics, and that phonological overlap between target words decreases learning (Ge et al., in press). However, there were no differences between groups in the CSWL task, and perceptual discrimination accuracy and word learning outcomes were not correlated. That is, perceptual training significantly improved participants' ability to discriminate the target sounds, yet this did not transfer to the learning of words that contain

---

\*Speaker

these contrasts. In our discussion, we focus on why accurate perception

**Keywords:** Implicit learning, statistical learning, cross, situational learning, vocabulary, phonology, phonetic training

# Focus last and Agent first in basic varieties : what happens when the agent is in the focus?

Isabel Repiso \*† <sup>1</sup>, Cyrille Granget \* ‡ <sup>2</sup>

<sup>1</sup> Paris-Lodron-Universität Salzburg – Austria

<sup>2</sup> Laboratoire de NeuroPsychoLinguistique – Université Toulouse - Jean Jaurès – France

One of the results of the research *The ecology of adult language acquisition* is that, at the basic stage, L2 discourse is made up of utterances governed by two principles: an informational structuring principle according to which focus information is in the last position (thus topical information, e.g. a protagonist, in the first position) and a semantic principle according to which the agentive entity is in the first position (Klein and Perdue 1997). In L2 discourse, these two principles converge when the topical entity is agentive: the topical agent is in first position, the focus information in last. But they come into conflict when the topical entity undergoes an action: the informational structuring principle predicts that the undergoing entity is in first position, whereas the semantic principle doesn't allow it. In L2, there are no studies devoted to the passive apart from Lee & Doherty (2019) and Gerolimich (2013). The first one shows a higher cost of processing passive than active utterances in L1 and L2 Spanish. Gerolimich (2013) is the only one examining production data, i.e. the linguistic means used in passive written contexts by Italian L1 learners of French L2 at a postbasic level. The study shows that Italian learners of L2 French use "patient first" utterances in 88% of the passive contexts. However, the study does not focus specifically on learners in the initial stages and in contexts of competition between semantic and informational principles. Our study aims to fill this gap by analyzing the means used by beginner to intermediate learners in the critical contexts mentioned above.

An 8-image description task was carried out with 37 beginner to intermediate learners of L2 French with L1 Spanish and German, known as having different syntactic means for passive sentences. In 4 images where the protagonist undergoes an action the informational and semantic principles guiding the basic variety predict that the topical entity is in first position "patient first" with regard to the informational structuring principle, but in last position with regard to the semantic control principle "patient last". Control data in French, Spanish and German L1 were used to measure the effect of L1 on the choice of syntax for the French L2 utterance.

The results show that in the critical contexts "patient first or last?", there are significant individual differences in the data analyzed in L1: patient first utterances are more numerous in French (56%) than in German (40%) and Spanish (8%). Nevertheless oral "patient first" French L2 utterances are significantly less numerous than in Gerolimich's (2013) study, accounting for only 8-9% of the utterances. For Spanish L1 learners, the semantic agent-first principle prevails over the informational topic-first principle, with 51 % of utterances in which the patient has a syntactic object function, while for German L1 learners, the informational and semantic principles are reconciled in a third way, where the patient is a topical entity in the first position of experiencer or proto-patient. We will show that these different L2 preferences can be explained

---

\*Speaker

†Corresponding author: isabel.repiso@sbg.ac.at

‡Corresponding author: cyrille.granget@univ-tlse2.fr



by first-language usage.

**Keywords:** basic variety, agent first, focus last, L1 Spanish, L1 German, French L2

# Perceptual curiosity in L2 learning: construct development and scale validation

Céline Rocher Hahlin <sup>\*† 1</sup>, Alastair Henry<sup>‡ 2,3</sup>, Joost Van De Weijer<sup>§ 2</sup>,  
Jonas Granfeldt <sup>\* ¶ 2</sup>

<sup>1</sup> Dalarna University – Sweden

<sup>2</sup> Lund University – Sweden

<sup>3</sup> University West – Sweden

## Background

Research into language learners' emotions has grown rapidly in recent years. However, despite many studies, special issues and conference symposia, emotions have rarely been examined beyond the language classroom (Myhre et al., 2022). While the range of emotions currently investigated is greatly expanded, constructs reflect this bias. Curiosity (Mahmoodzadeh & Khajavy, 2018) is one example. Curiosity is generating growing interest. However, conceptualization and investigation have been limited to the construct's epistemic dimensions, and emotions experienced in relation to L2 tasks and knowledge generation (e.g., Mahmoodzadeh & Khajavy, 2018; Nakamura et al., 2022; Takkaç-Tulgar, 2018). In contrast, perceptual curiosity involves interest in, and the direction of attention to novel stimulation, and motivates visual and sensory inspection (Berlyne, 1954). For languages rarely encountered beyond the classroom, a learner's perceptual curiosity can play an important role in generating motivation.

## Scale development, Method and findings

With a focus on foreign languages learnt in addition to English, and in a Swedish context, the scale was designed to capture perceptual curiosity when encountering a target language (French, German or Spanish) (i) in an online context (4 items), and (ii) in situations providing learning affordances (3 items).

The 7-item scale was developed and administered to 212 students in Sweden learning French, German or Spanish (languages rarely encountered beyond the classroom). On ordinal five-point scales, participants indicated the likelihood that curiosity would be aroused in various out-of-class situations. Responses varied across the entire range of the scale categories, with an internal consistency (Cronbach's alpha) equal to 0.80. All items had loadings larger than 0.5 on a single factor, and the overall fit of the model was good (Comparative Fit Index = 0.987).

---

\*Speaker

†Corresponding author: [crc@du.se](mailto:crc@du.se)

‡Corresponding author: [alastair.henry@hv.se](mailto:alastair.henry@hv.se)

§Corresponding author: [joost.van\\_de\\_weijer@humlab.lu.se](mailto:joost.van_de_weijer@humlab.lu.se)

¶Corresponding author: [jonas.granfeldt@rom.lu.se](mailto:jonas.granfeldt@rom.lu.se)

## Conclusion

In an age where L2 learning takes place in networks where the L2 classroom is but one node (Reinders & Benson, 2017), this study demonstrates the importance of perceptual stimulation. It suggests that a better understanding of L2 learners' curiosity can be gained when both epistemic and perceptual forms are investigated. In contexts beyond the classroom, and for infrequently encountered languages, the scale can provide an important means for measuring the learner's perceptual response to encounters with linguistic content.

## References

- Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology*, 45, 180–191.
- Mahmoodzadeh, M., & Khajavy, G. H. (2018). Towards conceptualizing language learning curiosity in SLA: An empirical study. *Journal of Psycholinguistic Research*, 48(2), 333–351.
- Myhre, T. S., Dewaele, J.-M., Fiskum, T. A., & Holand, A. M. (2022). Anxiety and enjoyment among young teenagers learning English as a foreign language outdoors: A mixed-methods study. *Innovation in Language Learning and Teaching*, 17(4), 827-844.
- Nakamura, S., Reinders, H. & Darasawang, P. (2022). A Classroom-Based Study on the Antecedents of Epistemic Curiosity in L2 Learning *Journal of Psycholinguistic Research* (2022) 51:293–308.
- Reinders, H., & Benson, P. (2017). Research agenda: Language learning beyond the classroom. *Language Teaching*, 50(4), 561–578.
- Takkaç-Tulgar, A. (2018). The effects of curiosity on second language learning in terms of linguistic, social-cultural and pragmatic development. *Journal of Educational Sciences*, 9(2), 59–72.

**Keywords:** Motivation, emotions, foreign language learning, curiosity, construct development

# Aptitude-treatment interactions in explicit instruction

Karen Roehr-Brackin <sup>\*† 1</sup>, Karolina Baranowska <sup>2</sup>, Renato Pavlekovic <sup>1</sup>,  
Pawel Scheffler <sup>2</sup>

<sup>1</sup> University of Essex – United Kingdom

<sup>2</sup> Adam Mickiewicz University in Poznań – Poland

Aptitude-treatment interaction (ATI) research is of both theoretical and practical interest to the field of second language (L2) learning (DeKeyser, 2012; Kidd et al., 2018). This is because it provides insights into the processes that link learner-internal individual difference factors and learner-external contextual factors including instructional approach – variables which jointly determine L2 outcomes. Recent ATI studies have tended to focus on the role of selected cognitive capacities, such as working memory, and have mostly manipulated specific instructional components, such as the nature of corrective feedback, or the type and timing of practice activities (see, e.g., chapters in DeKeyser, 2021). By contrast, there are no up-to-date studies which have compared ecologically valid instructional conditions as well as employed a full range of aptitude measures in accordance with current theorising (Li, 2022).

We addressed this gap by investigating the role of aptitude for explicit learning, aptitude for implicit learning, working memory and prior language learning experience in the context of different explicit instructional treatments comprising presentation and controlled practice activities. We developed a suite of four online language lessons in beginners' Polish that targeted two morphological features. The lessons were either deductive or inductive, and presented either in a single input modality (auditory or written) or in a mixed modality, resulting in four treatment conditions: auditory inductive, written inductive, mixed inductive, and mixed deductive. A total of 136 international adult volunteers with no prior knowledge of any Slavic languages completed the lessons and two post-tests. We administered the LLAMA test battery (Meara & Rogers, 2019) to measure aptitude for explicit learning, a probabilistic serial reaction time task (Kaufman et al., 2010) to measure implicit sequence learning ability, and an operation span task to measure executive working memory.

Our findings show that participants' phonetic and language-analytic abilities (Granena & Yilmaz, 2018; Li, 2015; Skehan, 2015, 2016), level of multilingualism and age predicted learning of the targeted morphological features. A cluster analysis identified four distinct learner profiles: high-aptitude, low-aptitude, memory-oriented and analytically oriented, mirroring previous work carried out under very different conditions (Ranta, 2002). We further found that deductive instruction seemingly neutralised individual differences in aptitude, thereby replicating previous results (Erlam, 2005; Li et al., 2019; Sanz et al., 2016). By contrast, ATI effects were observed in the single-modality conditions, with auditory input favouring high-aptitude learners and written input favouring high-aptitude, analytic and memory-oriented learners.

We discuss the theoretical and practical implications of these findings by highlighting the 'capital' afforded by prior language learning experience, over and above the role of cognitive ability. In addition to the inductive/deductive contrast in explicit instruction, we emphasise the importance of input modality which has hitherto been neglected in the field (but see Kim & Godfroid, 2019). Specifically, we argue that future work in an ATI paradigm should take into account

---

\*Speaker

†Corresponding author: kroehr@essex.ac.uk

the role of learners' perceptual abilities and/or preferences (auditory, visual), in addition to the role of cognitive abilities (memory, language analysis). In conclusion, we put forward some suggestions for instructional practice.

**Keywords:** aptitude treatment interactions, individual differences, language learning aptitude, explicit instruction, learner profiles

# Developments in LLAMA test validation

Vivienne Rogers \*<sup>1</sup>, Lars Bokander \*

<sup>2</sup>, Tesni Galvin<sup>1</sup>, Rhea Ray<sup>1</sup>, Brian Rogers<sup>1</sup>, Paul Meara<sup>1</sup>

<sup>1</sup> swansea university – United Kingdom

<sup>2</sup> Jonkoping University – Sweden

Language learning aptitude has received increasing attention in recent years. However, in the absence of freely available test batteries, researchers have turned to the LLAMA tests (Meara, 2005) despite issues with their lack of reliability and validation (Bokander & Bylund, 2020). The LLAMA tests have since undergone a number of significant changes in terms of items and test layout (Rogers, Meara & Rogers, 2023) but still comprise 4 tests: LLAMA D (sound recognition), LLAMA B (vocabulary learning), LLAMA E (sound-symbol correspondence) and LLAMA F (grammatical inferencing). With the new LLAMA v.3 tests in use by the research community, it is therefore timely to revisit the reliability and validity concerns highlighted by Bokander & Bylund (2020) with the new tests. This paper investigates some of these concerns in three distinct ways.

The first approach used LLAMA data taken from the lognostics website. The following inclusion criteria resulted in 1370 participants:

All incomplete tests removed,

Duplicate IDs – first complete test used,

Blanks and "anon" IDs removed,

IDs matched across four subtests.

We adopted the methods used in Bokander & Bylund (2020) and carried out internal consistency and item-analyses on the data. Internal consistency reliability (alpha, omega) was much improved compared to the first LLAMA version (Meara, 2005) in all subtests, with coefficients around .90 and just slightly lower in LLAMA D. Factor analysis of the test items revealed a two-dimensional structure in LLAMA D, reflecting the difference between familiar and unfamiliar sound stimuli. LLAMA B, E, and F displayed a one-dimensional structure, thus measuring single constructs. Classical item analysis and Rasch modelling showed that almost all items contributed to good measurement. The few that did not (four items in LLAMA D) will be subject to revision.

The second approach involved a test-retest method. 100 participants (50M, 50F) were recruited through prolific.com. Screeners included: L1 English, fluent in another language, education no higher than undergraduate degree and five age-groups (n=20) equally divided between 25-75

---

\*Speaker

( $M=49.7$ ,  $SD=14.42$ ). Participants took the LLAMA tests (reprogrammed into gorilla.sc) twice with 5 additional items in LLAMA F, four weeks apart (T1 & T2). Spearman's correlations with 94 participants showed significant medium correlations ( $p < .001$ ) between T1 and T2 (D:  $r_s=.478$ ,  $z=.520$ ; B:  $r_s=.570$ ,  $z=.648$ ; E:  $r_s=.577$ ,  $z=.658$ ; F:  $r_s=.654$ ,  $z=.782$ ). Linear regression showed no significant effect of age on any test.

The third approach looked at test-taking strategies and included qualitative interviews with 46 participants after taking the LLAMA tests in person. Interview transcripts were coded using a bottom-up approach in N.vivo (v.14). Strategies varied by task although story creation, linking to known things, guessing and rehearsal were common themes. LLAMA F strategies divided between participants identifying rules versus meaning. There was no consensus on which task was easiest or hardest.

Overall, we argue that the new LLAMA v.3 tests are an improvement over the original versions and would encourage researchers to use these versions, with the caveats mentioned by Meara (2005) and Rogers, Meara, & Rogers (2023) in mind

**Keywords:** aptitude, LLAMA tests, validation

# Acquiring grammar through input-based tasks: a quasi-experimental study with beginner learners of Italian in Austria

Anna Romano \* <sup>1</sup>

<sup>1</sup> University of Innsbruck, Department for Subject-Specific Education (Unit of Language Education) – Austria

Previous studies focusing on the application of Task-Based Language Teaching (TBLT) to beginner learners have shown encouraging findings regarding the acquisition of grammar knowledge (Erlam & Ellis, 2019; Shintani, 2016; Shintani & Ellis, 2010). However, results regarding the benefits of TBLT beginners' classes for the productive use of grammatical forms remain unclear, leaving this aspect a research desideratum.

Furthermore, there is a lack of research in this field concerning authentic classroom settings and foreign languages other than English (Bygate, 2020; Ellis, 2020; Erlam & Tolosa, 2022). To date, there are no studies on TBLT instruction for beginner learners of Italian. Gaining more insights into the teaching of Italian as a foreign language is particularly important for the Tyrolean context (Austria) investigated in this ongoing project, where Italian is taught in many schools as a second foreign language (after English). In addition, this Ph.D. project advocates for the advantages of TBLT in beginners' classrooms in general, where its potential has not yet been fully unlocked.

The above research gaps are addressed by investigating the effectiveness of a two-week TBLT program (4 lessons) for beginners of Italian in a real classroom setting in Tyrol. The grammatical focus of the study is on the plural marking of masculine and feminine nouns. The results of the TBLT classes are going to be compared with those of PPP (present – practice – produce) groups, making this project a comparative method study with the following research questions:

- Do beginner learners acquire receptive and productive knowledge of the Italian plural markers as a result of performing input- and comprehension-based tasks?
- Do beginner learners acquire receptive and productive knowledge of the Italian plural markers as a result of following output-based and explicit grammar teaching lessons?
- To what extent do the two groups differ?

A quasi-experimental study with four intact school classes ( $N = 60$ ), two of which receive the same type of instruction, is going to be conducted. The two experimental groups work with a task-based approach based on input- and comprehension-based tasks, while the other two groups function as control groups and are taught following a more explicit and output-based grammar teaching method (PPP-Method). All groups are going to take a pre-, post-, and delayed posttest

---

\*Speaker



on receptive and productive grammar knowledge. The receptive grammar knowledge is going to be assessed using a grammar listening task, the productive usage of the forms is going to be tested in an elicited imitation test and a spot-the-difference task. The latter is aimed at testing the production of the plural forms in an authentic communicative setting. The data is going to be analyzed quantitatively and qualitatively mainly on the basis of between-group comparisons. In addition to the grammar tests, a written questionnaire on the participants' language background and their contact with Italian outside the classroom is going to be conducted. My presentation for the doctoral workshops is going to focus mainly on the above presented design of the quasi-experimental study.

**Keywords:** grammar acquisition, TBLT, beginner learners, Italian as a foreign language, comparative method study, plural morphology

# The effects of CLIL intensity, extramural exposure and gender on young learners' reading and listening skills

Hanne Roothoof <sup>\*</sup> <sup>1</sup>, Amparo Lázaro Ibarrola <sup>\*</sup>

1

<sup>1</sup> Public University of Navarra – Spain

The number of content and language integrated (CLIL) programmes continues to grow in schools worldwide (Enever, 2018). At the same time, a large body of studies have compared the English proficiency of CLIL learners to those who only learn English as a foreign language, and the results mostly show an advantage for the former, both in primary and secondary school (e.g., Lasagabaster 2011; Lorenzo et al., 2010; Madrid & Barrios, 2018). However, conflicting findings have been found, especially with regards to receptive skills, and more particularly for the listening skill, which does not always seem to benefit as much from the increased exposure to English (Navés, 2011; Prieto-Arranz et al., 2015). Moreover, CLIL programmes can take many different shapes, and the number of hours devoted to CLIL differ widely between studies. Some studies have shown that only two hours of CLIL classes per week on top of the EFL classes already lead to increased English proficiency (Artieda et al., 2020), while in other studies CLIL learners receive 5 (Pérez-Cañado, 2018) or even 8 hours (Merino & Lasagabaster, 2017) of extra exposure to English. Only very few studies have compared different amounts of exposure to CLIL, or CLIL intensity, in order to determine how many hours are needed to make a difference (Merino & Lasagabaster, 2017). Previous studies have also received criticism for not considering possible moderating variables, such as socioeconomic status or extramural exposure to English (Bruton, 2013; Pérez-Cañado, 2016; Sunqvist & Sylvén, 2016). For these reasons, the present study compared two different CLIL programmes, a high intensity group receiving 5-7 hours of CLIL per week (n=85) and a low intensity one with 2 hours of CLIL (n=82), with a non-CLIL group (n=52) of primary school learners who only received 5 hours of EFL. The learners in the three groups were comparable in terms of socioeconomic status. They carried out a reading and listening test (Cambridge English: Flyers). Moreover, questionnaires were administered to gather information on the amount and type of extramural exposure via informal activities (listening to music, reading, etc.). Results show that the high-CLIL group outperformed both the low- and non-CLIL groups on the reading test, whereas no differences were found between the low- and the non-CLIL groups' reading scores. Moreover, where listening is concerned, no differences were found between the groups. Regarding extramural exposure, a weak but significant correlation was found between reading in English at home and the reading test scores, and listening to music and the listening scores. Some differences were also observed concerning gender, as watching tv and using the internet had an impact on reading for boys, while the correlation between listening to music and listening comprehension was only observed in girls.

---

\*Speaker

**Keywords:** CLIL, EFL, young learners, reading, listening, extramural exposure

# Declarative and procedural memory effects in second language acquisition: An aptitude–treatment interaction study

Simón Ruiz <sup>\*†</sup> <sup>1,2</sup>, Wensi Zhang <sup>3</sup>, Chaofan Jiang <sup>3</sup>, Jiayi Li <sup>3</sup>, Siqi Yang <sup>3</sup>,  
Padraic Monaghan <sup>3</sup>, Patrick Rebuschat <sup>2,3</sup>

<sup>1</sup> Universidad de Cantabria [Santander] – Spain

<sup>2</sup> University of Tübingen – Germany

<sup>3</sup> Lancaster University – United Kingdom

Meta-analytic research on the effectiveness of second language (L2) instruction has consistently shown that instruction is beneficial for L2 learning, and that different types of instruction produce differential effects on the development of L2 proficiency (Kang et al., 2018). While there is substantial evidence for the role of individual difference (ID) factors in L2 learning (Li et al., 2022), significantly less is known about possible interactions between IDs and educational treatments (i.e., aptitude-treatment interactions (ATIs); Cronbach & Snow, 1977). Examining possible ATIs in L2 acquisition is essential because instructional effectiveness depends largely on learners’ characteristics (DeKeyser, 2021). Critically, it is urgent to understand ATIs because for the first time adaptive L2 treatments can be provided on scale via Intelligent computer-assisted language learning systems (Ruiz et al., 2024).

We investigated 111 native speakers of Chinese who were exposed to an artificial language, to be acquired through cross-situational statistics, under three different learning conditions. All learners were asked to determine which of two scenes the sentence referred, and could acquire this through learning cross-situational associations between words and elements of the scene. In the explicit rule-search condition, they were also instructed to find out the rules of the language; in the explicit-presentation condition, they were further provided with grammatical information about the language; and in the implicit condition, learners were given no additional instruction. The data analysis was preregistered and conducted with mixed-effects regression modeling.

We found that both explicit learning conditions yielded larger learning gains than the implicit learning condition, with the explicit rule-search learning condition having the largest impact. Regarding ATIs, due to logistical reasons, we were unable to collect ID data from the explicit rule-presentation group; we therefore here report results concerning only the implicit and explicit rule-search groups. We found that declarative memory (DM) was positively related to learning outcomes in the explicit rule-search group, particularly for learners with relatively high DM ability, which support the theoretical assumptions and previous findings suggesting that DM is involved in L2 acquisition in explicit learning contexts (see Morgan-Short et al., 2022, for a recent review). Overall, the findings of this study contribute to our understanding of the underlying cognitive mechanisms of language learning by examining how declarative and procedural memory, two widely-studied and important cognitive ID variables, affect language acquisition, as well as the role of IDs in L2 development under different learning contexts, and further confirm the relevance and feasibility of investigating interaction effects in L2 research (DeKeyser, 2012, 2021).

---

\*Speaker

†Corresponding author: sruiz@sfs.uni-tuebingen.de

**Keywords:** individual differences, declarative memory, procedural memory, implicit and explicit treatments, aptitude by treatment interactions, crosssituational learning

# Investigating the relationship between gaze-contingent eye-tracking and L2 vocabulary development: The effects of reading proficiency and vocabulary size

Andrea Révész \*<sup>1</sup>, Matthew Stainer<sup>2</sup>, Ana Pellicer-Sanchez<sup>1</sup>, Marije Michel<sup>3</sup>, Yoojin Chung<sup>1</sup>, Mary Elisabeth Martin<sup>1</sup>, Sanne Van Eijsden<sup>3</sup>

<sup>1</sup> University College, London – United Kingdom

<sup>2</sup> Griffith University – Australia

<sup>3</sup> University of Groningen – Netherlands

Eye-tracking methodology is becoming increasingly popular as a tool to capture cognitive processes. Through this technology, however, it is also possible to design systems that can interact with eye movements. This gaze-contingency affordance holds particular promise in fostering second language (L2) acquisition (Révész et al., 2023). It is generally assumed that, to optimize L2 development, attention needs to be drawn to linguistic features. Some researchers (e.g., Long & Robinson, 1998) have also argued that this can be effectively achieved when linguistic information is supplied in context in response to linguistic performance. This hypothesis, however, has been the object of little empirical research in the context of reading. Therefore, our project aims to explore the extent to which interactive, gaze-contingent eye-tracking can facilitate attention to and development in L2 vocabulary knowledge while reading. Additionally, we explore the extent to which reading proficiency and vocabulary size might influence links between eye-gaze contingent eye-tracking and vocabulary learning.

The project forms a multi-site experiment (Hungary, Netherlands), involving 180 English as a foreign language children. The participating children are secondary school students at CEFR A2-B1 levels. As part of the treatment, they complete two reading tasks under one of three conditions: target words always highlighted, never highlighted, or highlighted interactively through gaze-contingent eye-tracking (i.e., target words are only highlighted when children gaze at the words while reading). Each reading task is followed by comprehension questions. The target items include low-frequency words, 10 nouns and 10 adjectives, each occurring twice over the two texts. We measure participants' attention to the target words through eye fixation duration and count measures and assess their learning of the target vocabulary through form recognition, meaning recognition and meaning recall pre- and post-tests. Reading proficiency is established through the reading component of the Cambridge Preliminary test, and vocabulary size is assessed with the XLex test (Meara & Milton, 2003).

Our pilot results from mixed effects modelling indicate that the interactive, gaze contingent condition leads to most attention to and development in vocabulary knowledge. Those with greater reading proficiency and vocabulary size benefit more from gaze-contingent highlighting. Participants also do better on the form recognition than meaning recognition and recall tests. The implications of these results will be discussed with respect to attentional models and optimal pedagogical interventions for vocabulary learning through reading, considering learners' reading proficiency and vocabulary size. Based on our results, we will also consider the potential

---

\*Speaker

of eye-gaze contingency in technology-mediated second language learning.

## References

Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition*. Cambridge University Press.

Meara, P., & Milton, J. (2003). X-Lex, the Swansea Levels Test. Express.

Révész, A., Stainer, M., Jung, J., Lee, M., & Michel, M. (2023). Using eye-tracking as a tool to develop lexical knowledge. *Language Learning & Technology*, 27(1), 1–22. <https://hdl.handle.net/10125/73537>

**Keywords:** gaze, contingent eye, tracking, vocabulary learning, attention

# Research trajectories on the interplay of L2 teachers' language aptitude and their students' L2 learning development

Olivia Rütli-Joy \* <sup>1</sup>

<sup>1</sup> Faculté des lettres et des sciences humaines - Université de Fribourg – Switzerland

The extent and pace at which an individual acquires an L2 in an instructional setting depends on multiple individual difference variables and environmental factors. Language aptitude (LA) is considered a key component of L2 learning success (Stansfield & Winke, 2008), correlating with general L2 proficiency at around .50 and explaining roughly 25% of the variance (Berthele & Udry, 2021; Li, 2016, 2019). While traditionally conceptualised as a stable ability (Carroll, 1993; Skehan, 1998), recent research has shown that LA may develop with increasing experience (Eisenstein, 1980; Grigorenko et al., 2000; McLaughlin, 1990; Sáfár & Kormos, 2008). Teachers' dispositions are equally crucial for L2 learning (Hattie, 2023; Seidel, 2014). Good teachers can significantly impact their students' learning progress (Rivkin et al., 2005), and meta-analytic results show that teachers are an important source of variance in their learners' development (Hattie, 2009). In the L2 classroom, teachers' language competences (TLC, cf. Rütli-Joy, 2022) are an important predictor for their students' L2 attainment (Burke, 2015; Elder & Kim, 2014; Kuster et al., 2014; Rütli-Joy, 2022). Further critical elements include teachers' professional knowledge (Shulman, 1986, 1987), pedagogical beliefs, values and thinking (Öztürk & Yildirim, 2015; Richards, 2010), and teachers' prior language learning experience (Borg, 2006). From the perspectives of socio-constructivist learning theory, lifelong learning (cf. Billett, 2018; Poquet & de Laat, 2021) and continuous professional development (Billett, 2018; Kuster et al., 2014), L2 teachers remain L2 learners who continuously develop their L2 abilities, learning capabilities, and teaching competences. Based on the premises that 1) prior L2 learning experience can influence 1a) LA scores, 1b) TLC, and 1c) L2 teaching effectiveness, and that 2) TLC can influence 2a) L2 teaching effectiveness, the question arises in what ways these variables interconnect, and whether teachers' LA scores correlate with their learners' L2 development. This poster presentation outlines a correlational study that seeks to investigate the following research questions: RQ1: What correlations can be identified between L2 teachers' LA scores, TLC, and their learners' L2 development?

RQ2: To what extent is the LA construct measured on the sample malleable?

RQ3: To what extent does teaching quality mediate L2 teachers' and their learners' LA and L2 proficiency over the course of one academic year?

Adolescent upper secondary school students and their respective teachers (adult "learners") constitute the sample for this study. Teachers' and learners' LA scores and L2 proficiency and TLC, respectively, are to be measured at the beginning (t1) and at the end of the school year (t2). In addition, a student questionnaire on teaching quality (Fauth et al., 2019) and a teacher questionnaire on applied teaching materials and methods are to be administered at t2 to control for some of the potential moderating variables. Correlational analyses are to be conducted post

---

\*Speaker



data collection. This poster outlines the state of the current research on the present research desideratum and deduces the rationale and study design of the envisaged study.

**Keywords:** Language aptitude, foreign language education, language attainment, language proficiency, teacher competences, individual differences

# Bilingual cognition and the effects of L2 category: colour and spatial concepts

Miho Sasaki \*<sup>1</sup>, Rune Okada<sup>1</sup>, Riku Yamane<sup>1</sup>

<sup>1</sup> Keio University – Japan

Recent studies on bilingual cognition have demonstrated cross-linguistic influence in various domains such as colour, number, grammatical gender and motion events (Cook & Li, 2016). Furthermore, the factors modulating the effect of language on bilingual cognition have been investigated considerably and we now know that a number of learner factors such as L2 proficiency, language usage, length of stay in the L2 environment, age of L2 acquisition are involved in bilingual cognitive restructuring. Based on those findings, the current study investigates how Japanese users of English are affected by their L2 in their categorisation of the two domains, colour and spatial concepts and compares the results.

Colour is the domain that is known to be influenced by the L2 category (Athanasopoulos, 2009; Athanasopoulos et al., 2011). While the previous studies demonstrated some integration of the two L1 colour concepts (e.g. *ao* "blue" and *mizuiro* "light blue" in Japanese) in L2 (e.g. *blue* in English), the current experiment examines how a single L1 colour concept (*murasaki* "purple" in Japanese) can be categorised into the two concepts in L2 (*purple* and *violet* in English). The other cognitive domain we examine is the classification of spatial events. Previously, Korean-English bilinguals demonstrated the L1 contrast between tight-fit and loose-fit relation differently from the monolinguals (Park & Ziegler, 2014). Online similarity judgment tasks contrasting several Japanese and English expressions (e.g. *hameru* "put on" and *ireru* "put in") were designed to investigate whether Japanese-English bilinguals would be affected by the L2 category according to their L2 proficiency or other factors.

19 Japanese users of English participated in the Japanese and English colour tasks and the spatial event task in Japanese and they were divided into the two proficiency groups. The colour task was conducted using Colour-aid colour chips same as Athanasopoulos et al. (2011). The results show that the advanced bilinguals choose the focus of *murasaki* differently from the intermediate users while both groups choose various colour chips for the English colour terms, which indicate their L2 foci are not consistent. The similarity judgments show that the advanced group tends to distinguish the near colour cross-category pairs more differently than the intermediate group. In the spatial event similarity judgment task where they see photos or videos showing some spatial movements, the advanced group judge them based on the L2 category as expected while the response times are not affected. The post-experiment interviews reveal reasons for individual variations, which are often found in bilingual cognition studies.

Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition*, 12(1), 83-95.

Athanasopoulos, P., Damjanovic, L., Krajciová, A., & Sasaki, M. (2011). Representation of colour concepts in bilingual cognition: The case of Japanese blues. *Bilingualism: Language and Cognition*, 14(1), 9-17.

---

\*Speaker

- Cook, V. J. and Li, W. (Eds.) (2016). *The Cambridge Handbook of Linguistic Multicompetence*. Cambridge: CUP.
- Park, H. & Ziegler, N. (2014). 'Cognitive shift in the bilingual mind: Spatial concepts in Korean–English bilinguals'. *Bilingualism: Language and Cognition*, 17 (2), 410–430

**Keywords:** Bilingual cognition, Categorisation, Colour, Spatial concepts, Japanese, Multicompetence

# When the dative becomes less reliable – L1 attrition in a multilingual context

Judith Schlenter <sup>\*†</sup> <sup>1</sup>, Marit Westergaard <sup>\*</sup>

1

<sup>1</sup> The Arctic University of Norway [Tromsø, Norway] – Norway

In two experiments, we investigated the use of morphological case for argument interpretation in a group of German L1 speakers (n=34) in Norway. All expats were regularly exposed to their later-learned languages Norwegian (~43%) and English (~30%). Unlike German, both Norwegian and English lack case marking on full NPs, and word order is a reliable cue to argument interpretation. Like in German, (+definite) is realized on pre-nominal articles in English, while in Norwegian it is realized post-nominally as a suffix. Following models such as the *Competition Model* (1) or the *Attrition via Acquisition* model (2), cross-linguistic influence from additional languages on the L1 should occur especially in contact situations in which languages display similar, but also slightly different, properties.

Participants were presented with two scenes, one showing an event as described by a transitive or ditransitive sentence and another showing the same event, but the roles of agent and patient (transitive events) or recipient and theme (ditransitive events) were reversed. Transitive sentences either followed the order subject-verb-object or object-verb-subject (OVS). In ditransitive sentences, the order of objects was manipulated. The task was to select the target scene. Picture selection was combined with the recording of eye movements (Experiment 1) or mouse movements (Experiment 2). In both experiments, we found no strong evidence of an influence from Norwegian and/or English, i.e., no overreliance on word order. When presented with an OVS sentence, participants used the case cue to revise any initial expectation towards an NP1-as-agent interpretation. Experiment 1 showed that the expats used accusative case in transitive and ditransitive sentences to predictively fixate on the target scene. However, unlike a previously tested group of L1 speakers in Germany (3), they did not or to a limited extent use dative case as a predictive cue, indicating effects of attrition specific to dative morphology.

Our finding that dative morphology is prone to L1 attrition is in line with previous research showing a specific vulnerability of the dative in (child) L2 attrition (4) as well as heritage German (e.g., 5). Importantly, unlike in previous research, in our study this cannot be explained in terms of incomplete or differential acquisition, as our participants were adults when they moved away from a German-speaking environment.

(1) MacWhinney, B. (2019). Language attrition and the Competition Model. In M. S. Schmid & Köpke (Eds.), *The Oxford Handbook of Language Attrition* (pp. 6–17). Oxford University Press. (2) Hicks, G., & Domínguez, L. (2020). A model for L1 grammatical attrition. *Second Language Research*, 36(2), 143–165. (3) Authors (under review). What eye and mouse movements tell us about expectations towards argument order: An eye- and mouse-tracking study in German. (4) Flores, C. (2015). Losing a language in childhood: A longitudinal case study on language attrition. *Journal of Child Language*, 42(3), 562–590. (5) Yager, L., Hellmold, N., Joo,

---

\*Speaker

†Corresponding author: judith.schlenter@uit.no

H.-A., Putnam, M. T., Rossi, E., Stafford, C., & Salmons, J. (2015). New structural patterns in moribund grammar: Case marking in heritage German. *Frontiers in Psychology, 6*, 1716.

**Keywords:** L1 attrition, multilingualism, prediction

# The effects of word frequency and noise on listening effort in bilinguals' L1 and L2: A pupillometry study

Jens Schmidtke \* <sup>1</sup>, Dana Bsharat-Maalouf<sup>†</sup> <sup>2</sup>, Tamar Degani <sup>2</sup>, Hanin Karawani <sup>2</sup>

<sup>1</sup> Haifa Center for German and European Studies, University of Haifa – Israel

<sup>2</sup> Department of Communication Sciences and Disorders, Faculty of Social Welfare and Health Sciences, University of Haifa [Haifa] – Israel

While listening in one's first language (L1) is a seemingly effortless process, listening in a second language (L2) can be significantly more effortful, for example, in the presence of background noise (e.g., Bsharat-Maalouf & Karawani, *Cognition*: 2022). Another factor thought to influence listening effort is word frequency (Schmidtke, *Frontiers in Psychology*: 2014). Word frequency effects are well established in the literature with word recognition of low frequency words being slower, especially in an L2 (e.g., Duyck et al., *Psychonomic Bulletin & Review*: 2008). In this study, we explored the impacts of background noise and word frequency during listening in both L1 and L2, utilizing pupillometry as a tool to quantify listening effort with the purpose of gaining a better understanding for why listening in an L2 is difficult. An improvement over previous studies lies in our testing of the same individuals in their L1 (Hebrew) and L2 (English), thereby accounting for general individual differences in the pupil response. Participants were 47 normal-hearing university students. Stimuli comprised 80 words in each language, half of which were presented in quiet and the remainder in noise (SNR: 0 dB). Word frequency information was derived from corpora of subtitles, representative of spoken language. Each trial consisted of a one second period to establish baseline pupil dilation, the presentation of the word (mean duration = 830ms), and a three second pause to measure post-processing pupil dilation. Participants were then prompted to recall the heard word. Overall, accuracy was high in all conditions (> 92%), confirming participants' ability to perform the task. The pupil data was pre-processed and all analyses were performed on baseline-corrected values, following standard procedures. Data were binned in bins of 50ms to reduce the number of computations. Subsequently, mixed-effects models were run on each bin with main effects for language, condition, frequency, and their interactions. In addition, audio length was a control variable.

An effect for language emerged starting at 600ms after word onset. Critically, a language by frequency interaction emerged between 1400 and 3100ms, with larger pupil dilation in L2. Subsequent analyses for each language revealed an absence of the effect in the L1 data, while in the L2 data, the frequency effect was significant from 1350ms to 2850ms: lower frequency was associated with larger pupil values. The effect of noise was such that the difference between high and low frequency words was diminished.

The absence of a frequency effect in L1 aligns with existing literature (e.g., Mor and Prior, *International Journal of Bilingualism*: 2021). Pupillometry adds information regarding the timing of the effect in L2. While the language effect emerged early, suggesting more efficient lexical access for L1 words, the frequency effect emerged relatively late (ca. 500ms after word offset).

---

\*Speaker

<sup>†</sup>Corresponding author: danabsharat5@gmail.com

Implications for current models of speech understanding will be discussed.

**Keywords:** pupillometry, listening in noise, frequency effects

# Minority language abilities and use boost L3-English abilities at the early stages of L3 acquisition

Marta Segura <sup>\*† 1</sup>, Adriana Soto-Corominas <sup>1</sup>

<sup>1</sup> Universitat Internacional de Catalunya [Barcelona] – Spain

Researchers have attempted to extend the theory of Linguistic Interdependence (Cummins, 1979) to third language (L3) acquisition (e.g., Bertele & Udry, 2022; Cenoz, 2003). However, studies that investigate large samples of L3 acquirers with homogeneous L1/L2 backgrounds remain scarce, especially at the early stages of L3 acquisition, where transfer of skills and knowledge from L1/L2 may be more directly observable. This longitudinal study investigates L3-English development in a bilingual community, Catalonia (Spain), where Catalan, the minority language, and Spanish, the majority language, are official languages. Participants (N=185) were tested three times: at the beginning of Grade 1 (Time 1 of testing; T1), at the end of Grade 1 (T2) and at the end of Grade 2 (T3). At the three time points, participants completed a receptive lexical and grammar task in Catalan, Spanish, and English. Parents completed a background questionnaire yielding participant background and linguistic information. We asked two main questions: 1-Over time, once L3 exposure and L3 age of onset of acquisition (AOA) are controlled for, is there any additional variance in participants' L3 abilities explained by L1/L2 skills? 2-Do abilities in one language predict L3 abilities more closely? If Linguistic Interdependence applied, participants with stronger L1 and L2 skills should show increased L3 performance, all else (L3 exposure/AOA) being equal.

To answer these questions, we employed two negative binomial models where English lexical scores (Model 1) and grammar scores (Model 2) were predicted by Testing time (T1, T2, T3), by the respective Catalan and Spanish scores and other covariates capturing English exposure, English AOA, and family socioeconomic status.

Whereas the model for vocabulary only found the Catalan score to be significantly and positively associated with the English score, the grammar model found a positive association between the English score and both the Catalan and Spanish scores, even though the association was stronger with Catalan. To probe into this apparent minority language advantage, a post-hoc model which additionally controlled for Catalan and Spanish language use in the home found that more Catalan use with family was associated with higher English scores overall.

In conclusion, results point to clear bilingualism effects in the L3-English development of receptive skills but not towards a strong theory of interdependence of L1/L2 skills. Evidence of an advantage was found for those with higher skills in or most frequent use of the minority language. These advantages are unlikely to stem from any direct transfer between languages, as Catalan is not closer to English than Spanish in its grammar or lexicon. Alternatively, we propose that because Spanish is the majority language in Catalonia, more minority language use and stronger minority language abilities foster additive and balanced bilingualism, yielding more balanced patterns of language use altogether. More balanced bilinguals may attain higher bilin-

---

\*Speaker

†Corresponding author: mseguram@uic.cat



gual abilities overall, which may in turn benefit L3 learning. Indeed, unplanned post-hoc models indeed confirm that bilinguals with more minority language exposure had higher performance in the three languages.

**Keywords:** L3, receptive skills, bilingualism, minority language, primary

# How much L2 grammar practice is needed to gain long-term knowledge?

Jonathan Serfaty \* <sup>1</sup>

<sup>1</sup> Université Côte d'Azur – UMR BCL – France

Second language (L2) grammar is a skill that can only be mastered through practice. Much research has already been conducted to determine how many times vocabulary should be retrieved before it can be remembered (e.g., Rawson & Dunlosky, 2022), but no previous research has determined the quantity of practice necessary to obtain durable and accurate knowledge of L2 grammar. Skill acquisition theory predicts that continuous productive application of the grammar rule in L2 sentences would allow learners to proceduralize a specific grammar feature, resulting in durable knowledge (DeKeyser, 2020). This process is predicted to be complete when learners no longer commit errors during practice (Kim et al., 2013). The aim of the following research was to quantify the amount of practice necessary to complete this proceduralization process.

I conducted these experiments using digital flashcards for grammar, a method that had previously been used to investigate the optimal distribution of practice sessions (Serfaty & Serrano, 2022). Each flashcard prompts the learner to type an L2 sentence. Collectively the sentences allow learners to induce and practice a particular grammatical rule, and a session ends when all the flashcards have been typed perfectly.

The first experiment tested whether the number of errors committed during a session indeed predicted durable memory, and explored how many sessions are necessary to achieve this level of knowledge. Participants (N = 119) learned an artificial language over two, three, four, or five daily sessions and were tested after 14 days. Posttest scores were similar after two or three sessions, then substantially increased after four sessions. No advantage was gained from the fifth session. It was also at the fourth session that participants ceased to improve their error rate, seeming to confirm that the error rate predicts proceduralization. The concept of a minimum-trials session (MTS) emerged, defined as a session in which a learner committed no more errors than the mean minimum number of errors in any session. In this case, it was two errors on 12 items. When reanalyzed by the number of MTSs, rather than the total number of sessions, I found that memory was durable after a learner performed two MTSs, regardless of how many sessions they performed in total.

The second experiment tested the hypothesis that two MTSs predicted durable knowledge by repeating the original experimental design with English language learners (N = 76) in a high school, using two different grammatical structures, and extending the posttest delay to two months. This experiment reproduced the findings from the first experiment.

These results expand our understanding of L2 skill acquisition by providing a possible behavioural marker for the completion of the proceduralization process, at least for decontextualized explicit rule application. The findings also provide tentative guidelines for L2 learners to measure their grammar knowledge.

---

\*Speaker

**Keywords:** grammar, flashcards, practice, memory

# Investigating patterns of teacher-student interaction in secondary education classes with recent migrant students

Shauny Seynhaeve \* <sup>1</sup>

<sup>1</sup> Universiteit Gent = Ghent University – Belgium

Classroom interaction has been put forward as an important aspect of learning, both in the field of applied linguistics as well as general educational studies. On the one hand, classroom interaction is claimed to facilitate language learning in the second language classroom (Tsui, 1985). On the other hand, learner-teacher interaction has been pointed at as an important condition for subject learning in general (Lo & Macaro, 2012). This makes classroom interaction a crucial aspect of schooling in various contexts, with particular relevance, perhaps, for Newly Arrived Migrant Students (NAMS).

Educational research has shown that NAMS have a reduced probability of attaining educational success: this growing student population shows higher dropout and grade retention rates and more enrollment in programs with lower social status (Lüdemann & Schwerdt, 2013). Various studies point to language - and more specifically: a lack of proficiency in the language of instruction - as an explanation for these inequalities (OECD, 2018; Vanblaere, 2012). While classroom interaction has the potential to serve as a supportive tool for the learning process of these vulnerable students, there is a lack of understanding regarding the specific interactional practices that NAMS encounter.

The objective of this research is to identify the interactional patterns between teachers and students in secondary education classes with NAMS. The study is set in Flanders, where a rising number of migrant students is enrolled in separately organised reception education classes for NAMS designed to help the students with the learning of Dutch and reinforce their social integration process. After approximately 1 year, most high school-aged NAMS transition into mainstream secondary education, where they continue their language learning processes while concurrently focusing on general schooling. Given that NAMS are integrated with mainstream students in regular secondary education classes, this context is particularly pertinent for examining interactional practices.

Three regular secondary education schools in Flanders were selected through convenience sampling. Across these schools, nine teachers were observed over 28 hours of teaching classes including NAMS. The teacher-whole class interaction in these lessons was transcribed verbatim using EasyTranscript, after which qualitative analysis was performed in order to identify teacher-student interactions. Then, four parameters based on An et al. (2021) were used to further analyze these interactions: (1) time percentage of teacher talk and student talk, (2) turn-taking behaviour between the teacher and students, (3) frequency of IRF sequences, and (4) time proportion, length and frequency of teacher monologue. Preliminary results show that whole-class interaction was highly teacher-dominated. Moreover, there were little opportunities for NAMS to take part in these interactions.

By focusing on NAMS, this study addresses a gap in the literature on classroom discourse in

---

\*Speaker

a specific but relevant context. Our study hopes to find good-practice examples of classroom interaction for NAMS, while also pinpointing areas where interaction patterns with NAMS allow for improvement. That way, our study will have implications for all teachers who deal with pupils from a different language background in their classroom.

**Keywords:** interaction, migrant students, classroom, secondary education

# L2 proficiency and orthographic effects on phonology across writing systems

Louise Shepperd \* <sup>1</sup>

<sup>1</sup> Radboud university [Nijmegen] – Netherlands

Adult language learners are typically exposed to large quantities of written input from the earliest stages of learning, with varied and far-reaching effects on L2 phonology (Bassetti, 2023; Hayes-Harb & Barrios, 2021). Most research investigating this topic focuses on learning across languages that share the same orthographic script. Additionally, many studies target first exposure and naïve participants, meaning that the role of proficiency has received limited attention in this field (Veivo et al., 2018). The present study contributes to emerging research into the relationship between proficiency and cross-scriptal orthographic influence on L2 phonological development. Specifically, it investigated the influence of written input on adult Arabic-speakers' accuracy in lexically-encoding a difficult English phonological contrast. Participant performance, perceptions, and strategies were then explored in relation to L2 proficiency.

A hierarchical cluster analysis was conducted, using four measures of L2 English proficiency for 109 L1 Arabic-speaking learners, indicating that participants were best understood in terms of a higher and lower proficiency cluster. These proficiency clusters were then analysed in relation to performance on an online word study, as well as reported perceptions and strategies in a posttest questionnaire. Word learning consisted of 12 English pseudowords, which were presented auditorily, accompanied by an image and written input. Target words were minimal pairs beginning with /f-v/, which is a difficult contrast for L1 Arabic-speakers, and /m-n/, which is a well-established contrast in both Arabic and English. Target words were presented with either Arabic spelling, English spelling or no written input. Word learning was tested using an audio-picture matching task, where half the trials matched and the other half mismatched target words with the images of their minimal pair item. Participants then reported their perceptions related to orthographic influence and language learning strategies used during the study.

Statistical analysis revealed that the two clusters differed significantly in their performance and strategy usage. The higher proficiency group were more accurate and reported less frequent usage of Arabic orthographic knowledge during word learning, compared to the lower proficiency group. Across both clusters, there was a consistent preference for exposure to English spelling and a belief that written input was important during early word learning. Offering complementary insight, four learner profiles from each cluster were formed, further examining learner backgrounds and individual differences. These profiles highlighted the diversity in learners' experiences, attention, and intentions, as well as demonstrating that orthographically-induced interference persists into high levels of proficiency. Taken together, these findings invite future research to look more closely at individual differences, such as cognitive control, auditory processing ability, working memory, and (language) learning experience, in relation to proficiency and cross-scriptal orthographic influence.

Bassetti, B. (2023). *Effects of Orthography on Second Language Phonology: Learning, Awareness, Perception and Production* (1st ed.). Routledge.

---

\*Speaker

- Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, 1–30.
- Veivo, O., Porretta, V., Hyönä, J., & Järvikivi, J. (2018). Spoken second language words activate native language orthographic information in late second language learners. *Applied Psycholinguistics*, 39(5), 1011–1032.

**Keywords:** L2 phonology, orthographic influence, proficiency, individual differences

# Progression and Task-based Variability of Linguistic Complexity across Grade Levels: A Case Study of the Secondary School Reading Texts in Hong Kong

Leung Shue Sum \* <sup>1</sup>

<sup>1</sup> University of Cambridge – United Kingdom

**Introduction.** Textbooks play a critical role in second language learning, particularly in preparing students for high-stakes exams. It is thus important to ensure that reading materials in textbooks progress and eventually match the examinations. One dimension to gauge such progression is through linguistic complexity, which has been shown to correlate with growing proficiency. The recent advancement of natural language processing tools has enabled automatic analysis of linguistic complexity and thus studies evaluating English-as-a-Second-Language (ESL) textbook passages.

**Prior studies.** Past studies on ESL textbook passages can be divided into two groups based on their objectives. Some studies (e.g. Gedik & Kosal, 2022) focused on whether textbook materials matched examination texts in terms of linguistic complexity, whereas others inspected the progression of linguistic complexity of textbook materials alone. Interestingly, all these studies disregarded the role of progression when comparing textbooks and exams, neglecting the possibility that textbook materials can increasingly resemble exam papers as the series progresses. More importantly, these studies overlooked the effects of tasks (i.e. the communicative purposes of texts), despite research evidence that task types (e.g. narration vs. description) can influence different aspects of linguistic complexity (e.g. Alexoupoulou et al., 2017; Michel et al., 2019).

**Aims.** This study inspects the progression of textbook materials used to prepare candidates for exams while accounting for task effects. There are three research questions: (i) whether linguistic complexity in textbooks progresses across grades, (ii) how the linguistic complexity in textbooks compares with that in exams, and (iii) whether task types influence linguistic complexity with progression controlled for.

**Data.** The study focuses on secondary school textbooks in Hong Kong, as they are specifically designed to prepare students for college entrance exams from an early stage (starting in grade 7). A small-scale corpus was compiled from over 400 ESL reading passages in secondary school English textbook series (grades 7 to 12) and the English Language college entrance examinations taken by twelfth graders. 46 linguistic variables from two categories (syntax: 14, lexicon: 32) were extracted with Kyle's (2023) Suite of Automatic Linguistics Analysis Tools.

**Findings.** Rank-based ANOVA on the data revealed the following: (i) all textbook series progressed but only partially, e.g. lexical sophistication increased from grade 7 to 12 but syntactic complexity plateaued after grade 9; (ii) linguistic complexity of textbook passages for grades 10 to 12 eventually matched (and at times exceeded) that of the exam texts; (iii) descriptive texts were typically more linguistically complex/sophisticated than narrative texts, e.g.

---

\*Speaker



more subordination and academic vocabulary; however, each task type exhibited unique characteristics, e.g. narrative texts featured more long-distance coordination (i.e. more T-units per sentence) and low-frequency collocations attested in fictional texts.

**Implications.** Pedagogically, the findings highlight the need for students to read across task types to familiarise themselves with task-type-specific features to develop as all-rounded readers. Methodologically, they also suggest that textbook studies should consider progression when comparing textbook content and exam materials, and account for task effects when pooling textbook passages together for analyses.

**Keywords:** textbook analysis, text complexity, progression, task effects

# Using structural priming to boost the production of a less preferred L2 syntactic structure

Edwige Sijyeniyo <sup>1</sup>, Sarah Schroyens \* <sup>1</sup>, Rob Hartsuiker <sup>2</sup>, Sarah Bernolet \*

1

<sup>1</sup> University of Antwerp – Belgium

<sup>2</sup> Ghent University – Belgium

This study investigated the effect of input frequency on syntactic priming effects for Dutch passives in late learners of Dutch (L2) with French as their native language (L1). The study builds on the results of Sijyeniyo et al. (subm.), who investigated abstract structural priming within Dutch passive alternants (PP-final passives: ”*De man wordt gevolgd door de hond*” – ”The man is being followed by the dog” and PP-medial passives: ”*De man wordt door de hond gevolgd*” – \*The man is by the dog being followed”) in a comparable group of L2 learners. They found that explicit language instruction affected the production preferences of L2 learners such that they showed a baseline preference for the PP-medial passive, which is dissimilar between the L1 and the L2 and doesn’t exist in the L1, rather than for the PP-final passive (”*L’homme est suivi par le chien*”). This general preference differed from the preference that was measured for native speakers of Dutch in the same study, who produced no PP-medial passives in the pre-experimental baseline and had a clear preference for PP-final passives in the baseline condition in the syntactic priming experiment. Because of these baseline differences between both groups, the priming effects that were obtained differed as well: Whereas the native speakers only showed priming for PP-medial passives, the learners of Dutch showed comparable priming effects for PP-medial and PP-final passives. However, we found significant interactions between L2 proficiency and production preferences and priming effects in the learners: Lower L2 proficiency speakers mostly produced PP-medial passives and showed strong PP-final priming, whereas higher proficiency L2 speakers produced PP-medial and PP-final passives, creating room for PP-medial priming. These results suggest a) that the observed priming effects are closely linked to the participants’ baseline preferences and b) that production patterns align more with these of the native Dutch speakers with increasing proficiency. The current study is a replication of Sijyeniyo et al.’s L2 priming study, but, here, we increased the proportion of PP-final passive primes to 50% (vs. 33.33% PP-medial passive primes and 16.67% baseline items) to boost the usage of this least preferred alternative. We found that learners produced more PP-final passives in the baseline condition than in Sijyeniyo et al.’s study (55 vs 40%). Moreover, we observed stronger PP-medial priming than PP-final priming. Again, we found significant interactions between L2 proficiency and production patterns: PP-final priming was weaker for advanced bilinguals than for less advanced bilinguals, because the proportion of PP-final passives in the baseline condition increased together with participants’ proficiency. We suggest that our manipulation may be useful to L2 teaching practices, especially when specific structures are underused.

---

\*Speaker

**Keywords:** syntactic priming, input frequency, syntactic preferences

# Writing as a Bridge to Speaking: Effects of a Task-based Writing Intervention on Speaking Skills in German of Pre-university Students in the Netherlands

Fred Sikkens \* <sup>1</sup>

<sup>1</sup> University of Groningen [Groningen] – Netherlands

In a second language (L2) writing offers several advantages compared to speaking. Given its slower pace and permanency, L2 writers have greater control over their output (Michel et al., 2022) allowing them to focus on form as well as meaning (Cumming, 1989). "More learner control over attentional resources (lexis and syntax) as well as more need and opportunity to language during and after production" (Williams, 2012), enables L2 learners to plan, monitor and review their production. These two characteristics of writing (slower pace and planning) have been found to play a significant role in the development of L2 speaking competence (Kang, 2020). Consequently, it might be that frequently engaging in writing tasks, facilitates subsequent oral production. That is, it might be that L2 learners can experiment and insert "new" and "more complex" forms, that they have learned and practiced during writing, in follow-up oral performances (Rubin & Kang, 2008).

To investigate the relationship between writing and speaking in an L2, we conducted a pilot study with 29 pre-academic fourth-grade students of German L2 in the Netherlands. As part of their regular foreign language class in school, participants were randomly allocated to a writing vs. speaking group. Drawing on task-based principles (Ellis, 2003), they either engaged in a writing or speaking task of one single task-cycle (pre/task/post). All oral performances were transcribed and then, like the written performances, coded for CAF measures (cf. Michel, 2017): linguistic complexity (TTR, Guiraud Index), accuracy (Weighted Clause Ratio) (WCR) (Foster and Wigglesworth, 2016), fluency (speech rate, text length, pauses, repairs). In addition, we measured functional adequacy using the 6-point rating scale of Kuiken and Vedder (2017). Data analyses demonstrate a varied and changing picture of the speaking and writing performances, with some indications that participants of the writing group scored higher on accuracy (WCR) and fluency compared to those of the speaking group. This allows for the tentative conclusion, that indeed, writing tasks may positively affect speaking skills. This support, however, is limited to the pilot study, which only looked at a single task cycle comprehending an interval of only three weeks, and more, and particular, longer interventions are needed.

The current paper addresses this need as it will present data of a six-month longitudinal intervention with a larger data set. The data collection of this study commenced in September 2023 and is planned to conclude in March 2024. At the outset, it involved N=111 participants from four different secondary schools in the Netherlands. All students participated in pre- and post-tests on writing and speaking to evaluate the outcomes of the intervention. The intervention excludes a control group (n = 24), while the experimental group (n = 87) was divided into a writing or speaking cohort. All experimental cohorts completed six task cycles, including writing vs. speaking tasks in German. Written and oral performances will be coded and analysed for

---

\*Speaker

relevant CAF measures. During the doctoral workshop, I hope to receive feedback on the type of measures and suggestions for further analyses.

**Keywords:** Task based research, L2 writing, L2 speaking, L2 German, linguistic complexity, accuracy, fluency (CAF), functional adequacy (FA)

# Learning the Pronunciation of English Words from Textual Input: Should We Listen First?

Liwen Situ \* <sup>1</sup>

<sup>1</sup> University of Western Ontario – Canada

There are no straightforward rules guiding learners in making reliable predictions of the pronunciation of English words. For example, the words *through*, *thought*, *though*, *tough*, and *plough* share the same *-ough* spelling, but each is pronounced differently. Also unpredictable is stress placement in various English words. In *precursor* and *pretentious*, the primary stress is on the second syllable, whereas in *predator* and *prevalent*, the stress is on the first syllable. Owing to such inconsistencies in spelling-sound correspondences, silent reading may leave representations in readers' minds of the pronunciation of words that will later need to be rectified as they hear the same words in speech. The question addressed in this study is whether such initial inaccurate representations are readily replaced this way.

This question is of practical relevance because, for many learners of English as a foreign language, reading remains the principal mode of exposure to the language. If we were to find that inaccurate representations of the pronunciation of words are *not* easily replaced by the correct pronunciation, then it may be more judicious to first engage the learners with spoken discourse featuring the same words. One could also argue, however, that silent reading can pique learners' curiosity about the pronunciation of certain words, and that this will make them pay special attention to these words when they encounter them in their spoken form. In this study (in progress), we therefore compared two sequences of engagement with the same textual material: silent reading followed by listening and listening followed by silent reading.

The study adopted a counterbalanced within-participant design with high-intermediate ESL learners ( $N \geq 30$ ) at a North American university. In a one-on-one meeting with the researcher, half of the participants read the first half of a text silently, then read it aloud, and then listened to the audio-recording. For the second half of the text, they first listened to the audio-recording, then read the passage silently, and then read it aloud. The other half of the participants performed the same tasks but in the reversed order. The text included 16 preselected target words that, according to piloting, are often mispronounced. As a post-test, the participants read the entire text aloud once more, allowing a comparison with how they pronounced the target words when they read the passages aloud previously. Because it is likely that individual differences, including learning styles, matter, the participants were asked to answer relevant items of Cohen et al.'s (2009) *Learning Style Survey* to help us to distinguish auditory from non-auditory learners. The students were also interviewed after the reading tasks about their experience of the tasks, and about whether they imagined what words sound like during silent reading.

The study contributes to a strand of L2 vocabulary research that compares the benefits of trial-and-error versus study-plus-retrieval procedures (e.g., Elgort, 2017; Strong & Boers, 2019) but it is the first (to our knowledge) to focus on pronunciation as a facet of vocabulary knowledge.

---

\*Speaker

**Keywords:** pronunciation, feedback, memory, learning styles

# Assessing the Impact of Typology and Proficiency on Pronoun Usage in SLA: A Large-Scale Analysis of Null-Subject Language Learners

Jye Smallwood \* <sup>1</sup>

<sup>1</sup> University of Cambridge – United Kingdom

This project is a large-scale investigation on the role of typology in language transfer by learners of null-subject languages. The project aims to explore transfer effects by language type. A null-subject language is one where subjects are either optional or unnecessary. In these languages, (e.g., Spanish, Italian, Arabic, Mandarin Chinese, Russian, Swahili, and many others) dropping a subject is permitted. However, English is a non-null-subject language, and a subject is required in virtually all contexts. The aim of the present study is to determine if this difference in subject parameters means that learners are more prone to omit pronouns in second language (L2) English production. The empirical evidence suggests that L1 influences the acquisition of certain features in L2. The data also suggests that language development is non-linear, and learners may struggle with transfer at the A2/B1 threshold in Common European Reference Framework (CEFR). The majority of the research on these questions, however, has focused on small-scale studies of first language (L1) Spanish or Italian learners of English. I will use data from the EFCAMDAT learner corpus to sample a large and typologically diverse set of learners that represents a broader range of the world's languages. This corpus contains over a million writing samples from 174,000 learners. This allows me to sample writing samples from learners of 11 typologically diverse L1s. Grouping these L1s in accordance with Roberts' (2019), this project thus explores whether the 4 established language types correspond to 4 distinct learner profiles. I will evaluate this by observing the suppliance of different types of subjects, objects, potential of postverbal subjects production in the learners' writing. Thus, the primary contribution of this project is to consider a larger range of language types in probing language transfer in second language acquisition (SLA).

**Keywords:** null subjects, typology, corpus, SLA, syntax

---

\*Speaker



# Exploring the role of task modality and task complexity in L2 performance in EFL classes

Tugce Smith <sup>\*† 1</sup>, Sebnem Yalcin <sup>\* ‡ 1</sup>

<sup>1</sup> Boğaziçi üniversitesi = Boğaziçi University [Istanbul] – Turkey

The current study investigates how L2 performance is affected by the interplay of task complexity (simple versus complex) and modality (speaking versus writing). Addressing a gap in task complexity research, it responds to the underexplored theorizing effects of writing on task complexity in L2 learning and performance. Previous studies primarily focused on the effects of writing and speaking on L2 performances (Bulté & Housen, 2009; Ellis, 1987; Ellis & Yuan, 2005; Ferrari & Nuzzo, 2009; Kormos, 2014; Yu, 2009; Vasylets et al., 2019) and the differing impacts of task complexity in these modes (Kormos & Trebits, 2012; Kuiken & Vedder, 2011; Tavakoli, 2014; Vasylets et al., 2017; Zalbidea, 2017). However, the combined effects of task complexity and modality have been largely overlooked, with Vasylets et al. (2017) being a notable exception.

The study's unique approach included investigating the interaction between task complexity and modality, using the same participants for all conditions and counterbalancing for task complexity and modality, a methodology not previously employed (Kuiken & Vedder, 2011; Tavakoli, 2014). It involved 52 seventh-grade students, a demographic rarely studied in task complexity and modality research, thereby providing fresh insights into this age group's language learning processes. The participants were randomly assigned to two modality orders: writing-speaking (N = 23) and speaking-writing (N = 29), and the effects of task modality and complexity on their syntactic and lexical complexity, as well as accuracy, were examined through a 2x2 factorial within-subjects design.

Interaction analyses, as unique part of this study, revealed a clear parallel increase between task complexity and task modality for phrasal and lexical complexity. This led to the conclusion that task complexity and modality affect phrasal and lexical complexity in the same manner. Unit accuracy, however, showed a statistically significant interaction stemming from simple tasks. Clausal accuracy results revealed a high likelihood of an interaction with a similar trend although the current sample did not detect one. These results raised questions about how learners allocate their attentional resources freed by the availability of more time in writing. It was concluded that tasks with lower conceptualization and organizational demands might free up more time and resources for the linguistic encoding and monitoring of language output in writing whereas tasks with higher cognitive demands might limit opportunities left for monitoring as learners might choose to be engaged in conceptualization and content. This amplifies the effects of learners' preferences in task complexity and modality.

Overall, the study underscores that task modality substantially influences L2 performance, with task complexity being a crucial variable. The findings suggest that the effects of task complexity differ between writing and speaking. Pedagogically, this implies that tasks with diverse cognitive demands across different modes can enhance various aspects of L2 performance. The study emphasizes the need for further research into the combined effects of task complexity

---

\*Speaker

†Corresponding author: betul.aktas1@boun.edu.tr

‡Corresponding author: sebnem.yalcin@boun.edu.tr

and modality, to guide language teaching and curriculum development towards a more balanced language learning approach.

**Keywords:** speaking, writing, task complexity, task modality, L2 performance

# The effect of morphological mismatch on the processing of L2 compounds: Evidence from eye-tracking

Suhad Sonbul <sup>\*† 1</sup>, Dina El-Dakhs <sup>‡ 2</sup>, Anna Siyanova-Chanturia <sup>§ 3,4</sup>

<sup>1</sup> Umm Al-Qura University – Saudi Arabia

<sup>2</sup> Prince Sultan University [Riyad] – Saudi Arabia

<sup>3</sup> Victoria University of Wellington – New Zealand

<sup>4</sup> Ocean University of China – China

Several studies have assessed the influence of first language (L1) – second language (L2) lexical mismatching on processing. While most of the available evidence was related to lexicalization (i.e., word-concept mappings), little effort has been made to examine the same phenomenon in relation to morphological mismatch when processing compounds. To illustrate, the Arabic equivalent for ‘railway’ is a compound (match) while the Arabic equivalent of ‘bookstore’ is a monomorphemic word (mismatch). Two models have been proposed to account for the effect of mismatching on L2 lexical processing (Potter et al., 1984): *word association model* and *concept mediation model*. In the context of morphological processing, the first model assumes that ‘match’ compounds are processed faster than ‘mismatch’ compounds, while the second model assumes no differences between the two types of compounds. Only a few studies have attempted to evaluate these assumptions using decontextualized measures either off-line (e.g., El-Dakhs et al., 2020) or on-line (e.g., Cheng et al., 2011). Through employing the eye-tracking technology, we aim to examine the processing of ‘match’/‘mismatch’ processing as it proceeds naturally during natural reading.

To this end, we tracked the eye movements of L1 speakers and L2 speakers (L1 Arabic) as they read three types of words in English: match (compounds in both languages, e.g., ‘railway’), mismatch (compounds in English only, e.g., ‘bookstore’), and control words (monomorphemic in both languages, e.g., ‘restaurant’). The target items included 18 match compounds, 18 mismatch compounds, and 18 control monomorphemic words embedded in sentence contexts (triplets):

Our apartment is close to the railway and the new car park. (match)

Our apartment is close to the bookstore and the new car park. (mismatch)

Our apartment is close to the restaurant and the new car park. (control)

Thirty-nine L1 speakers and 42 L2 speakers took part in the study. Both groups of participants took the V\_YesNo online vocabulary test (Meara & Miralpeix, 2017) as a rough measure of proficiency. Mixed-effects models were fitted in R to examine any differences in processing between L1 and L2 speakers and any potential effect of rough proficiency on processing. Separate models were constructed for four eye movement measures representing both early (first fixation

---

\*Speaker

†Corresponding author: sssonbul@uqu.edu.sa

‡Corresponding author: ddakhs@psu.edu.sa

§Corresponding author: anna.siyanova@vuw.ac.nz

duration and first-pass reading time) and late (total reading time and fixation count) processing. Length and frequency of lexical items were included as covariates in the mixed-effects models to partial out for their effect prior to examining the main factors.

Results showed that L1 speakers did not show significant differences between the three types of items, indicating baseline performance. L2 speakers, however, showed an early processing disadvantage (first-pass reading time) for 'mismatch' items relative to 'controls', with the effect being stronger for less proficient L2 readers. The effect, however, disappeared as proficiency increased. The 'match' items were processed similar to controls by the L2 group. These results can be interpreted in line with partial L1 interference. We discuss the findings in relation to theories of the bilingual mental lexicon and morphological processing.

**Keywords:** Compounds, proficiency, eye, tracking, L1 effect, morphological processing, bilingual mental lexicon

# On the Nature of the Morphological Mismatch in the Comprehension of Relative Clauses in Monolingual and L2 English Children

Adriana Soto-Corominas \* <sup>1</sup>, Evangelia Daskalaki \*

<sup>2</sup>, Johanne Paradis <sup>2</sup>

<sup>1</sup> Universitat Internacional de Catalunya [Barcelona] – Spain

<sup>2</sup> University of Alberta – Canada

Two main findings emerge from research on first language (L1) English comprehension/processing of relative clauses in children: First, subject relative (SR) clauses (1) are easier to understand than object relative (OR) clauses (2). Second, a mismatch in morphological features, such as number, between the two DPs (3-4) facilitates OR comprehension (Lau and Tanaka, 2021).

(1) Show me the rabbit (that \_\_\_ chases the monkey).

(2) Show me the rabbit (that the monkey carries \_\_\_).

(3) Show me the rabbit (that the monkeys carry- $\emptyset$  \_\_\_).

(4) Show me the rabbits (that the monkey carries \_\_\_).

Few studies have investigated whether the observed asymmetries apply in child L2-English acquisition (cf., Cilibriasi et al, 2022), and whether the mismatching advantage is independent from the morphological marking of the verb (3rd person singular *-s* vs. plural  $\emptyset$ ). We address these gaps by investigating children whose L1 is Syrian Arabic, a language with rich subject-verb agreement.

We asked: 1-Do L1-Syrian Arabic children acquiring L2-English behave like L1-English speakers in showing higher accuracy on SRs over ORs and in sentences with mismatching features over matching ones? 2-Does the type of feature mismatch (i.e., singular-plural with  $\emptyset$ -marking on the verb of the RC (3) vs. plural-singular with overt marking (*-s*) on the verb (4)) affect the comprehension of relative clauses?

Monolingual (N=38; *Mean*=10.69, *SD*=2.71) and bilingual children (N=79; *Mean*=11.73, *SD*=1.84) participated in this study. Bilingual children had been exposed to English for a mean of 4.47 years (*SD*=0.55) after immigrating to Canada, and they demonstrated knowledge of English 3rd person *-s* in a sentence repetition task. Participants completed a sentence-picture matching task where they had to match an auditory stimulus (1-4) to one of two pictures. For both SRs and ORs, 6 items included two DPs that matched in features (1-2), 6 items included a singular-plural mismatch (3), and 6 included a plural-singular mismatch (4).

---

\*Speaker

Regarding RQ1, a logistic regression showed a main effect of group, relative clause type, and matchingness, but no interactions, indicating that monolinguals were overall more accurate than bilinguals but both groups were more accurate with SRs than ORs, and with mismatching than matching relative clauses. Regarding RQ2, a logistic regression including only ORs (given that SRs were at ceiling) and relevant post-hoc contrasts found that bilingual participants were most accurate in sentences with a plural-singular mismatch, whereas monolinguals performed best on relative clauses with a singular-plural mismatch.

Overall, bilinguals show the same asymmetries (SR > OR, mismatching > matching) as monolinguals in the comprehension of relative clauses. Despite both groups benefiting from number mismatches, the nature of the advantage differed by group. In monolinguals, the advantage was triggered by the singular-plural condition (3), which may stem from pragmatic effects, as plural entities are more likely to be assigned subjecthood/agenthood. Bilinguals' advantage was granted by the plural-singular condition, where morphology is overt and clearly disambiguates the correct candidate for subjecthood. We speculate that this advantage may stem from bilinguals' L1, which is morphologically rich in subject-verb agreement.

**Keywords:** L2 English, Relative clauses, Feature mismatch

# Investigating the relation between second language proficiency and study success using a causal inference approach

Sybren Spit \* <sup>1</sup>, Sible Andringa <sup>2</sup>, Oisín Ryan <sup>3</sup>

<sup>1</sup> University of Amsterdam – Netherlands

<sup>2</sup> University of Amsterdam – Netherlands

<sup>3</sup> UMC Utrecht – Netherlands

There are many factors that determine whether students successfully complete a post-high school (or tertiary) educational program. The linguistic skills students possess is one such factor (Brown & Hudson, 2002; De Wachter & Heeren, 2013; Van Dyk, 2015). This is not unsurprising, given that language is the medium that is used to transmit knowledge, and that reading, writing, listening and speaking are important skills in any educational setting (Kieffer, 2008; Prevoo, Malda, Mesman & van IJzendoorn, 2016). There is some research that shows that language skills also play an important role in international students' study success. Students from abroad are generally required to demonstrate that they possess a sufficient level of proficiency in the language of instruction (their second language, or L2) for entering their favored programme (Daller & Phelan, 2013; Daller & Xue, 2009; Roche & Harrington, 2013), and there is evidence to suggest that educational success depends on a sufficient level of L2 proficiency (Hu & Trenkic, 2019; Trenkic & Warmington, 2019).

However, the number of studies investigating the relation between L2 proficiency and study success is limited. Most studies focus on university students, and most studies are observational in nature (i.e., non-experimental). They mostly do report that L2 proficiency and study success are correlated. As a result, it is unclear how this relation varies across different educational settings. A second problem is that there are questions about causality: it seems likely that L2 proficiency causally affects study success, but the evidence is mostly indirect. The present study used a novel graphical modeling approach to causal inference (Hernán & Robins, 2020; Pearl, 2009), which provides guidance when estimating causal effects from non-experimental studies. We aim to answer two questions with this study: is there a relation between language proficiency and study success, and does this relation differ across educational contexts? And can we use observational data to make inferences about a possible causal relation between language proficiency and study success?

We made use of a large dataset retrieved from the national Office of Education to answer these two questions. This dataset contains records from 2011 and onwards for all international students who took the official L2 State Exam, required to enter the educational system, as well as data on recorded study success as registered with the national Office of Education. The study success data pertain to a range of educational programs from vocational to academic programs. This dataset enables estimation of the causal effect that language proficiency would have on students' study success. In this poster presentation, we will present the outcomes of these analyses and discuss what these mean for developing L2 policy in educational settings.

---

\*Speaker

**Keywords:** Study Success, State Exam Dutch as a Second Language, Causal Inference



# Audiovisual processing of vowels in second language learners of Swedish and German

Helene Springer \*<sup>1</sup>, Frida Splendido<sup>1</sup>, Marianne Gullberg<sup>1</sup>

<sup>1</sup> Lund University – Sweden

Speech is a multimodal phenomenon that incorporates a variety of acoustic and visual cues that are part of our daily interactions. A wide range of studies has shown that visual information from articulatory gestures and lip movements facilitate the processing of speech sounds under both adverse and clear conditions, in both L1 and L2 listeners (Arnold & Hill, 2001; Drijvers & Özyürek, 2020; Hazan et al., 2010; Ross et al., 2007; Sumbly & Pollack, 1954; see Traunmüller & Öhrström, 2007 for Swedish). These studies suggest, however, that this is not a constant effect. My PhD project focuses on the modulation of the audiovisual benefit in second language listening. Research on how this effect is modulated by different linguistic and individual factors within and across different levels of processing is scarce, especially concerning a systematic treatment of variables such as visual salience, in combination with the distance of L1 and L2 sound inventories, and individual listener variables.

Thus far, my work has focused on operationalizing the notion of visual salience. Fine-grained statements on this variable would allow for vowel contrasts to be grouped into high- and low-salient on continuous scales (rather than binary phonological features). Such statements can be made by using a computer vision approach to extract visual parameters of vowels in Swedish from the addressee's perspective. The measured parameters will be used as reference material to define visual salience in the forthcoming perception experiments.

The experiments planned for my PhD project are centered around vowel perception in two specific groups, namely Swedish learners of German, and German learners of Swedish. These groups offer a particularly interesting scenario since German learners of Swedish encounter an additional vowel phoneme, whereas Swedish learners of German must adjust to a vowel system with one less vowel category compared to their L1. The experiments target three different levels of speech processing, namely discrimination, classification, and identification of vowels in three conditions: audio-only, visual-only, and audiovisual. More specifically, the experiments probe what auditory and visual features L2 listeners use for discriminating, classifying, and identifying vowels in the auditory and visual modality; whether audiovisual speech input facilitates their processing of L2 vowels; and if so, which linguistic and individual listener-related factors modulate this benefit. This will be addressed by examining the effect of acoustic and visual salience within the L2 vowel system and the L2 vowels' distance to the L1 vowels. Moreover, cognitive measurements of general working memory, phonological and visuospatial short-term memory will be included, and potential correlations of proficiency-related variables (language background, L2 listening proficiency, L1 lipreading, categorical perception, etc.) with response times and patterns, such as confusion matrices, etc.

Potential implications include educational contexts (teaching strategies, visual instructions, etc.), the implementation of visual aids and communication tools in multimodal and multilingual communication, the increase of accessibility for vulnerable listener populations (including hearing-impaired and L2 populations), the development of systems recognizing and synthesizing

---

\*Speaker

speech. Finally, it can provide insights into the interaction of the auditory and visual system in multimodal L2 perception.

**Keywords:** L2 sound acquisition, L2 vowel processing, audiovisual speech perception, visual salience

# Exploring TBLT practices in Hong Kong senior secondary schools through the framework of Intended Constructive Alignment

Cheung Suet-Sin \* <sup>1</sup>

<sup>1</sup> Institute of Education – United Kingdom

Task-based language teaching (TBLT) has gained increased popularity as a pedagogical framework for teaching second languages in the past few decades. The growing adoption of TBLT has naturally been coupled with an enhanced interest in task-based language assessment (TBLA). From a language testing perspective, imposing TBLA promotes positive washback as the assessment tasks and pedagogic tasks share the same theoretical underpinnings. This is also referred to as constructive alignment, wherein curriculum objectives and examination goals are intentionally aligned. While it is generally assumed that, when curriculum and examination goals are aligned, positive washback occurs, little is known about whether the introduction of TBLA indeed leads to greater reliance on TBLT teaching principles. To bridge this gap, the current doctoral project endeavours to investigate the extent of alignment between a task-based assessment and accompanying pedagogical approaches in the Hong Kong secondary context. Specifically, I will explore whether and how the integrated assessment task in the Hong Kong senior secondary exam may influence the adoption of TBLT through observing teachers' planning and use of textbook materials.

This project employs a qualitative research design obtaining three major sources of data: (1) exam-related documents, (2) textbook materials designed to prepare learners for the integrated task in the HK secondary exam, and (3) verbal protocol and interview data from local Hong Kong secondary English teachers. First, all published Hong Kong Diploma of Secondary Education English Language exam tasks, rubrics and reports will be evaluated against task definitions extracted from previous TBLT literature (East, 2021; Ellis, 2003; Ellis & Shintani, 2012; Long, 1985; Skehan, 1998) to gauge the extent of alignment between the

assessment tasks and task definitions. Second, with a view to understanding whether textbook writers use tasks and develop teaching sequences according to TBLT methodological principles (Long, 2015), I will analyse the learning activities and sequences in the textbooks through the lens of TBLT. Eventually, to explore teachers' TBLT practices, six teachers will be invited to video-record their lessons when delivering examination tasks feedback, adopting the textbooks and other learning activities. These lessons will be observed and triangulated with stimulated recall interviews. Online think-alouds requiring participating teachers to plan a lesson with a given unit selected from the textbooks will also be conducted. Lesson observation videos, stimulated recall interviews, and think-aloud sessions will all be coded and rated according to Long's (2015) TBLT methodological principles.

Results of this project will serve a dual purpose. Locally, it will inform officials about the ex-

---

\*Speaker

tent to which TBLT is implemented in secondary English language education holistically. This information will help evaluate the suitability of the current language curriculum and language teacher training programmes. More broadly, TBLT researchers and assessment experts will gain a deeper understanding of whether and how TBLT may materialise when TBLT is promoted via government mandate, high-stakes examination, and teaching and exam preparation materials. In this poster presentation, document analysis and teachers' think-aloud results will be shared.

**Keywords:** Language Testing, Integrated tasks, TBLT

# We Have Friend – Exploring Preference in Article Use in L1 Finnish Adult Learners of English and Swedish

Lari-Valtteri Suhonen <sup>\*† 1</sup>, Anders Agebjörn <sup>2</sup>

<sup>1</sup> University of Borås – Sweden

<sup>2</sup> Malmö University – Sweden

Article use poses a challenge for second-language learners (Chrabaszcz & Jiang, 2014). Previous research suggests that having acquired articles in an L2 facilitates the acquisition of articles in a third language (e.g., Aribas & Cele, 2021; Jaensch, 2009), and that article acquisition is to some extent determined by frequency-based patterns in the input (e.g., Ogawa, 2015). However, both suggestions suffer from methodological limitations. The present study addresses those limitations by investigating Finnish-speaking learners of English and Swedish, considering both cross-linguistic influence and frequency-based regularities in the input. Finnish students learn English and Swedish at school and Finnish learners (whose L1 lacks articles) might to some extent rely on their knowledge of English in Swedish article use (Jarvis, 2002; Nyqvist, 2016). For the present study, acceptability data (Likert 1–5) was collected from Finnish adults ( $n = 57$ ) as well as Swedish ( $N=133$ ) and English ( $N=30$ ) control groups. The stimuli consisted of 32 base sentences, each appearing in four versions – with and without an article in English and in Swedish – resulting in 128 unique variations of each sentence. No participant rated the same base sentences in multiple variations. We also collected self-ratings of proficiency. Blogmix (Språkbanken, 2017 (1.67M)) and enTenTen (Sketch Engine, 2022 (52G)), both building on colloquial language, were used to establish corpus frequency and article probability (i.e., the probability that the phrase occurs with an article).

For each item, the learners received a target-likeness score based on their deviance from the L1 groups' mean score. The target-likeness score was used as an outcome variable in two mixed-effects models, one with Swedish and one with English as the target language (TL). Both models included the participants' relative proficiency as well as item frequency, article probability score, and TL proficiency as fixed effects. Item and participant were included as random effects.

Both self-reported proficiency and participants' ratings indicate that the Finnish participants were generally more proficient in English than in Swedish. In our preliminary results, TL proficiency and article probability show as predictors of target-likeness in both mixed-effects models. However, there are clear differences between the models. Absolute frequency is only a significant predictor in the English model, meaning that the more frequent the item was, the more native-like the participants' ratings were. In line with the hypothesis about English affecting Swedish article use, we do see effects of comparative English-Swedish proficiency in the Swedish model.

## Selected references

---

\*Speaker

†Corresponding author: lari-valtteri.suhonen@hb.se

Aribaş, D. & Cele, F. (2021). Acquisition of articles in L2 and L3 English: the influence of L2 proficiency on positive transfer from L2 to L3. *Journal of multilingual and multicultural development* 42(1).

Chrabaszczyk, A. & Jiang, N. (2014). The role of the native language in the use of the English nongeneric definite article by L2 learners: A cross-linguistic comparison. *Second language Research* 30.

Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in SLA* 24(3).

Ogawa, M. (2015). Input frequency of article-noun combinations as another factor for L2 article use. *Ars linguistica: linguistic studies of Shizuoka* 21.

**Keywords:** L2, L3, articles, definiteness, Finnish, Swedish, English, cross, linguistic influence, transfer, frequency, corpus

# A lexical analysis of commercial-off-the-shelf games in English

Iwarin Suprapas \*<sup>1</sup>, Beatriz González-Fernández<sup>1</sup>

<sup>1</sup> The University of Sheffield – United Kingdom

Out-of-classroom (extramural) media has been recommended as a complement to compensate for the limited exposure to second language (L2) input, especially in English as a Foreign Language (EFL) contexts where classroom teaching time is constrained (Peters, 2018). In addition to movies, TV series, and podcasts, commercial-off-the-shelf (CoTS) games are among the most popular types of media that L2 learners engage with outside of the classroom, and have been speculated to have greater potential to promote L2 incidental vocabulary learning (Rodgers & Heidt, 2021). However, to date, the lexical features of gaming language has yet to be extensively and thoroughly. This research aims to build a lexical profile of games in the English language, and shed light on the vocabulary knowledge learners need to comprehend gaming language.

A corpus of gaming language was created, comprising language data extracted from 34 popular English CoTS games from a variety of gameplay genres (e.g., adventure, strategy, and role-playing games). Then the corpus was lexically analysed using Nation's (2017) 25000 BNC/COCA word lists as the reference list. Results indicate a lexical demand of at least 6000 most frequent word families with proper nouns and marginal words for 95% and more than 25000 for 98% coverage which is higher than other extramural input types (e.g., movies & TV series) (Webb & Rodgers, 2009a; 2009b). From the main corpus, various sub-corpora were created for each game genre. Each subcorpus was lexically profiled, and results were compared across sub-corpora to explore potential lexical differences across genres. Results reveal adventure games have the lowest and strategy games have the highest lexical demand. However, when analysing games within the same genre, the lexical demand varies; thus other factors, such as game content might be a better indicator of lexical demand than gameplay genre. Implications for vocabulary learning and recommendations for further research will be discussed.

## References

- Nation, I. S. P. (2017). The BNC/COCA Level 6 word family lists (version 2.0.0).
- Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL - International Journal of Applied Linguistics*, 169(1).
- Rodgers, M. P. H., & Heidt, J. (2021). Levelling up comprehensible input and vocabulary learning. In *Pop Culture in Language Education: Theory, Research, Practice*. Routledge.
- Webb, S., & Rodgers, M. P. H. (2009a). The lexical coverage of movies. *Applied Linguistics*, 30(3), 407–427. <https://doi.org/10.1093/applin/amp010>
- Webb, S., & Rodgers, M. P. H. (2009b). Vocabulary demands of television programs. *Lan-*

---

\*Speaker

**Keywords:** Input, extramural, lexical analysis, lexical profile, lexical demand, games, vocabulary



# How well do I select words? Exploring L2 vocabulary self-assessment

Yui Suzukida \*<sup>1,2</sup>, Takumi Uchihara<sup>3</sup>

<sup>1</sup> Juntendo University – Japan

<sup>2</sup> University College of London – United Kingdom

<sup>3</sup> Tohoku University – Japan

Second language (L2) learners' self-assessment has become the center of scholarly attention due to its impact on learner autonomy, self-regulatory learning, and learners' awareness of their own performance, as well as their confidence (Isbell, 2021). According to meta-analyses (Li & Zhang, 2020), self-evaluation has been explored across various linguistic domains, including listening, reading, writing (Yaghoubi-Notash, 2012), and speaking (Trofimovich et al., 2016). However, only a few studies have explored learners' self-assessment of vocabulary use (Gaffney, 2018). Given the crucial role that the accurate use of vocabulary plays in successful communication (e.g., speech comprehensibility; Saito et al., 2016), it is of utmost importance to understand how L2 learners perceive their vocabulary knowledge in contrast to their actual lexical proficiency. To address the current knowledge gap, this study set out to investigate (a) whether there are any discrepancies between objective assessments and self-assessments of vocabulary knowledge, and (b) factors affecting L2 learners' self-perception of vocabulary knowledge.

A total of 96 Japanese learners of English in Japan engaged in a semi-spontaneous speaking task, various cognitive and linguistic tasks (e.g., auditory processing tests, working memory, Lextale, TOEIC listening test), and a background questionnaire (e.g., length of learning English in classroom settings, frequency of conversation in English). Following the speaking task, they were asked to replay the recorded speech and self-assess the accuracy and richness of the lexical items they used. The same speech was then evaluated by native listeners on the same scales. The raw rating scores from the native listeners were then subtracted from those of the participants to compute overconfidence scores. The difference between these assessments was used to calculate 'overconfidence scores,' indicating how much participants overrated or underrated their vocabulary skills (appropriate use and richness) in English.

Regarding the first research question, the results of the correlation analysis and computation of effect size revealed a small-to-medium discrepancies between self- and other-assessment in terms of accuracy of vocabulary use ( $r = .22$ , Cohen's  $d = .37$ ) and richness of vocabulary ( $r = .23$ , Cohen's  $d = .52$ ). Further correlation analyses between native speakers' evaluation scores and the participants' overconfidence scores showed that learners who underestimated their ability were in fact better in terms of appropriate selection of lexical items ( $r = -.64$ ,  $p < .001$ ) and richness in the use of lexical items ( $r = -.40$ ,  $p < .001$ ). In terms of the second research question, multiple regression analyses demonstrated that the participants' self-assessment of appropriate use of lexical items and richness of the items they used in the speech were affected by aural lexical knowledge (e.g., scores from the auditory yes/no test). Based on these findings, pedagogical implications are suggested to help learners improve their accurate perception of their own vocabulary knowledge and usage.

---

\*Speaker

**Keywords:** selfassessment, vocabulary, speech proficiency, individual differences

# The role of idiomaticity in fluent speech: The case of listener-based judgement in an argumentative speech task

Kotaro Takizawa \* <sup>1</sup>

<sup>1</sup> Waseda University [Tokyo, Japan] – Japan

Oral fluency is a major determinant of speaking proficiency (Iwashita et al., 2008). In particular, a subdimension of oral fluency, *perceived fluency* (PF), is a relevant construct in language assessment as how raters perceive speaker's fluency directly leads to speaking scores. It has been suggested that PF is influenced largely by speaker's utterance fluency (UF) (Suzuki et al., 2021). What is less known, however, is lexicogrammatical factors employed in speech that influence PF (e.g., Rossiter, 2009). Recently, a growing amount of attention has been paid to the role of multiword sequences (MWS) in oral fluency. Prior studies suggest that MWS, as either underlying cognitive fluency or produced by being elicited in speech tasks, is associated with UF (e.g., Kahng, 2020; Tavakoli & Uchihara, 2019). Yet, how speaker's use of MWS in speech affects rater's perception of fluency is relatively underexplored. This study responded to Foster's (2020) claim that "perception of speaker fluency is...a perception of idiomaticity" (p. 449) by taking into account the influence of speaker's UF in statistical prediction.

A total of 102 Japanese university students gave an argumentative speech. The speech samples were then normalized and cut from the beginning for 60-second excerpts for rating. To ensure that the speech samples were full range of fluency levels at each 3-day session, a cluster analysis was conducted for speech rate, resulting in two groups. Speech samples ( $n = 17$ ) were extracted from each group to each day, resulting in 34 speech samples at each session. Raters consisted of 11 graduate students in Applied linguistics and a teacher with a doctoral degree in Applied linguistics. The experiment was conducted online using Praat script (Boersma & Weenink, 2020). Rater training was conducted to help understand the definition of fluency (speed, breakdown, and repair fluency; Tavakoli & Skehan, 2005) and help them focus only on fluency features to ensure validity and reliability. 6-point scale was used for rating, considering the level range of the speakers (Suzuki et al., 2021). UF was computed based on Tavakoli and Skehan's speed, breakdown, and repair fluency. As a measure of MWS, bigrams and trigrams of speech texts were extracted and the frequency, mutual information, and proportion 30K were computed with reference to the spoken subsection of Corpus of Contemporary American English using TAALES (Kyle & Crossley, 2015).

PF rating was internally consistent (Cronbach alpha = .975). Correlation analyses showed the significant association between UF and PF ( $\rho = -.524-.916-$ ), UF and bigram/trigram proportion ( $\rho = -.197-.294-$ ), and PF and bigram/trigram proportion ( $\rho = .268, .339$ ). Partial correlations between PF and ngram proportion, while controlling for the effects of UF, demonstrated that PF and bigram proportion were still significantly related ( $\rho = .234-.407$ ). Mixed-effects regression analyses, with speaker and rater being random effects, further showed that bigram proportion significantly predicted PF, taking into account significant UF predictors. The current findings support Foster's (2020) argument that "perception of speaker fluency is...a perception of idiomaticity" (p. 449). Implications for teaching and assessment of fluency will be discussed with audience.

---

\*Speaker

**Keywords:** multiword sequence, oral fluency, perceived fluency

# L2 vowel and stress acquisition: results from a perceptual training at the onset of learning

Gabriela Tavares \* <sup>1</sup>, Susana Correia<sup>†</sup> <sup>1</sup>, Andrea Deme<sup>‡</sup> <sup>2</sup>

<sup>1</sup> Centro de Linguística da Universidade Nova de Lisboa – Portugal

<sup>2</sup> Eötvös Loránd University - Department of Applied Linguistics and Phonetics – Hungary

Research in L2 speech learning has been witnessing an increasing number of studies on perceptual training (1). However, the effects of perceptual training on segmental and suprasegmental properties remain understudied in L2 speech learning. While most training studies focus on segmental acquisition (vowels and consonants) (2-8), or, at a suprasegmental level, on tone (9-12), word stress is rarely investigated. Additionally, studies on L2 word stress rarely approach vowel reduction. Another relevant aspect is that perceptual training has been conducted mainly with learners at intermediate or advanced L2 levels (2-9). The present study addresses these questions by conducting an online high variability perceptual training with adult Hungarian speakers learning European Portuguese (EP).

Sixty-two Hungarian university students enrolled in Portuguese beginner courses were recruited and randomly assigned to three groups. Group *Vowels* (n=20) received training on vowels, Group *Stress* (n=22) was trained on word stress, and Group *Vowels & Stress* (n=24) completed a training on word stress combined with stress related vowel reduction, a crucial feature to perception by EP native speakers (13). Each training program included six sessions conducted over six weeks, consisting of AX and AXB discrimination tasks with feedback. Before and after the training, participants' perception in EP vowels and EP stress contrasts was tested by means of oddity discrimination tasks. After the training, other than the repetition of the pretest trials, participants were also presented with generalization trials, that included new stimuli produced by novel talkers.

Comparisons between pretest, post-test and generalization trials showed no effect of *group* in vowel discrimination, that is, the performance of the three groups was similar in the three testing moments. As for the perception of stress contrasts, we found a significant effect of *group* from pretest to post-test ( $\chi^2(4, N = 7) = 32.07, p < .001$ ), with groups *Stress* and *Vowels & Stress* performing significantly better than Group *Vowels*. The results of the generalization test showed that only Group *Stress* improved, although the comparison between this and the other groups was not significant.

Our results suggest, firstly, that perceptual training in vowel discrimination may not be effective at the onset of learning, contrary to the results of studies with more advanced learners (2-5). Secondly, perceptual training focusing on stress contrasts can be effective, suggesting that stress 'deafness' may not be persistent, contradicting the results of previous studies (14-15). Finally, training with stimuli focused exclusively on stress contrasts may be more effective than training combining stress contrasts with vowel reduction.

---

\*Speaker

<sup>†</sup>Corresponding author: susanacorreia@fsh.unl.pt

<sup>‡</sup>Corresponding author: deme.andrea@btk.elte.hu

The results obtained in this study have important implications for L2 speech learning and classroom interventions. Among other aspects, at the onset of L2 speech learning, perceptual training seems to be effective in suprasegmental features, but not in vowel acquisition. This suggests a similar learning path between L1 and L2 phonological acquisition, since in L1, suprasegmental features are also acquired first (16). The lack of improvement in the generalization trials indicates that a six-session perceptual training is probably not sufficient to develop abstract representations, transferable to novel items.

**Keywords:** SLA, perceptual training, vowels, word stress, vowel reduction, onset of learning.

# A study on the three-way interaction between speech perception, production, and orthography in L2 German

Megumi Terada \* <sup>1</sup>

<sup>1</sup> Humboldt Universität zu Berlin – Germany

Numerous studies have shown that the orthographic form influences both speech perception and production in second language acquisition, which is referred to as the *orthographic effect* (Bassetti 2023). While a significant amount of research has been conducted on English as the target language, relatively little is known about other languages. For German, research projects have identified both negative and positive effects of orthography on speech perception and production in terms of final devoicing (e.g. Hayes-Harb et al., 2018) as well as vowel duration and quality (e.g. Nimz, 2016). Nevertheless, further systematic research is needed.

In my doctoral research, I will investigate how the perception and production are related on both orthographic and auditory-acoustic levels, since it remains unclear whether the non-target-like pronunciation can be explained by the orthographic effect or by the difficulties in sound perception and/or production (cf. Flege, 1995). For this purpose, different tasks are combined in a within-speaker design: shadowing task (with and without additional orthographic input), dictation task, reading aloud, and spontaneous speech. It is crucial to include both auditory and written input in the study design, as multiple studies have demonstrated that the input modality affects the production (Kato & Baese-Bark, 2020; Uchihara et al., 2022). Furthermore, as Bassetti (2023: 22) points out, there is limited research on the orthographic effect in spontaneous speech.

The participants of this study are experienced German learners with two L1 groups: American English and Japanese (n=20 for each group). The targets are umlauts (< ä> , < ö> , < ü> ) in a stressed position and in different phonological contexts, such as onset, consonant cluster, and coda. Together with the speech data, their learning context, attitudes towards pronunciation, and subjective impressions of the experiment will be assessed (cf. Bassetti, 2023). Spoken output will be acoustically analysed on Praat, and the relationship between experimental results and speaker-related variables will be statistically tested using generalised mixed linear effect models. Finally, a rating study will be conducted to assess comprehensibility and accent-ness as reported by Uchihara et al. (2022).

References:

- Bassetti, B. (2023). *Effects of orthography on second language phonology: Learning, awareness, perception and production* (Routledge research in language education). New York.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 92, 233-277.
- Hayes-Harb, R., Brown, K., & Smith, B. L. (2018). Orthographic input and the acquisition of German final devoicing by native speakers of English. *Language and speech*, 61 (4), 547-564.
- Kato, M., & Baese-Berk, M. M. (2020). The effect of input prompts on the relationship between perception and production of non-native sounds. *Journal of Phonetics*, 79, 100964.

---

\*Speaker

Nimz, K. (2016). *Sound perception and production in a foreign language: Die Rolle der Orthographie in der fremdsprachlichen Lautperzeption und-produktion* (Vol. 9). Universitätsverlag Potsdam.

Uchihara, T., Webb, S., Saito, K., & Trofimovich, P. (2022). Does Mode of Input Affect How Second Language Learners Create Form–Meaning Connections and Pronounce Second Language Words?. *The Modern Language Journal*, 106 (2), 351-370.

**Keywords:** orthographic effect, pronunciation, L2 German



# The acquisition of gender in adolescent German learners of Spanish: Evidence from production and perception

Clara Terlaak <sup>\*† 1</sup>, Sarah Schimke <sup>\*</sup>

<sup>1</sup>, Johanna Wolf <sup>\*</sup>

1

<sup>1</sup> Ludwig Maximilian University [Munich] = Ludwig Maximilians Universität München – Germany

The acquisition of L2 gender assignment and agreement is a well-known challenge. Even highly proficient L2 learners are often struggling with the spontaneous production of gender agreement markings (Dewaele/Véronique 2001; Franceschina 2005; Ayoun 2018). Previous research has identified the transparency and salience of the agreement assignment and marking system in the respective L2 (Arnon/Ramskar 2012) as an important factor of influence. The current study looks at German-speaking instructed adolescent learners of Spanish, and focuses on salience of different agreement targets, as well as on the relation between language perception and production. We ask:

- whether agreement errors are more likely to disrupt processing when they are marked both on an adjective and an article, using an eyetracking-during-reading paradigm.
- whether markings that appear not to be salient during reading are more error-prone in spontaneous written production.
- whether learners who show more sensitivity to agreement markings during reading produce a higher number of correct markings in a gender production task.

To address these questions, we collected written frog story retellings in a group of adolescent learners of Spanish (n = 51). In a second study targeting the same population, adolescent learners of Spanish (n = 22 so far, data collection is ongoing) read short paragraphs of texts while their eye movements were registered, and completed a written gender assignment and agreement task with lexical items different from the reading task. In the stimuli of the reading task, we manipulated whether critical noun phrases (underlined) contained an agreement error (conditions 2,4) or not (conditions 1,3), and whether agreement was marked on the article only (conditions 1,2) or also on an adjective (conditions 3,4). Note that all items contained nouns with highly regular endings, so that gender assignment should be facilitated.

1) agreement, adjective invariable

Allí el escritor triste caminó durante dos horas.

---

\*Speaker

†Corresponding author: C.Terlaak@campus.lmu.de

2) no agreement, adjective invariable

Allí la escritor triste caminó durante dos horas.

3) agreement, adjective variable

Allí el escritor lindo caminó durante dos horas.

4) no agreement, adjective variable

Allí la escritor linda caminó durante dos horas.

”There the(MASC/FEM) sad(INVARIABLE) /cute(MASC/FEM) writer(MASC) walked for two hours.”

Preliminary analyses of data from twenty learners suggest that agreement errors lead to significantly longer first-pass and total reading times on the critical region, and that this effect was not modulated by the presence of agreement on the adjective. Moreover, there was a moderate positive relation between individual learner’s difference score (reading times for ungrammatical minus grammatical items) and performance in the production task. Finally, in line with the eyetracking results, a corpus analysis of the frog story data revealed particular high variability and error rates on adjectives.

We conclude that article-noun combinations are salient markers of agreement in L2 Spanish, and that learners who are sensitive to violations during reading also tend to be more successful in producing correct gender markings. We will discuss these results in light of different models of gender acquisition, and will also relate them to results from an ongoing parallel data collection on L2 French.

**Keywords:** gender assignment, gender agreement, adolescent learners, language perception, language production, eyetracking, foreign language learning

# The volatility of the moral foreign-language effect may rest in its making: A pupillometry approach

Dieter Thoma \*<sup>1</sup>, Pernelle Lorette \*

2

<sup>1</sup> University of Mannheim – Germany

<sup>2</sup> University of Mannheim – Germany

In moral decision-making, people normally have a deontological bias, i.e., they prefer options that comply with moral norms regardless of consequences (e.g., do not harm or kill humans) to utilitarian ones that maximize societal welfare (e.g., harm one for the benefit of many). Interestingly, this deontological bias seems to be reduced if unbalanced bilinguals make moral decisions in their weaker second language (L2), compared to their dominant first language (L1). However, the number of studies failing to replicate the so-called moral Foreign Language Effect (FLE) accrues.

To apprehend the volatility of the moral FLE and to better understand its causes, we investigated the language-dependent processes during the problem perception and decision-making process preceding the decision outcomes. We designed a 2 (language: L1 German — L2 English) by 2 (dilemma type: impersonal — personal) auditory pupillometry and decision experiment with language as between- and dilemma type as within-subject factors. German-English bilinguals ( $N = 84$ ) participated in the lab and self-assessed their proficiency in both languages. While looking at an eye-tracked fixation screen, they listened to 12 standard moral dilemmas (6 where the agent acted personally to induce the harm and 6 where the agent was only indirectly involved in the immoral act) as well as 4 trivial filler dilemmas. Each dilemma required a binary choice between a deontological and utilitarian option.

The decisions replicated a strong main effect of dilemma type with a deontological bias only in personal dilemmas, yet we found no moral FLE in decision outcomes. Changes in pupil size indexing emotional arousal while listening to the moral dilemmas mirrored the decision outcome, with higher arousal during personal vs. impersonal dilemmas across languages (see Figure 1). Eye-blinks further indicated that personal dilemmas recruited more attentional cognitive resources than impersonal ones. Although decision *outcomes* did not differ between languages, mixed-effects regression models revealed subtle language-dependent differences in the *process* of problem perception and decision-making. Pupil change and eye-blink data indicated an earlier resolution of the personal-impersonal contrast in L1 than in L2. We interpret this as an initially heightened sensitivity to moral differences in L1 due to more efficient semantic processing. When decision-makers are advanced L2ers with high task engagement, these initial processing differences may vanish at the level of decision outcomes because L2 semantic processing catches up in the evaluation phase. These findings call for additional research that goes beyond decision outcomes and investigates language-dependent problem perception and decision-making process

---

\*Speaker

to understand "bilingual morals" reflected in decision outcomes.

Figure 1. Pupil change during listening to personal vs. impersonal decision dilemmas (onset at 0 ms, differentiation visible between 22000 and 32000 ms).

**Keywords:** foreign language effect, morals, emotion, cognition, pupillometry

# Enriching L2 writing data with community-driven vs. crowdsourced comparative judgement: a validity study

Peter Thwaites \* <sup>1</sup>, Pauline Jadoulle \*

<sup>1</sup>, Magali Paquot <sup>1</sup>

<sup>1</sup> UCLouvain – Belgium

Comparative Judgement (CJ) is a form of assessment in which teams of judges perform pairwise comparisons of learner performances rather than evaluate each item in isolation. Several studies have argued that CJ's reliability and efficiency make it a potentially valuable L2 writing assessment tool (Paquot et al., 2022; Sims et al., 2020). One particular possible application of CJ is for the enrichment of data used in SLA studies, where proficiency measurement and reporting practices have been the subject of recent criticism (Park et al., 2022).

In CJ, judges make decisions based on their own understandings of a target competence, rather than being guided by a rubric. CJ scholars argue that this increases the method's validity by avoiding the "hyper-specificity" of some rubrics (Pinot de Moira et al., 2022). However, the lack of guidance given to CJ judges has also led to a criticism: since it is not possible to know what criteria each judge uses for selecting one item over another, it is also not possible to evaluate CJ's validity for the assessment of the target construct.

Several studies of L1 writing assessment have explored judges' decision-making processes, generally reporting that they consider a wide range of construct-relevant criteria (e.g. Lesterhuis et al., 2022). However, few studies have explored judges' decisions in an L2 context. Such research would help to clarify CJ's validity for assessing L2 writing. Additionally, no existing study explores how judges' expertise affects their decision-making. This is another important question, since CJ studies sometimes report recruiting judges who lack expertise in the target construct (e.g. via crowdsourcing platforms). While research suggests that such judges can produce evaluations of similar reliability and concurrent validity to those produced by experts (Sims et al., 2020), little is known about how these judges make their decisions.

Here we report on a study involving the analysis of notes that judges make while conducting CJ on L2 argumentative essays. Two distinct data collections are currently being conducted, each involving a different group of judges: one of applied linguists recruited through a community driven approach; the other of users of the crowdsourcing platform Prolific. In each study, 25 essays are evaluated by 30 judges, each of whom performs 10 judgements. After each judgement, the judge makes notes explaining their decision. We independently calculate the reliability and concurrent validity of each group's judgements, then categorise each comment made by each judge according to the CEFR C4 writing rubric (Council of Europe, 2009) criteria of argumentation, range, accuracy, or coherence, or as referring to a construct-irrelevant text feature. We then examine whether the two judging groups (a) consider the full range of CEFR criteria in their judgements, and (b) base their decisions on construct-relevant textual features. Through these analyses, the study provides detailed insight into the reliability and

---

\*Speaker

validity of both community-driven and crowdsourced approaches to using CJ. The study therefore addresses concerns regarding CJ's validity in general, while also helping SLA researchers to choose a form of language assessment suitable for their own research needs.

**Keywords:** Comparative judgement, Writing, Proficiency, Research methods, Assessment

# ChatGPT for delivering tasks in less commonly taught languages

Mireia Toda Cosi \* <sup>1</sup>

<sup>1</sup> University of Maryland [College Park] – United States

Interaction plays a pivotal role in language learning (Long, 1996; Ellis, 2003). However, when learning a less commonly taught language (LCTL) in a non-immersive setting, opportunities for interaction are often limited. While language exchanges are an option, the scarcity of suitable conversational partners in the target language remains. We assess the suitability of ChatGPT for facilitating conversational practice in Catalan. Prior research has highlighted the potential of Artificial Intelligence (AI) tools for language learning (Bibauw et al., 2023). AI has a potential that calls for a careful and critical evaluation of its usability as a learning tool, with both its positive and negative implications, especially when it may increase access among LCTLs learners.

We assess the suitability of ChatGPT as a conversational partner at two levels: (1) ecological validity in its interactions and (2) its ability to effectively adjust the difficulty of its interactions. Ten participants at the B1 CEFR level in Catalan with diverse language backgrounds will engage in three 10-15 minute chat interactions with ChatGPT. These interactions are structured as tasks based on participants' self-reported needs. The tasks consist of discussing a YouTube video, engaging in a debate on a polarizing topic over a family chat, and making dinner arrangements with a friend with particular dietary needs. Each participant will complete all tasks, each of which is randomly assigned to a difficulty manipulation. Difficulty is manipulated as: unspecified proficiency, B1 level, and B2 level. The task order and difficulty is semi-randomized to ensure a balanced representation of the task by difficulty combinations. Along with the tasks, participants will fill out a demographics, motivation (Nagel, 2023), and technology use and attitude questionnaires (Berghal et al., 2023). Additionally, an exit interview will gather qualitative data on participants' reflections of the experience. Furthermore, the interactions with ChatGPT will be evaluated by three native Catalan speakers. They will review the learners' conversation and provide feedback on how they would have responded, assess the naturalness of the responses, task completion rates based on a rubric, and share their impressions about the interaction quality.

Taking together learner reports, chat scripts, and native speaker ratings, we will assess the extent to which ChatGPT is a suitable and ecologically valid conversational partner and the impact, or lack thereof, of matching the AI's proficiency level with the learners'. A thematic analysis will identify recurring themes in the interviews, while chat scripts will undergo linguistic complexity analysis following Bulté and Housen (2012). The results will take into account the motivation and technology attitudes reported by participants.

We anticipate that ChatGPT will be an adequate conversational partner, albeit with somewhat unnatural responses. Participants' experiences and contributions will likely be influenced by their attitudes towards technology, with those holding positive attitudes towards technology reporting a more favorable experience. We will discuss the results within the Input Hypothesis (Krashen, 1982), wherein learners are expected to benefit more from interactions at a level slightly above their own, as measured by increased engagement, interest, and the use of more complex language structures.

---

\*Speaker

**Keywords:** ChatGPT, Catalan, TBLT, interaction



# Individual variation in heritage language morphosyntactic processing: Causal modeling of sensitivity to clitic placement in Bosnian and Serbian in Norway

Aleksandra Tomic \* <sup>1</sup>, Yulia Rodina<sup>†</sup> <sup>1</sup>, Fatih Bayram <sup>2</sup>, Cecile De Cat<sup>‡</sup> <sup>3</sup>

<sup>1</sup> University of Tromsø – Norway

<sup>2</sup> The Arctic University of Norway [Tromsø, Norway] – Norway

<sup>3</sup> University of Leeds – United Kingdom

Heritage language (HL) bilinguals often exhibit more variable morphosyntactic comprehension and production patterns than monolinguals (Scontras & Polinsky, 2019), likely due to exposure to more variable input, more restricted language experience, and processing pressures. Studies of individual variation in HL representation and processing have most frequently included overarching, sweeping predictors of variation, such as language dominance and proficiency. To pinpoint the exact facets of individual language experience which most strongly and directly contribute to morphosyntactic (in)variability in early heritage and later bilinguals, we performed a comprehensive causal inference analysis considering how these facets cause each other, and ultimately the morphosyntactic (in)variability in processing. The morphosyntactic (in)variability was measured by the differences in the processing of marked vs. unmarked pronominal clitic order, in a sample of Bosnian and Serbian heritage speakers (HSs) and later bilinguals in Norway.

Bosnian/Serbian pronominal clitics cluster after the first word/phrase in a clause in the so-called Wackernagel clitic-second position, regardless of the position of the verb to which they serve as arguments. This surface independence from the verb might represent a processing challenge and/or could be under the influence of societal language (SL) with post-verbal object pronoun placement. Bilinguals and potential attriters of HLs with the same or related clitic placement in English-speaking environments indeed produce or accept examples of post-verbal clitics in the marked third position (Serbian, Dimitrijević-Savić, 2008; Bulgarian, Ivanova-Sullivan et al., 2022).

In an online self-paced listening task, Bosnian- and Serbian-Norwegian HSs and potential attriters ( $n = 72$ , 43 born in Norway or immigrated by the age of 6) listened to sentences with fragments containing clitics in the unmarked 2nd position and marked 3rd position and judged each sentence on its well-formedness. Based on the results of the corresponding pilot experiment with speakers in Bosnia, sensitivity to clitic placement was expected to be manifested by slower listening times and decreased acceptability judgments for the fragments with the marked order.

Participants in Norway completed a newly developed comprehensive Heritage Language Questionnaire (HeLEx) for us to obtain comprehensive individual language experience profiles in terms of: onset of SL exposure (cont.), extent of multilingualism, self-rated proficiency in oracy

---

\*Speaker

<sup>†</sup>Corresponding author: yulia.rodina@uit.no

<sup>‡</sup>Corresponding author: c.decat@leeds.ac.uk

and literacy, dominance, media experience in HL, community of HL and SL interlocutors, current amount of HL experience, cumulative HL experience, quantity of code-switching, subjective ethnicity, formal education and literacy in HL.

The results of a complex Structural Equation Model including the hypotheses on all causal relationships among these variables, and ultimately how they all cause morphosyntactic variability, suggest that the later age of SL exposure, leading to more cumulative HL experience, and increased frequency of reading HL and visits to the HL country, lead to the maintenance of the HL and increased sensitivity to the marked 3rd position order.

**Keywords:** heritage language, structural equation modeling, causal inference, clitics

# Introducing multilingual assessment in English-medium instruction schools in India

Jacopo Torregrossa \*<sup>1</sup>, Ianthi Tsimpli<sup>2</sup>, Lina Mukhopadhyay<sup>3</sup>, Vrishali Patil-Ingle<sup>3</sup>

<sup>1</sup> Goethe-University Frankfurt – Germany

<sup>2</sup> University of Cambridge – United Kingdom

<sup>3</sup> The English Foreign Languages University, Hyderabad – India

Some studies have shown that English-medium instruction (EMI) may leave space for the use of pupils' multiple languages during learning (Lin & He, 2017; Sa & Li, 2020). Recent studies demonstrated that this pedagogic approach may even enhance English learning (Anderson & Lightfoot, 2018; Cenoz & Gorter, 2022). However, an inconsistency is often observed between the integration of multilingualism in teaching and the way in which pupils' language abilities in English are assessed, namely in English only. Pupils are thus not allowed to rely on their multilingual repertoire to its fullest potential. We examine the extent to which the integration of multilingual elements in assessment benefits pupils' learning outcomes, focusing on abilities involved in text comprehension.

The study was conducted in six fourth-grade classrooms of low-SES government primary EMI schools in Hyderabad. We tested 205 pupils divided into two groups, a control (CG,  $n=111$ ; *age-range*:8-15; *M*:9;6) and an experimental (EG,  $n=94$ ; *age-range*:8-15; *M*:9;4). All classrooms were multilingual, with all participants speaking at least Telugu (the societal language). Each group was administered two multiple-choice tests assessing comprehension of a narrative and an expository text (environmental science) in English (in two different sessions). The multiple-choice questions involved the ability to retrieve factual information and draw local inferences (i.e., integration of information across sentences). The expository text featured also questions on vocabulary, whereas the narrative text questions involving global inference (i.e., the link between textual information and world-knowledge). The CG was administered questions in English only, the EG received for each question type, half of the questions in Telugu and half in English.

For each test, we conducted a generalized-linear-mixed-effects analysis with response accuracy as dependent variable and the interaction between question type and language mode (one-language vs. two-language) as independent one. For the expository text, we found a significant effect of language mode (better performance in the two-language mode), and a significant language-mode x question-type interaction (better results in the two-language mode in association with vocabulary and factual questions). For the narrative text, we found a better performance in the two-language mode with questions involving global and local inferencing, whereas performance was better in the one-language mode with factual questions. The same pattern was observed with both questions that were in English in both groups and questions that were in English in the CG and Telugu in the EG.

The study reveals that pupils exhibit better text-comprehension abilities if they are allowed to rely on their bi-/multilingual resources during testing. The involvement of Telugu boosted the comprehension of technical, academic vocabulary with expository texts. Narrative texts do not necessarily involve technical vocabulary, which explains the lack of beneficial effects of

---

\*Speaker

the two-language mode on factual questions. However, the two-language mode encouraged the development of complex discourse representations integrating both texts, as reflected in a better performance with inference-making questions. The results speak in favour of the use of multilingual assessment in both language and content learning classrooms, especially in contexts that are highly multilingual like India.

**Keywords:** multilingual assessment, English medium instruction, India, text comprehension, multilingual pedagogies

# Title: From SLA and TBLT constructs to actual decision making: Automated task design for EFL teaching in primary schools

Vera Trager \* <sup>1</sup>, Roger Gilabert \*

2

<sup>1</sup> University of Barcelona – Spain

<sup>2</sup> University of Barcelona – Spain

The global emphasis on mandatory second language education in primary schools has highlighted the importance of understanding Task-Based Language Teaching (TBLT) impact on young learners. Despite TBLT being a focal point of study for over four decades, a notable gap exists in comprehending the design of language learning tasks, especially for primary school children. Little is known about how SLA (e.g. input, output, interaction) and TBLT constructs (e.g. task variables, focus on form) actually transfer to decision making by teachers and designers during task design. To bridge this gap, a web-based task design tool (taskGen) has been developed that combines knowledge from second language acquisition theories, TBLT research findings, as well as natural language processing, and learning design in order to assist teachers and task designers in crafting student-centred task designs. The study aims at analysing the ongoing tool pilot evaluation with a focus on primary school EFL teachers, and it does so in terms of usability, acceptance, and adaptation by delving into their design choices, patterns and processes. Employing a mixed-methods approach, the study follows an iterative process by incorporating three rounds of data collection and analysis. The qualitative aspect involves piloting and evaluating non-functional and functional prototypes through interviews, background questionnaires, and think-aloud sessions-all subjected to thematic analysis using AtlasTI. Working closely with around 50 stakeholders, teachers and pre-service teachers from Catalonia, Austria and Hungary, this method allowed to improve the tool while developing it and adapt it to the teacher's needs. The quantitative analysis features a survey on perceived usefulness and ease following Davis' Technology Acceptance Model (1989), as well as a global database tracking micro-decisions and cataloguing design choices made by teachers. Data will be extracted by means of data analytics, generalized mixed-effects models will be conducted using SPSS and R in order to analyse, explore and visualize teachers' preferred task design choices and processes. System tracking provides the opportunity to observe and analyse these micro design choices on a large scale, including favoured task design options, the duration spent on task design, the choice (or lack thereof) of linguistic features, and choices related to focus on form through input enhancement using text editing tools. This exploration serves as a crucial step in filling a key research gap by enabling a deeper understanding of teachers' approaches to task design. The discussion will revolve around how the analysis of automated task design may shed light on how knowledge coming from SLA and TBLT is instantiated in specific choices during task design.

---

\*Speaker

**Keywords:** Task based language teaching, Task Design, EFL, Primary School

# Event conceptualisation by French-English, Russian-English and Russian-English-French speakers: insights from verbal and non-verbal tasks

Alina Tsikulina \* <sup>1</sup>

<sup>1</sup> Savoirs, Textes, Langage (STL) - UMR 8163 – Université de Lille, Centre National de la Recherche Scientifique – France

In the domain of spatial language, and especially in the domain of motion event encoding, languages follow very different lexicalization patterns (1-2): In some languages, such as French (FR), speakers tend to lexicalize Path information in the verb leaving the expression of Manner optional or peripheral (*verb-framed* type). In others, such as English (EN) and Russian (RU), speakers prefer compact and dense structures, where Manner and Path are encoded together, the one lexicalized in the verb, the other encoded in satellites (*satellite-framed* type) (3-4). Such differences seem to constrain not only how speakers organize information verbally but also how they represent spatial components cognitively (5-7).

For monolinguals, recent experimental work has shown that such linguistic asymmetries influence both verbal and non-verbal behavior (8-9). For second language (L2/L3) learners, previous studies have led to inconsistent conclusions (10): some support that even advanced L2 speakers are mainly guided by their first language (L1) (11); others suggest that advanced learners are capable to reconceptualize events and learn new encoding patterns regardless their L1 (12-13). This study involves a crosslinguistic (French, English, Russian) and cross-sectional (monolinguals, advanced bilinguals and trilinguals) design, and investigates (a) whether the typological distance of the involved languages influences the learnability of such patterns in L2/L3 event encoding, and (b) to what extent such crosslinguistic variability affect cognitive processing beyond language use.

25 participants were tested in one verbal (Production task) and one non-verbal (Similarity Judgement) task. For the production task, participants had to describe video-clips involving voluntary motion events (e.g., *a man walking across a street*). For the similarity judgement task, participants had to watch a target video-clip of a scene involving a specific Path and a specific Manner of motion (e.g., *jumping into a house*), and choose as fast as possible among two variants involving either a change of Manner (*running into*) or a change of Path (*jumping out of a house*) the one that looked most similar to the target. Data were collected from five different groups: 2 groups of monolingual speakers (RU, FR) and 3 groups of advanced learners: L2-EN bilinguals (RU-EN, FR-EN) and L2-EN/L3-FR trilinguals (RU-EN-FR).

The results show that: besides the strong typological differences in the verbal strategies of monolinguals (*verb-framed* in FR, *satellite-framed* in RU), bilinguals and trilinguals, overall, do not show strong L1 transfers when using motion verbs in L2. However, FR-EN bilinguals, although capable to express Manner in the verb and Path in the periphery, they show traces of L1-transfer by double-marking Manner, expressed also in the periphery (not prototypical in EN).

---

\*Speaker

With respect to the non-verbal data: although RU monolinguals show no significant Manner- vs. Path-congruent differences, and bilingual and trilingual RU speakers also manage to adopt EN-like behavior, FR monolinguals show an overall preference for Path-congruence, a tendency that also persists in FR-EN participants' choices.

This study explores the behavioral and algorithmic implications the linguistic constraints and typological distance may have for L2/L3 learning, and contributes more generally to the debate about universal vs. language-specific dimensions of spatial cognition.

**Keywords:** Motion events, Crosslinguistic influence, Psycholinguistics, L2 and L3 acquisition, Cognitive linguistics



# The Effects of Input Modality in Processing L2 Vocabulary and Developing Different Vocabulary Knowledge Types

Ayşen Tuzcu \* <sup>1</sup>

<sup>1</sup> Michigan State University [East Lansing] – United States

Researchers have investigated the promise of bimodal input, i.e., aural and written input presented simultaneously, in enhancing vocabulary learning from meaning-focused activities. Audio support in bimodal input, in particular, has been argued to speed up the detection of words in written input, direct L2 learners' attention to words, and enhance the form-meaning links for new vocabulary (e.g., Long, 2017). However, so far, the research focus has been on the learning outcomes, and there is a lack of studies showing the processing of bimodal input for vocabulary learning. Moreover, the learning outcomes have traditionally been assessed through written, untimed accuracy tests that tap into conscious, verbalizable knowledge of target words. How bimodal input contributes to fluent form and meaning retrieval of the newly learned vocabulary during a real-time activity (e.g., reading) is still unknown and warrants attention. Therefore, this study aimed to address these research gaps by comparing the effectiveness of unimodal and bimodal input in facilitating different cognitive processes and developing different types of vocabulary knowledge.

Sixty-three adult L2 English speakers were randomly assigned to unimodal (i.e., written input only) and bimodal input (i.e., written and aural input) groups, and they read a 9500-word graded reader containing 24 target pseudowords (2 to 16 frequency of occurrences) over two days. Reading processes were recorded using eye-tracking, and four eye-tracking measures tapping into early and late reading processes (i.e., gaze duration, regression path duration, rereading time, and total reading time) were used to investigate participants' attention levels. Learning outcomes were measured with three written tests (form recognition, meaning recall, and meaning recognition) and one sentence-reading test that indicate fluent retrieval of word meaning in real time.

Eye-tracking data were analyzed using growth curve models, and learning gains on vocabulary tests were analyzed using mixed logistic models (form recognition, meaning recall, and meaning recognition) and linear mixed effects models (sentence-reading test). For both input groups, the growth curve models indicated a nonlinear decrease in attention levels from the first encounter with target pseudowords to the last encounter, corroborating the findings of previous research (e.g., Godfroid et al., 2018). However, the two groups showed different reading behaviors during the early and late reading processes. The bimodal input group allocated increased attention to target pseudowords in early reading processes as reflected in gaze durations. The unimodal input group, however, had increased attention to target pseudowords in late reading processes as reflected in rereading times. Moreover, the bimodal input group had significantly higher form recognition scores, particularly on the delayed test. This group also had more robust lexical representations for the target pseudowords as reflected in the reading times in the sentence-reading test. The learning gains of the two groups on the meaning recall and meaning recognition tests were comparable. Overall, these results indicated that aural input influences the reading patterns of L2 readers, and having access to phonologic forms of new words through audio has

---

\*Speaker

an additional positive effect in vocabulary learning. The findings are discussed in the light of theories of SLA and attention.

**Keywords:** input modality, incidental vocabulary learning, processing, eye, tracking

# Secondary analysis in second language research: Examining (un)exploited analysis potentials and preventing potential pitfalls

Leonie Regina Twente \* <sup>1</sup>

<sup>1</sup> Universität zu Köln – Germany

Secondary analysis is not a new concept, especially in the social sciences, where the use of existing data to answer new research questions is common practice. Recently, a wider range of disciplines, including education and psychology, have shown renewed or novel interest in secondary analysis approaches. The generally increasing importance of this research strategy is also in line with efforts to improve research data infrastructures, institutionalize open science, and enable replicability (cf. Weston et al., 2019).

Apart from corpus research, which has a long history of sharing and re-using data, secondary analysis has so far been of limited relevance in (second) language research. However, this is slowly changing, as recent methodological and substantive contributions indicate (e.g., re-analyses in Gudmestad & Edmonds, 2018; Marx & Caspari, 2022). Given the high costs and ethical issues associated with data collection, it is imperative to advance secondary analysis and harness the potentials it offers for our work.

In my dissertation project *Secondary analyses on the language development of newly immigrated school children and adolescents* I address this issue, pursuing two goals.

The first objective is a review of previous, current and potential use of secondary analysis in research focusing on German as a second language in the newly immigrated student population. Expanding work that shows current limitations of secondary analysis in this field (Caspari & Marx, 2022), I examine the availability and eligibility of data from a variety of sources, and the extent to which the analysis potential has been exploited or remains unexploited. The second objective is to conduct secondary analyses to answer questions on the development of German as a second language in newly arrived student populations, and the factors influencing it.

In my presentation, I focus on the first part of my project, specifically answering the following questions as pertaining to research using secondary analysis to investigate language skills in newly arrived students in Germany:

- What (types of) questions have already been addressed?
- How was the secondary analysis implemented?
- Which (types of) data were used?
- What is the scope, eligibility, and quality of the data sources used?

---

\*Speaker

I then discuss the implications emerging from my results for the planned second, empirical part of my dissertation. This will raise issues of interest to (secondary) data analysis more generally, including selection and appraisal of data, identification of target groups in datasets, as well as challenges and potentials of analyzing and combining different kinds of data.

## References

- Caspari, C., & Marx, N. (2022). *Challenges with Immigrant Students' Data in NEPS*. <https://osf.io/m2n9x>
- Gudmestad, A., & Edmonds, A. (Eds.). (2018). *Critical Reflections on Data in Second Language Acquisition* (Vol. 51). John Benjamins Publishing Company. <https://doi.org/10.1075/lllt.51>
- Marx, N., & Caspari, C. (submitted). *Language Skills of Immigrant Schoolchildren in Germany: A Re-Analysis of the NEPS Panel Study Data*. <https://doi.org/10.17605/OSF.IO/3VQE6>
- Weston, S. J., Ritchie, S. J., Rohrer, J. M., & Przybylski, A. K. (2019). Recommendations for Increasing the Transparency of Analysis of Preexisting Data Sets. *Advances in Methods and Practices in Psychological Science*, 2(3), 214–227. <https://doi.org/10.1177/2515245919848684>

**Keywords:** secondary analysis, newly arrived students, German as a second language

# Exploring the compatibility of TBLT and morphologically complex languages: an analysis of learner performance on an L2 Lithuanian task

Richard Udes \* <sup>1</sup>

<sup>1</sup> Vilnius University [Vilnius] – Lithuania

The acquisition of case marking is considered one of the greatest challenges for L2 learners of morphologically complex languages such as Lithuanian (Savickienė, 2006). Task-Based Language Teaching (TBLT) might offer a solution. Within the TBLT framework, the sequencing of case introduction and exposure to inflected forms needs to be based on students' communicative needs in the form of pedagogic tasks (Long, 2015). TBLT, thereby, prioritizes task accomplishment through communication, relegating grammatical analyses to the post-task phase. This should foster both explicit and implicit knowledge of the language and its grammar (Ellis et al., 2020). Nevertheless, previous research suggests that applying TBLT to the teaching of a morphologically complex language does not necessarily lead to advancements in the acquisition of grammatical cases (Markina, 2018). Clearly, code complexity (i.e. the intricacies inherent in the linguistic code) needs to be considered (Gilabert & Castellví, 2019). The present study aims to contribute to testing the viability of the TBLT approach for the teaching of L2 Lithuanian. To pilot a task and tests developed for the experiment in TBLT, ten A1-A2 level students were engaged. The purpose of this pilot study was to trial materials on learners for the target proficiency level (A2) and establish whether any immediate case-related learning gains were observable. The study followed a pre-test – intervention – post-test design. The participants completed a form recognition and a form recall test of explicit knowledge and an elicited oral imitation test measuring their implicit knowledge (Erlam, 2009). As an intervention, they completed one carefully prepared L2 Lithuanian pedagogic task according to the TBLT principles and specifically addressing the target group in terms of topic relevance and use of nominative, genitive, and accusative case marking. One group (n = 5) completed a version of the task with a pronounced focus on grammar, the other (n = 5) without. The intervention lasted 90 minutes. The gains in the post-tests were analysed statistically. This preliminary analysis showed minor learning gains within the group that completed the task with a more pronounced focus on grammar. This is an expected result, although the full task cycle of my PhD project might also stimulate learning gains for the group exposed to less grammar-focused instruction. Furthermore, the study only investigated immediate learning gains. This leaves open the possibility for lasting acquisition which needs to be investigated through delayed post-tests following a longer experimental intervention, i.e. the completion of more tasks.

**Keywords:** TBLT, SLA, Lithuanian, Grammatical case

---

\*Speaker

# Vocabulary learning with online and paper dictionaries in the context of meaning-focused reading

Isabelle Udry <sup>\*†</sup> <sup>1</sup>, Raphael Berthele <sup>1</sup>

<sup>1</sup> Université de Fribourg = University of Fribourg – Switzerland

Incidental vocabulary learning from meaning-focused input can contribute to the acquisition of the vocabulary needed for successful communication in a foreign language (Nation, 2022). One of the frameworks for designing meaning-focused vocabulary learning tasks is the Involvement Load Hypothesis (ILH) put forward by Laufer and Hulstijn (2001) and developed further by Yanagisawa and Webb (2021). The ILH posits that task effectiveness relies on three components inherent in the task: *need*, *search*, and *evaluation*. While the ILH has been studied extensively from various angles (Hazrat & Read, 2021), there is comparatively little research on the search component, i.e. the process of looking up word meaning in a dictionary (for exceptions see e.g. Hazrat, 2020 or Zou, 2012). Re-examining the search component seems justified given that novel language supporting devices, including online dictionaries and translators, have transformed the way in which meanings of words can be accessed. Although these tools are widely used (Niño, 2020), their contribution to vocabulary learning through meaning-focused input remains unclear. Furthermore, education policy may be disregarding technological change, and the promotion of paper dictionaries persists despite their waning presence in modern classrooms (for the Swiss context see Udry & Berthele, 2023).

While paper dictionaries as a *medium* are outdated, their use could remain legitimate in terms of *learning* if they can help optimize vocabulary retention. Based on the assumption that different search conditions induce different levels of cognitive engagement, it can be hypothesized that the method used for accessing meaning affects the extent of vocabulary retention. The present study investigates this claim in the context of reading: 370 intermediate learners of English (L1 German or French) completed a meaning-focused vocabulary learning task in which the ILH index varied only in terms of the search component. Participants read three short texts, each followed by a set of comprehension questions. Before answering the questions, 15 target words highlighted as relevant for comprehension were translated from L2 to L1 in one of three search conditions: online dictionary (n=126), paper dictionary (n=121), or on-screen glossary (n=123). Following the reading task, participants completed an unannounced vocabulary post-test for meaning recall, both immediately and two weeks later.

Regression analyses with mixed effects models reveal a significant impact of the search condition on vocabulary retention in the immediate post-test. The dictionary conditions demonstrate greater learning gains than the glossary condition, with use of online dictionaries leading to the highest level of vocabulary retention followed by paper dictionaries. These effects are not maintained in the delayed post-test.

The results suggest that the use of paper dictionaries does not lead to better retention, and therefore, cannot be supported as a beneficial resource in the classroom, at least in terms of vocabulary learning. Also, the degree of *search* varies depending on the medium when early

---

\*Speaker

†Corresponding author: isabelle.udry@unifr.ch

stages of vocabulary learning are considered. Moreover, the decrease in vocabulary test scores between the immediate and delayed post-test in all conditions emphasises the need to consolidate vocabulary knowledge.

**Keywords:** language pedagogy, digital learning, vocabulary learning

# Exploring the active and passive dimension of subtitling in L3-Dutch teaching for German schools and its impact on incidental vocabulary acquisition in comparison

Lukas Urbanek \* <sup>1</sup>

<sup>1</sup> University of Münster – Germany

In recent years, research has examined two different approaches to audiovisual input with on-screen text: "Subtitles as a Support" and "Subtitling as a Task" (Talaván, 2010). These approaches have the potential to improve a variety of foreign language skills and have been studied specifically in higher education settings concerning English as a source (L1) or target (L2) language in combination with predominantly Romance languages (see Lertola, 2019a; Montero Perez, 2022 for reviews). Talaván (2006) assumes that vocabulary acquisition "is enhanced when subtitles are used not in a passive but in an active mode, being the students themselves the ones who create them". This is consistent with the Involvement-Load Hypothesis (Laufer & Hulstijn, 2001), which receives strong empirical support, particularly in reading activities (e.g., Yanagisawa & Webb, 2021).

This study presents two sub-studies that aim to promote incidental vocabulary acquisition by over 300 L3-Dutch learners in German schools. Talaván's (2006) assumption is addressed by having participants watch the Dutch short film "Spelletjesavond" twice under different conditions, without or with activated subtitles (in L1-German, L2-English, L3-Dutch, or L1-German/L3-Dutch combined) in sub-study 1. In sub-study 2, participants engaged in active interlingual subtitling of an excerpt from the Dutch short film "Alles mag" utilizing a browser-based subtitle editor. Pre- and post-tests on form and meaning recognition in addition to meaning recall were administered to assess learners' vocabulary progress.

The results indicate that learners in sub-study 1 significantly increased their vocabulary in all conditions after watching the short film, although to varying extents (in line with previous research by Wang & Pellicer-Sánchez, 2022). Moreover, active subtitling in sub-study 2 had a significant impact on vocabulary acquisition (in line with previous research by Lertola, 2019b). Bringing both sub-studies together and leaving aside the divergent audiovisual input and the different target vocabulary items tested, the effect sizes per sub-study and the direct statistical comparison indicate that output-based active subtitling is more beneficial for incidental vocabulary acquisition compared to input-based viewing of the short film without additional activities. This supports the Involvement-Load Hypothesis, which has already been discussed for subtitling and language learning (Lertola, 2019b), but has scarcely been tested empirically in this context.

**Keywords:** Audiovisual translation, Subtitling, Task, Support, Comparison, Involvement load, Dutch as a foreign language, German school setting

---

\*Speaker



# Extramural English in the Turkish context: Scale development and the relationship with English proficiency

Mehmet Sercan Uztosun <sup>\*† 1</sup>, Pia Sundqvist <sup>‡ 2</sup>

<sup>1</sup> Norwegian University of Science and Technology – Norway

<sup>2</sup> University of Oslo – Norway

Extramural English (EE) has recently become an emerging and evolving research field in SLA research (Schwarz, 2020). The concept provides a theoretical basis for out-of-class and self-initiated English activities that may contribute English language learning either incidentally or intentionally (Sundqvist, 2009). Previous research studies showed that EE plays an important role in language learning as it correlates with language proficiency, including vocabulary knowledge (e.g. De Wilde & Eyckmans, 2017; Sundqvist, 2009, 2019), speaking competence (Sundqvist & Wikström, 2015) and reading and listening comprehension skills (Sylvén & Sundqvist, 2012). Most studies measured EE using questionnaires (Lee, 2022), which presents a methodological limitation because questionnaires may not provide reliable and valid scores (Dörnyei and Taguchi, 2010). Hence, there is a need for the development of scales that measure EE in a reliable and valid ways. This need was addressed by Sundqvist and Uztosun (in press) who developed EE scale in the Scandinavian context and called for a need to develop EE scales in different contexts. This study was designed to address this call. It aimed to develop the Turkish version of the EE scale (TEE scale) and reveal whether EE engagement is related to English as a foreign language (EFL) proficiency.

TEE scale was developed in two stages: exploratory factor analysis ( $N = 308$ ) and confirmatory factor analysis ( $N = 310$ ), which were followed by measuring its test-retest reliability ( $N = 57$ ) and known-groups validity ( $N = 102$ ). The data were collected from English language teaching departments at seven universities in Türkiye. The analyses led to a 47-item TEE Scale that is loaded onto nine factors: 1) EE Digital Creativity, 2) EE Gaming, 3) EE Music, 4) EE Social Interaction, 5) EE Reading and listening, 6) EE Internalised, 7) EE Writing, 8) EE Social reading and listening, 9) EE googling and viewing.

TEE scale was administered to 56 students at a university in Istanbul who took preparatory exemption examination that comprised three tests: reading/listening, writing, and speaking. In addition to completing the scale, the participants were invited to submit their test scores. The data were analysed using Pearson correlation analysis.

The finding revealed that EE correlated with three test scores at moderate levels: total exam score  $r(56) = 0.48$ ;  $p < .01$ ), reading and listening ( $r(56) = 0.38$ ;  $p < .01$ ) and speaking  $r(56) = 0.46$ ;  $p < .01$ ). Strong correlations were found between speaking test scores and EE internalized  $r(56) = 0.60$ ;  $p < .01$ ) and total exam score and EE internalized  $r(56) = 0.52$ ;  $p < .01$ ). Further significant positive correlations were found between test scores and different

---

\*Speaker

†Corresponding author: mehmet.s.uztosun@ntnu.no

‡Corresponding author: pia.sundqvist@ils.uio.no

factors in the TEE scale.

These findings revealed that EE is related with English language proficiency yet different EE activities play distinct roles. The findings also confirmed that EE is context-specific in nature because the TEE scale led different factors than EE scale. Implications are discussed regarding the need for developing EE scales in different contexts and role of EE in language learning.

**Keywords:** Extramural English, incidental L2 learning, informal language learning, language proficiency, scale development

# Measuring the impact of prosody awareness raising: A training programme for adult inexperienced L2-English learners

Stella Ville \* <sup>1</sup>

<sup>1</sup> University of the Balearic Islands – Spain

Prosody has become more popular in recent years in SLA research (Alazard, 2010; Frost, 2016). Because of its physical nature, pedagogy often concentrates on embodied approaches (Carrera-Sabaté, 2021; Baills, 2022). A few studies focus on inexperienced adult learners. This study addresses the question of raising awareness to prosodic features with this set of learners. Teachers, students, and researchers acknowledge the importance of pronunciation. Since Guberina (1965) postulated its primacy, prosody has proven essential to L1 and L2 acquisition (Kjellin, 1999) and perception (Hahn, 2004; DeMeo, 2012), which led authors to suggest it should be prioritised when teaching pronunciation (Gilbert, 2008). However, students generally have little awareness about prosody, and teachers tend not to view it as essential (McNerney & Mendelsohn, 1992). Prosody is considered the most difficult L2 feature to acquire (Brown, 2000) and to teach (Burgess & Spencer, 2000), and as teachers generally lack training in teaching pronunciation (Levis, 2021), the gap between research and practices understandably remains.

Pronunciation instructions generally yield positive results (Thomson & Derwing, 2015), but often rely on skills that learners may not have developed yet. If all methods share the goal of learners' acquisition of language-specific pronunciation features, an interesting starting point is initial awareness-raising (Carlet & DeSouza, 2018). To that end, I designed and implemented a training programme to work on perception and production of key features of English and challenge learners' conception of the language learning process. "Investigations of instruction-awareness links are uncommon" (Kennedy, 2014). The research question is whether low-level learners can benefit from a multimodal programme aimed at raising awareness on prosody, in terms of performance and attitude, thus laying favourable foundations for further learning. I hypothesised that:

H1. The training programme will raise learners' awareness.

H2. Raising awareness will:

- a) Lead to improved oral production at T2;
- b) Help students reach a more accurate understanding of their current performances;
- c) Positively impact attitudes;

H3. The A1 groups will show greater changes than the A2 group.

The study was carried out in Spain, where 85 inexperienced adult Spanish-native participants received a 9-hour training over 11 weeks and completed a 3-task pre- and post-test: reading

---

\*Speaker

aloud, picture description and utterance repetition. I collected data from questionnaires and interviews (attitudes, habits, self-assessment, stress placement) to explore the dynamic connections between students self-reported and actual performances and investigate the impact of the approach. I am still exploring the data, but preliminary results are encouraging. Further analyses include T1-T2 compared acoustic measures to evaluate learning, contrasted with qualitative data to assess awareness and attitudes; and compared A1 vs. A2 groups results, adding to the idea that "the first-year context leads to greatest improvement" (Munro & Derwing, 2015). The goal of this research is to contribute to the development of teaching methods, informing as to the usefulness of prosody awareness-raising, and optimal ways and time to achieve it. Future research directions include replication with different languages, broader feedback collection, and a more longitudinal approach (more delayed post-tests, inclusion of a control group, etc.).

**Keywords:** prosody, SLA, embodied pronunciation, awareness, multimodality

# Individual versus collaborative pre-task planning in L2 English written production

Arturo Valera-Garcia \*<sup>1</sup>, Raquel Criado<sup>†</sup><sup>1</sup>

<sup>1</sup> University of Murcia – Spain

As Ellis (2021) claims, pre-task planning in L2 writing is grounded on both theoretical and pedagogical considerations. However, while research on pre-task planning has mostly focused on oral tasks, L2 writing empirical evidence has been scarce and mixed (McDonough & De Vleeschauwer, 2019; Tabari, 2023). These results could be attributed, inter alia, to the methodological variation of the previous studies regarding the control or not of the total time on task, the selection and operationalization of CAF measures, and the presence or absence of a pre-task training period (Wu & Ellis, 2023). Thus, the aim of this quasi-experimental study is to contribute to the body of research by comparing the effects of collaborative vs. individual pre-task planning on the quality of secondary-school students' L2 English written production in terms of CAF measures. The study adopted a pre-/post-test design in which group 1 (n=8) was assigned to the individual pre-task planning condition, while group 2 (n=10) was assigned to the collaborative pre-task planning condition. The data collection comprised three phases. In the first one, the participants completed the Oxford Placement Test (Allen, 2004) and they were asked to write a pre-(diagnostic) task. During the intervention stage (second phase), the participants were involved in two sessions of pre-task planning training. They were introduced to the role and importance of planning a written text, received guidelines on how to develop an outline and devised two structured outlines for an opinion essay on a planning worksheet (adapted from Neumann & McDonough, 2015). The opinion essay task belonged to the students' official L2 English syllabus. As for the post-test (third phase), the participants created an outline for a different opinion essay—individually or collaboratively—and wrote their final texts individually in both conditions. They were given 10 minutes to plan their texts and 30 minutes to compose them. Participants in the collaborative condition planned their texts in self-selected pairs. The texts were analysed considering CAF measures. Results will be discussed in light of the cognitive theoretical underpinnings of pre-task planning in L1 writing, which have usually been applied to L2 writing: the Overload Hypothesis and the Interaction Hypothesis (Kellogg, 1990). Pedagogical implications for secondary education classes will be included.

**Keywords:** Pre Task Planning, L2 Writing, CAF measures

---

\*Speaker

<sup>†</sup>Corresponding author: rcriado@um.es

# L1-to-L2 and L2-to-L1 cross-linguistic structural priming: the role of verb bias effects

Chantal Van Dijk <sup>\*† 1,2</sup>, Holger Hopp <sup>1</sup>

<sup>1</sup> English Linguistics, TU Braunschweig – Germany

<sup>2</sup> Institute of Linguistics, University of Stuttgart – Germany

A possible key mechanism behind structural priming is surprisal: less expected structures yielding larger priming according to error-based learning; e.g., Chang et al., 2006). However, its role in *cross-linguistic* structural priming is unclear. Surprisal should hold for L1-to-L2 and L2-to-L1 priming, but evidence is circumstantial: either only L1-to-L2 priming showing longer-term priming consistent with error-based models (Hwang & Shin, 2019; also see Unsworth (2023)), or only bidirectional short-term priming within L2 speakers was studied (Kootstra & Doedens, 2016). We investigate surprisal in L1-to-L2 and L2-to-L1 priming with the same prime (German) and target language (English) in L1-German *and* in L1-English speakers, examining the following questions:

RQ1) Is there evidence of short-term and longer-term priming?

RQ2) Is there evidence of surprisal due to prime-verb bias?

RQ3) Does priming strength depend on priming direction?

Ninety-two L1-German-L2-English adults and 89 L1-English-L2-German adults participated. A baseline task measured production preferences for English and German ditransitive verbs (Double Object (DO) structures: "She showed the pirate the painting" and Prepositional Object (PO) structures: "She showed the painting to the pirate") through picture descriptions. In a German-to-English priming task, a spoken DO or PO prime preceded each picture description. There were 4 English target verbs, 4 strongly-DO-biased and 4 weakly-DO-biased German prime verbs which were PO-biased in their English translation (Şafak, 2022). A post-test measured English DO/PO production afterwards. We predicted that if German prime-verb biases drive priming, surprisal and therefore PO is stronger than DO priming and L1-German participants are more sensitive to prime-verb biases, yielding stronger priming.

Generalized linear mixed effects models showed significant short-term and longer-term (i.e. baseline to post-test) priming. Prime-verb bias and priming direction significantly modulated priming. German prime-verb biases predicted short-term priming in L1-German but not in L1-English participants. Instead, English prime-verb translation-equivalents' biases predicted short-term priming in L1-English but not in L1-German participants. In line with these short-term effects, L1-German and L1-English participants showed longer-term PO and DO priming, respectively.

Our study is one of the first to show bidirectional short *and* longer-term cross-linguistic structural priming. Prime-verb bias modulated L1-to-L2 short-term priming, indicating that German prime-verb bias carries over in English. Strikingly, biases of L1-translation equivalents of German prime verbs modulated L2-to-L1 priming, indicating that in cross-linguistic priming L1

---

\*Speaker

†Corresponding author: c.vandijk@tu-braunschweig.de

prime-verb bias carries over, regardless of prime and target language. Regarding longer-term priming, participants showed sensitivity to L1 properties, in line with surprisal. German PO primes yielded cumulative priming in L1-German participants instead of less surprising DOs. Instead, English PO-biased prime-verb translation equivalents yielded cumulative DO priming in L1-English participants.

Summarizing, although participants received the same exposure, short-term and longer-term priming behaviour differed, indicating that participants' L1 baseline preferences are crucial for priming outcomes, in line with surprisal in error-based learning (e.g., Chang et al., 2006).

References

Chang et al. (2006). Becoming syntactic. *Psychological Review*.

Hwang & Shin (2019). Cumulative effects of syntactic experience. *Journal of Memory and Language*.

Kootstra & Doedens (2016). Bilingual syntactic choice. *Bilingualism: Language and Cognition*.

Unsworth (2023). Shared syntax and cross-linguistic influence. *Linguistic Approaches to Bilingualism*.

**Keywords:** crosslinguistic structural priming, verb bias, surprisal effects, dative alternation

# Crowdsourcing L2 proficiency assessment using comparative judgement: What is really important to novice judges?

Nathan Vandeweerd <sup>\*† 1</sup>, Peter Thwaites <sup>\* ‡ 2</sup>, Magali Paquot<sup>§ 2</sup>

<sup>1</sup> Radboud university [Nijmegen] – Netherlands

<sup>2</sup> Université Catholique de Louvain = Catholic University of Louvain – Belgium

Proficiency is one of the most important constructs in second language acquisition research, yet many studies continue to either under-report information about the proficiency of the learners in their studies, or use unreliable indicators such as institutional status (Park et al., 2022). Comparative Judgement (CJ) provides a quick, cost-effective, and reliable way to assess large numbers of learner texts (Wheadon et al., 2020), and might therefore be seen as a possible solution to this problem. One additional advantage of CJ in this context is that judges can be rapidly recruited through crowdsourcing, making the method extremely efficient (Paquot et al., 2022). Research has also suggested that novice judges can produce evaluations of acceptable reliability even when they lack extensive training or expertise in evaluating the target construct (Sims et al., 2020).

However, very little is known about what aspects of written competence novice judges pay attention to when evaluating L2 writing in a CJ task: whether for instance, they take different criteria into account compared to expert judges (Schoonen et al., 1997). In this presentation, we explored this issue from two perspectives. Firstly, we collected survey data exploring which aspects of written competence were most salient to non-experts recruited via the crowdsourcing platform Prolific ( $N=390$ ). An exploratory factor analysis of the survey responses revealed that there were three overarching elements that Prolific participants found important in learner writing. These overlapped with traditional constructs in SLA research, namely: complexity, accuracy and coherence. Mean scores for these three factors suggested that participants considered coherence to be the most important aspect of written competence, followed by accuracy, and then complexity.

Subsequently, we invited a subset of the survey participants ( $N=40$ ) to complete a CJ task comparing 50 texts from the ETS corpus (Blanchard et al., 2014), resulting in a rank scale that was highly reliable and significantly correlated with rubric-based scores obtained by expert judges. Then, to test the relationship between the aspects of written language that the judges claimed to be important in the survey against those most predictive of higher rank-order positions, we entered automatically calculated text-based measures of complexity (e.g., lexical diversity and frequency, subordination, noun-phrase elaboration), accuracy (no. grammatical, spelling, punctuation errors) fluency (length) and coherence (e.g., lemma overlap across adjacent sentences) into a linear regression model, with CJ rank order as the dependent variable. This model showed that the rank order of the texts was most strongly predicted by accuracy ( $n$  grammatical and spelling errors) and lexical complexity (lemma frequency in the BNC).

---

\*Speaker

†Corresponding author: nathan.vandeweerd@ru.nl

‡Corresponding author: peter.thwaites@uclouvain.be

§Corresponding author: magali.paquot@uclouvain.be



These results suggest that novice judges recruited from a crowdsourcing platform are indeed sensitive to characteristics of L2 production that are relevant for the evaluation of writing proficiency (e.g., accuracy and lexical complexity). This was evident both in their explicitly reported attitudes towards L2 writing and in the linguistic characteristics that were predictive of the CJ-based rank order, lending credence to the use of CJ and crowdsourcing as a method to reliably and validly elicit L2 proficiency data.

**Keywords:** crowdsourcing, comparative judgement, proficiency, L2 writing, learner corpus

# Exploring the Longitudinal Development of Lexical and Syntactic Complexity in Young L2 English Learners' speaking

De Wilde Vanessa \* <sup>1</sup>

<sup>1</sup> Ghent University – Belgium

An approach in which language development and the learning process have a central place is Complex Dynamic Systems Theory (CDST). In this approach language is seen as a complex system driven by internal and external resources and consisting of interrelated subsystems. Furthermore, development is considered to be highly individual and dependent on initial conditions (De Bot et al., 2007; Larsen-Freeman, 1997).

In this presentation, I will report on a study which explored the longitudinal development of lexical and syntactic complexity in young learners' L2 English speaking development. The study sheds light on how initial proficiency and task type impact the learning trajectory and whether periods of significant growth can be discerned.

The participants in this study were twelve L2 English learners attending the first year of secondary school in Flanders, where formal English lessons only start in the first or the second year of secondary school when learners are 11 to 13 years old but where exposure to English outside the classroom typically starts a lot sooner (De Wilde et al., 2020). The participants were all in the same school and class and received two hours of formal English instruction per week. To map individual learners' development, a 30-week longitudinal study was conducted. Every week, learners did two speaking tasks: a picture narration task and an informative task.

Based on previous CDST-studies (Chan et al., 2015; Yu & Lowie, 2019; Pfenninger, 2020) and the taxonomy of complexity which was proposed by Bulté and Housen (2012), we decided to look into five measures of linguistic complexity. Two syntactic and three lexical complexity measures were selected: length of AS-unit, verb-VAC frequency, word count, MTLD and word frequency, respectively tapping into syntactic complexity, syntactic sophistication, lexical productivity, lexical diversity, and lexical sophistication. To be able to answer the research questions, we used generalized additive models (GAMs). We modelled a GAM for each participant's individual learning trajectory.

Similar to previous studies (e.g. Chan et al., 2015; Yu & Lowie, 2019), this study showed that there were large differences between the learning trajectories of various individuals, even though these learners were quite similar (same age, same L1, same school, same English teacher...). The learners' proficiency at the start of formal L2 English instruction had a large impact on their learning trajectory. The syntactic and lexical complexity in the spoken tasks of the four learners who started the school year with a high initial proficiency hardly changed throughout the year. Other learners, with a lower starting proficiency went through periods in which the complexity measures significantly changed. Overall, more periods of development were observed in the narrative task than in the informative task and the growth in complexity was mainly observed for lexical measures. Further results and implications of the study will be discussed during the presentation.

---

\*Speaker

**Keywords:** lexical and syntactic complexity, speaking development, young learners, longitudinal study, CDST

# The assessment of functional adequacy in oral interactional tasks

Ineke Vedder \* <sup>1</sup>

<sup>1</sup> University of Amsterdam [Amsterdam] = Universiteit van Amsterdam – Netherlands

In the past decade functional adequacy (FA), viewed as successful task completion and adequacy of language in relation to task and context, has emerged as a key construct in SLA research. The importance of assessing FA as an essential component of L2 proficiency, in addition to complexity, accuracy and fluency (CAF), has been emphasized in a number of publications aiming to incorporate FA into the assessment of L2 performance and to establish relationships between the linguistic and the functional dimension of L2 proficiency (Hulstijn et al., 2012; Pallotti, 2009, 2022; Révész et al., 2016).

Within the framework of task-based language teaching (TBLT) and task-based language assessment (TBLA), Kuiken and Vedder (2017, 2018) proposed a rating scale of FA, inspired by Grice's conversational maxims (1975). The scale, developed for oral and written monologic tasks, comprises four dimensions: Task Requirements, Content, Comprehensibility, Coherence & Cohesion.

Several experimental studies have been conducted since, in order to test the applicability of the rating scale, with various learners, proficiency levels, task types and settings (Ekiert et al., 2018, 2022; Kuiken & Vedder, 2022; Strobl & Baaten, 2022). What these studies have shown is that the FA scale appears to be a reliable and valid assessment tool. An important issue for further investigation concerns, however, the potential applicability of the FA scale for non-monologic tasks (González-Lloret, 2022; Youn, 2015, 2018), on which the current presentation focuses.

In our talk we report on an experimental study among L2 learners of Italian and English, set up to explore the use of Kuiken and Vedder's rating scale for oral interactional tasks. To that end, we reformulated the Coherence & Cohesion dimension of the FA scale, in terms of the adequacy of speakers' engagement in situated communicative interaction. The scale descriptors of this dimension 'Interaction', derived from the CEFR (2001), refer to the interlocutors' ability of (i) topic management, (ii) taking/keeping the floor, (iii) mutual co-operation, (iv) asking for/providing clarification.

The following research questions were formulated:

1. How do the judgments of the raters on the dimension 'Interaction' of the FA scale correlate?
2. How are raters' judgments on the four dimensions of FA correlated?

25 dyads of intermediate L2 learners of Italian (different L1's) and 13 dyads of L2 learners of English (L1 Italian) had to perform an oral information-gap task, in which a joint decision had to be made with respect to the choice of a Bed & Breakfast. Data were assessed by two groups of raters (nine for Italian, four for English), after a three-hours training session.

---

\*Speaker

In the talk we will present the outcomes of the study and the implications of the findings for research on CAF and FA. We will also discuss the way in which FA measures can be integrated into the larger field of TBLT and TBLA.

**Keywords:** functional adequacy, TBLT/TBLA, language assessment, rating scale, oral interaction

# ‘Please sport’ or ‘police support’? Examining the effect of phonetic reduction on speech intelligibility for L2 listeners

Gil Verbeke <sup>\*†</sup> <sup>1</sup>, Holger Mitterer <sup>2</sup>, Ellen Simon <sup>\*</sup>

3

<sup>1</sup> Ghent University – Belgium

<sup>2</sup> University of Malta – Malta

<sup>3</sup> Ghent University – Belgium

Across languages, words have a different pronunciation in casual and careful speech, as words produced in casual speech generally contain more instances of phonetic reduction. In English, for example, speakers frequently elide schwas in high-frequency words, such as *memory* (mem’ri) or *support* (s’port). Differences in pronunciation also occur when speakers have a regional or non-native accent (cf. models of L2 speech perception/production). While L1 listeners are highly successful at dealing with variation in pronunciation, L2 listeners may be doubly burdened when phonetically reduced words occur in accented speech. This study aims to investigate how phonetic reduction in different accents of English affects intelligibility for L2 listeners.

120 L1 British English, Belgian Dutch and European Spanish participants completed an orthographic transcription task, in which they transcribed 120 sentences in English. Sentences were produced by speakers with a General British English (*acrolectal*), Newcastle (*regional*) and French (*non-native*) accent. Half of the sentences contained a phonetically unreduced target word (e.g., bakery for *bakery*); the other half a reduced target word (e.g., p’lice for *police*). Each sentence was controlled for speaking rate, lexical frequency and the contextual predictability of the target word. The effect of phonetic reduction on intelligibility was assessed using a generalized linear mixed-effects model, with transcription Accuracy as the outcome variable. Accent (General British, Newcastle, French), L1 (English, Dutch, Spanish) and Reduction (unreduced, reduced), including their interactions, were the fixed effects. Random intercepts were included for Participants and Items.

First, we hypothesized that L1 listeners will outperform L2 listeners, as L1 listening is generally easier than L2 listening. Secondly, we hypothesized that reduced words will be less intelligible to Spanish listeners than to Belgian Dutch listeners due to the absence of central vowels in Spanish, but the presence of such vowels in Dutch. Consequently, reduced words occur more often in Dutch, which may result in an intelligibility benefit for Dutch-speaking listeners. Thirdly, we hypothesized that General British English speakers will be most intelligible, and French speakers least intelligible, depending on listeners’ familiarity with these accents.

The results confirmed that L1 English listeners outperformed both groups of L2 listeners, and that L2 listeners were more strongly affected by phonetic reduction than L1 listeners. Despite

---

\*Speaker

†Corresponding author: gil.verbeke@ugent.be

the overall higher transcription accuracy of the Dutch-speaking participants, no significant difference in the reduction effect could be established between Belgian Dutch and Spanish listeners. The phonetic reduction effect was also similar across accents. Generalizing over reduction conditions, we found that L1 English listeners had a larger intelligibility advantage for the acrolectal and regional English accents than L2 listeners, and both L1 English and Belgian Dutch listeners were noticeably more troubled by French speaker accents than L1 Spanish listeners. That is, the advantage of English- and Dutch-speaking listeners over Spanish-speaking listeners is attenuated when listeners are exposed to French-accented English.

We are currently running a follow-up experiment with 80 Dutch- and Spanish-speaking participants. The results of both experiments will offer a deeper understanding of how phonetic reduction in different accents of English affects intelligibility.

**Keywords:** English, phonetic reduction, language variation, intelligibility, accents, L2 listening

# Learning English in secondary school in Reunion Island: biliteracy and age factor in foreign language learning.

Mallaury Viry \* 1

<sup>1</sup> Université Paul-Valéry Montpellier 3 - Faculté de Langues et cultures étrangères et régionales –  
Université Paul-Valéry - Montpellier 3 – France

In second language acquisition, the correlation between language acquisition and starting age is a subject vastly debated among researchers. There are two competing theories: "the younger is better", with the argument of a 'Critical Period' (Penfield and Roberts, 1959 as mentioned in Ortega, 2009) suggesting that there is an ideal time for language learning and that language acquisition is compromised after this period; versus "the older is better", where researchers propose that late learners of an additional language may catch-up and outperform early starters due to their advanced level of literacy in their first language, allowing a transfer of skills from the first language to the second language (Pfenninger & Singleton, 2017: 2019). Regarding the influence of the home language(s) on the acquisition of foreign language literacy, scholars claim that a positive transfer of knowledge is possible if learners have a sufficient level in their first language(s) (Cummins, 1979).

These two strands of research were taken up by Simone E. Pfenninger in the *Beyond Age Effects* (BAE) project, in which she reported the results of a five-year longitudinal study led in multilingual Switzerland. She examined the English development of a cohort of 636 secondary school students at two periods of time, the first in 2009 when the students were around age 13-14 and the second in 2014 at the age range 18-19. Her goal was to analyse the role of starting age and the influence of the home language(s) in learning a new language for different learner populations: monolinguals, simultaneous bilinguals, and sequential bilinguals. The results suggest that biliteracy and contextual effects have a predominant influence on the acquisition of English foreign language (Pfenninger, 2020).

Our study is an approximate replication of this research. We aim to add to the literature by looking at the effect of the age factor on the acquisition of an additional language, paying particular attention to French-Creole bilingualism in Reunion Island and comparing its specific features with those of other multilingual territories such as Switzerland. To do so, we carried out a research protocol on a sample of 71 participants, comprised of three Year 10 classes, i.e., learners aged between 13 and 15. All the participants could read and write in French. Regarding the English age of onset, we defined two age groups, a starting age of 3 to 6 years for the early starters, and an age group of 8 to 11 years for those who had learned English later. We adapted Pfenninger's research protocol to match the profile of our participants. We mainly focused on written activities and the assessment of written skills (biographical questionnaires, English written production, placement test, productive vocabulary test). The results indicate that age does not have a statistically significant influence on English language skills and that French-Creole bilinguals do not seem to benefit from a positive transfer of their knowledge of their home languages to the target language.

---

\*Speaker



**Keywords:** age factor, bilingualism, biliteracy, English, Creole, second language acquisition.

# Motivation for learning languages in English secondary schools: an exploratory study of beginner learners of Mandarin, French, Spanish and German

Robert Woore \*<sup>1</sup>, Laura Molway \*

<sup>1</sup>, Clare Savory \*

1

<sup>1</sup> Oxford University Department of Education – United Kingdom

Motivation, central to L2 learning, has seen a ‘surge’ in research in recent decades (Boo, Dörnyei & Ryan, 2015). However, studies have focussed predominantly on adult learners of English; school-age learners of other languages are underrepresented. Addressing this gap, we examined the motivation of beginner learners of Mandarin and European languages in five English secondary schools, against a backdrop of increasing uptake of Mandarin. This is an interesting context for L2 motivation research: (1) Language learning in England is widely seen as being in crisis, with persistent problems of low motivation, uptake and attainment; (2) Mandarin is typologically distant from English, acknowledged as difficult to learn and ‘likely to pose particular motivational challenges’ (Chan et al., 2022:16).

We administered a questionnaire – used in Coleman et al.’s (2007) national survey of English secondary schools, originally based on Gardner’s (2001) socio-educational model – to 997 students in their first year of language study. We also conducted 11 focus group interviews (n=43 students), to understand better the complex, dynamic nature of learners’ motivation from a ‘person-in-context’ perspective (Ushioda, 2009).

Quantitative data analysis found that, taking into account the clustering of the data within schools, overall L2 motivation levels did not differ significantly between Mandarin and European languages, and were similar to those reported by Coleman et al. (2007). However, exploratory factor analysis indicated different factor structures: for Mandarin, a three-factor structure, with most variance being explained by ‘integrative/instrumental motivation’; for European languages, a four-factor structure, with a ‘good student’ factor being most important. This may reflect students’ perception of European languages as similar to other school subjects, whereas Mandarin is something new and different.

We analysed the qualitative data by inductive coding, identifying seven over-arching themes. A range of theoretical perspectives from the SLA literature were found to illuminate different aspects of the interview data, without any one framework being adequate in itself. Key differences in motivation for learning Mandarin versus European languages were: a sense of novelty and difference, with the writing of Chinese characters being particularly attractive; a sense of

---

\*Speaker

being 'special' through studying a language which is difficult and unusual to learn; and a greater interest in China's culture and history. Pedagogical and theoretical implications are discussed.

#### References

- Boo, Z., Dörnyei, Z. and Ryan, S., 2015. L2 motivation research 2005–2014: Understanding a publication surge and a changing landscape. *System*, 55:145-157.
- Chan, J., Woore, R., Molway, L. and Mutton, T., 2022. Learning and teaching Chinese as a foreign language: A scoping review. *Review of Education*, 10(3):e3370.
- Coleman, J.A., Galaczi, Á. and Astruc, L., 2007. Motivation of UK school pupils towards foreign languages: a large-scale survey at Key Stage 3. *Language Learning Journal*, 35(2): 245-281.
- Gardner, R.C. (2001) Integrative motivation and second language acquisition. In Z. Dörnyei and R. Schmidt (eds.) *Motivation and Second Language Acquisition* (pp. 1-20).
- Ushioda, E., 2009. A person-in-context relational view of emergent motivation, self and identity, in Z. Dörnyei & E. Ushioda (eds.), *Motivation, language identity and the L2 self*, pp.215-228.

**Keywords:** motivation, Chinese as a Foreign Language (CFL), instructed language learning, school, aged learners

# The role of musical aptitude in the L2 acquisition of tonal words: Evidence from cross-situational statistical learning

Xia Wang <sup>\*† 1</sup>, Yuxin Ge <sup>1</sup>, Patrick Rebuschat <sup>1</sup>, Padraic Monaghan <sup>2</sup>

<sup>1</sup> Department of Linguistics and English Language, Lancaster University – United Kingdom

<sup>2</sup> Department of Psychology, Lancaster University – United Kingdom

Cross-domain transfer (Ong et al., 2016) occurs between music and language when listeners process overlapping acoustic cues between speech and music. One speech feature that shares overlapping acoustic dimensions with music is lexical tone. In tonal languages, lexical tones are realized through pitch variations and are used to distinguish word meanings; hence, they are crucial in tonal word learning. However, lexical tones usually pose challenges for L2 learners with non-tonal native languages in learning tonal words, especially without explicit instructions, due to the absence of tonal cues in their phonological representations. One hypothesis based on cross-domain transfer between music and language is that this tonal difficulty may be modulated by better musical aptitude. In this study, we tested this hypothesis by exploring whether and how music-to-language transfer facilitates L2 tonal word learning. We adopted a cross-situational word learning paradigm (CSWL), where participants learned word meanings through multiple exposures without explicit instructions and mapped sounds to objects based on statistical co-occurrences. Additionally, we investigated whether non-tonal language speakers could learn tonal words without explicit instructions.

English native speakers completed a CSWL task with 6 blocks, retrospective verbal reports, and five musical aptitude tests (Yoko, 2008, 2011). In CSWL, participants were asked to map 12 disyllabic Chinese pseudowords to corresponding referents. There were four types of CSWL trials: p-MP (minimal pairs contrasting on the pitch of tones), d-MP (minimal pairs contrasting on the duration of tone), dp-MP (minimal pairs contrasting on pitch and duration), and v-MP (minimal pairs contrasting on one vowel). These types allowed us to test speakers' ability to perceive phonological details when learning phonological contrasts. The retrospective verbal reports were used to assess participants' awareness of non-native sounds. The musical aptitude tests assessed listeners' sensitivity to timbre, tempo, pitch, loudness, and melody.

Results showed that participants steadily increased accuracy from the first to the final block across all trial types. Except for the first two blocks of the d-MP trials, accuracy across the other three trials was significantly above chance (0.5) for all blocks. Learners showed the highest accuracy in v-MP trials, and d-MP trials caused the lowest accuracy. Additionally, participants who were aware of the pitch/duration cues learned better in CSWL. The correlation test showed that the melodic test score was significantly correlated with the accuracy in the CSWL task in p-MP, dp-MP, and d-MP trials.

Our study contributes to the understanding of cross-domain transfer between music and language in L2 learning. Findings suggest that non-tonal speakers can use fine acoustic cues to learn tonal words by keeping track of cross-situational statistics, and that musical aptitude relates to tonal word learning for non-tonal speakers. Moreover, non-tonal learners could develop different

---

\*Speaker

†Corresponding author: x.wang116@lancaster.ac.uk

levels (full, partial, or minimal) of awareness regarding the pitch and duration of lexical tones, and awareness also predicted non-tonal learners' performance in CSWL. Further investigation on the role of musical aptitude in other aspects of language learning will provide greater insight into integrating the relationship between music and language for L2 learning.

**Keywords:** Musical aptitude, statistical learning, minimal pairs

# L2 grammar-for-interaction in Mandarin: Changes in Students' Use of *na* in Classroom Interaction

Xiaoyun Wang \* <sup>1</sup>

<sup>1</sup> University of Alberta – Canada

Grammar plays a crucial role in participants' interactional practice of accomplishing social actions (Couper-Kuhlen & Selting, 2018). However, the development of L2 grammar-for-interaction among learners has so far remained underexplored (Pekarek Doehler, 2018). Recent emergent studies in this field have investigated the use of L2 speakers' grammatical forms for interaction in English (Eskildsen, 2011; Pekarek Doehler & Balaman, 2021), French (Pekarek Doehler, 2018; Skogmyr Marian, 2021), Korean (Kim, 2009), and Japanese (Ishida, 2009; Masuda, 2011). To enrich our cross-linguistic understanding of L2 grammar-for-interaction, this study analyzes the use of token *na* in Chinese-as-a-Second-Language (CSL) classrooms.

The study uses the methodology of Interactional Linguistics with a cross-sectional design, comparing the students' use of *na* across different proficiency levels. The data are 20 hours of video recordings from 13 university-level CSL classrooms, encompassing various levels and language skills, including 4 advanced classes, 5 intermediate classes, and 4 beginner classes. The focused token, *na* in Mandarin, can function as a demonstrative pronoun, equivalent to 'that' in English (Li & Thompson, 1981), or as a discourse marker roughly glossed as 'so' or 'and then'. As a discourse marker, it can be used to manage topics (Biq, 1990; Li & Ren, 2020) and form a connection with prior speech (Fang, 2000).

Preliminary data analysis reveals a developmental diversification trend in the students' use of *na*. At the beginner level, students primarily use *na* as a demonstrative. In mid-beginner level classes, students begin using *na*-prefaced utterances as responses. At intermediate levels, *na* emerges as a discourse marker to preface declaratives that extend prior topics. In *na*-prefaced utterances, students express their opinions on prior topics (see the excerpt below). Advanced students diversify its function, using *na*-prefaced interrogatives to launch the next question on their interactional agenda. This observation offers evidence of how L2 speakers develop their use of discourse markers to organize their actions, demonstrating a growth in their interactional competence. This progression is marked by their enhanced ability to manage sequences and topics through the strategic use of grammatical constructions.

Excerpt:

01 Tch: you ziji de jiating you ziji de shiye.

'has their own family and their own career.'

02 chengjialiyе.

'to build a family and a career.'

03 Ann: **na** (.) zai zhongguo (-) meiyou chengjia de nanren.

-> '**then, a man in China who has not built a family,**'

04 keneng hui bei juede qiguai.

---

\*Speaker

**‘might be considered strange.’**

This study will also present descriptive statistical data to show how the distribution of *na* with different functions evolves with language proficiency. Its goal is to enrich cross-linguistic discussions on the development of L2 grammar-for-interaction, providing case-based insights into Mandarin, a non-European language that has largely been overlooked in previous studies. Additionally, this research aims to broaden our understanding of CSL teaching and learning in general.

**Keywords:** L2 grammar for interaction, classroom interaction, Mandarin Chinese, discourse markers

# The processing of cognates in idioms – a self-paced reading with lexical decision

Kristina Weissbecker \* <sup>1</sup>

<sup>1</sup> Universität Kassel – Germany

In bilingual, psycholinguistic research, cognates have been a popular and much used source as part of the experimental material (some prominent examples are Costa et al., 2000, or Peeters et al. 2013). A high amount of the conducted experiments, however, studied them in isolation rather than in sentence context, as was criticized, for example, by Li & Gollan (2018). Even fewer of these investigated cognates within fixed phrases such as idioms. Hendriks et al. (2021), are in fact among the first to do so. Research on idioms, on the other hand, has not yet sufficiently decided upon is whether the literal meaning of words within idioms can be accessed next to the idiomatic interpretation (see, e.g., Gibbs, 1993). If we assume this to be possible, then cognates within idioms should elicit the same facilitation effects that have been observed in non-idiomatic sentences. Correspondingly, the two research questions in this study are: a) can speakers access cognates that are embedded in idioms? And if yes, b) can we observe a cognate boost in idiomatic phrases, too, or does this effect arise only in literal, non-idiomatic sentences? In the present study (which is a follow-up to a conducted experiment with fewer participants and slightly different material) cognates and non-cognates are compared in both idiomatic and non-idiomatic sentences. The experimental design includes a self-paced reading part (SPR) followed by a lexical decision (LDT) and a subsequent content question. Thus, cognates and non-cognates are investigated in both a sentential context as well as in isolation. Participants are German-English participants and are being recruited via Prolific.com.

The preliminary study results ( $n = 64$ ) yielded the following: in the LDT, cognates in idioms produced the fastest RTs with 933 ms, and cognates in non-idiomatic sentences elicited 972 ms. Non-cognates were overall slower than cognates with non-cognates in idiomatic phrases producing RTs of 1059 ms and in non-idiomatic constructions around 1049 ms. A Welch t-test showed a significant difference between cognate and non-cognate items ( $p = 0,0019$ ), in other words, we can detect a cognate effect in both the idiomatic and non-idiomatic conditions. Comparing idiomatic with non-idiomatic items, however, no significant difference could be observed ( $p = 0,66$ ). So, whether or not a sentence contains an idiomatic expression generally does not seem to influence the responses in the LDT.

These and further results of the study will be presented with regard to the two main research questions from above. The experiment is furthermore embedded in a PhD project which aims to determine how cognate translations are stored in the bilingual mental lexicon and accessed during processing. Until now, only a small part of the research has dwelled on the question how cognate translations are stored. So, the study is also aimed to shed light on the question if cognate translations are stored in separate lexical entries or together in one.

---

\*Speaker



**Keywords:** cognates / bilingual / mental lexicon / processing / self, paced reading / lexical decision / reaction times

# Investigating CLI in multilingual acquisition through an artificial language

Marit Westergaard \* <sup>1,2</sup>, Chloe Castle \* <sup>† 1</sup>, Natalia Mitrofanova <sup>1</sup>

<sup>1</sup> The Arctic University of Norway [Tromsø, Norway] – Norway

<sup>2</sup> Norges teknisk-naturvitenskapelige universitet - NTNU – Norway

The Linguistic Proximity Model (LPM) argues that learners can draw on resources from both/all their previously acquired languages in developing a grammar of a new language (Ln) (Westergaard et al. 2017, Westergaard 2021). This is in contrast to the Typological Primacy Model (Rothman et al. 2019), which argues that at initial stages of acquisition, there is wholesale transfer of all grammatical representations and routines from the previously acquired language that is most similar to the Ln, according to a hierarchy where lexical similarity is considered first and syntactic similarity last.

Mitrofanova et al. (2023) used a semi-artificial language (lexically similar to Norwegian, but with case-marking on nouns, as in Russian) in a Sentence-Picture Matching Task (SPMT) with Russian-Norwegian, English-Norwegian and Greek-Norwegian participants (Greek has case-making on determiners, English has no case). Results showed that Russian-Norwegian bilinguals scored significantly higher than English-Norwegians and Greek-Norwegians. The authors conclude that syntactic similarity is facilitative at a very early stage, but only when the expression of a property also has superficial similarity to a previously acquired language.

The current SPMT study replicates the previous study with a twist, involving *two* artificial languages, both lexically similar to Norwegian. Language A has case-marking on nouns (similar to Polish) and Language B case-marking on determiners (different from both Norwegian and Polish). A subtractive language-groups design is used, with two participant groups: Polish-Norwegian-English multilinguals and Norwegian-English bilinguals. Participants were exposed to 20 grammatical sentences in the artificial language (Language A or Language B), 10 with SVO word order and 10 with OVS. The SPMT which followed consisted of 60 sentences, 15 SVO correct, 15 SVO incorrect, 15 OVS correct, and 15 OVS incorrect. For Language A we predicted that the Polish-Norwegian speakers would perform better than the Norwegian speakers in the two critical conditions (SVO incorrect and OVS correct), with facilitative CLI from Polish. For Language B, we predicted that, similar to the Greek-Norwegians above, there would be less of a facilitative effect for non-superficial structural similarity.

Results show that Polish-Norwegian multilinguals perform significantly better than Norwegian-English bilinguals in one of the critical conditions (OVS correct) in Language A (case on nouns, abstract + superficial similarity). This effect is not seen in Language B (case on articles, abstract structural similarity only). The difference between Polish-Norwegian multilinguals and Norwegian-English bilinguals in Language A indicates that multilingual learners do not transfer one of their previously acquired languages wholesale, but rather that both/all pre-existing languages influence the acquisition process. We suggest that other factors, including proficiency level, activation, SVO bias, and *yes*-bias, may also play a role and explain why there was a difference between the groups for only one of the critical conditions. The lack of a facilitative effect for the Polish-English-Norwegian group in the critical conditions in Language B provides

---

\*Speaker

<sup>†</sup>Corresponding author: chloe.castle@uit.no

support for Mitrofanova et al (2023)'s findings, suggesting that, at early stages of learning, syntactic similarity is facilitative only when the property is expressed in the same way in the previously acquired language.

**Keywords:** Cross, Linguistic Influence, language acquisition, artificial languages, Linguistic Proximity Model

# Pronoun interpretation in English: When native speakers 'misbehave'

Lydia White \*<sup>1</sup>, Heather Goad<sup>2</sup>, Guilherme Garcia<sup>3</sup>, Natália Guzzo<sup>3</sup>,  
Liz Smeets \*  
<sup>4</sup>, Jiajia Su<sup>5</sup>

<sup>1</sup> McGill University – Canada

<sup>2</sup> McGill University – Canada

<sup>3</sup> Université Laval – Canada

<sup>4</sup> York University – Canada

<sup>5</sup> Beijing Foreign Studies University – China

Pronoun interpretation in languages like English is open to ambiguity: the antecedent for an overt pronoun can be either subject or object. Nevertheless, native speakers are reported to prefer an interpretation where the antecedent is the subject (e.g., Contemori & Dussias 2020; Cunnings et al. 2017); see (1) (preferred interpretations underlined).

(1) Monica phoned Claudia when she was in the office.

(2) Monica phoned Claudia when SHE was in the office.

However, prosody can impact pronoun interpretation, leading to a shift in preferences. Unstressed pronouns in English prefer subject antecedents, while stressed pronouns prefer non-subject antecedents (Grimshaw & Rosen 1990); compare (1) and (2). Similarly, Gargiulo et al. (2019) report, for Swedish, that a shift in antecedent preferences is signalled in production by stress on a pronoun or a pause between clauses.

The role of prosody in L2ers' interpretations of English pronouns has not previously been considered. We examine antecedent choices of Italian-speaking L2ers and English native speakers. We hypothesize that both groups will show a preference for subject antecedents for unstressed pronouns and that there will be a shift away from this preference in cases involving stressed pronouns or a pause between clauses.

We report on an experiment, administered online, involving 21 Italian speakers (intermediate/advanced English proficiency) and 21 native speakers of English. Stimuli were 24 biclausal sentences (similar to (1) and (2)), presented auditorily. Stimuli manipulated presence/absence of pronoun stress and presence/absence of pause between clauses. Each sentence was preceded by a written context introducing potential referents. After listening to a sentence, participants indicated their preferred referent for the pronoun.

The L2ers showed sensitivity to prosody, as predicted: subjects were the preferred antecedents for unstressed pronouns while stress led to a significant decrease in subject antecedent choices, regardless of the presence of a pause ( $b^{\wedge} = -0.99$ , 95% HDI =  $(-1.89, -0.14)$  (no pause);  $b^{\wedge} = -1.60$ , 95% HDI =  $(-2.58, -0.69)$  (pause)). The results from the native speakers, on the other

---

\*Speaker

hand, were unexpected. Although the data show an overall preference for subject antecedents over objects ( $b^{\wedge}=0.56$ , 95% HDI = (-0.08, 1.18)), confirming earlier findings, neither stress nor pause led to a shift in antecedent choices.

The L2ers' preferences may be indirectly attributable to the L1: the two different pronoun types in Italian (null versus overt) differ in their antecedent preferences, which may make L2ers extra-sensitive to differences in the L2 (unstressed versus stressed). Native speakers may need richer contextual cues for prosody to have an effect.

## References

Contemori, C. & P. Dussias. 2020. The processing of subject pronouns in highly proficient L2 speakers of English. *Glossa* 5(1): 38.

Cunnings, I. et al. 2017. Anaphora resolution and reanalysis during L2 sentence processing. *SSLA* 39: 621-652.

Gargiulo, C. et al. 2019. The role of prosody in overt pronoun resolution in a null subject language and in a non-null subject language: A production study. *Glossa* 4(1): 135.

Grimshaw, J. & S. Rosen. 1990. Knowledge and obedience: the developmental status of the Binding Theory. *LI* 21: 187-222.

**Keywords:** prosody, pronoun interpretation, L2 English

# The Effectiveness of Cognitive Linguistics (CL)-inspired Instructions on L2 Vocabulary Learning: A systematic literature review

Injung Wi \* <sup>1</sup>, Frank Boers \* <sup>† 1</sup>

<sup>1</sup> Western University – Canada

Boers (2013) provided an inventory of 15 quasi-experimental intervention studies published between 1996 and 2010 that focused on the impact of cognitive linguistic-inspired teaching techniques on the retention of L2 vocabulary. He found that most of the included studies reported positive effects of cognitive linguistic-based instruction on L2 vocabulary retention. However, there has not been a comprehensive research synthesis on this topic since Boers (2013). It seems timely and appropriate to conduct another research synthesis on the effectiveness of cognitive linguistic-inspired instruction on L2 vocabulary acquisition to further ascertain its effects and to gain insight into how methodologies in this line of research have evolved over a decade. Cooper and Hedges (2009) suggest that a research synthesis is a way of scrutinizing existing studies and related theories, combining findings, and mapping the gaps for future research. Building on Boers (2013), therefore, a systematic review targeting studies published after 2010 was conducted as a means of synthesizing research that investigated the effectiveness of cognitive linguistic inspired approaches to L2 vocabulary teaching. Through a keyword search using databases such as ProQuest, ERIC, and Google Scholar, a search of journals such as *TESOL Quarterly* and *Language Learning*, and a citation search, I imported a total of 21 studies to Covidence. 18 studies were eligible for the full-text screening after 3 duplicates were removed. Based on a set of inclusion criteria informed by Boers (2013), a total of 9 studies were included in the final pool of the systematic literature review. This review provides an inventory of the 9 studies, outlining main research features such as target items, interventions, participants, testing methods, and results. This inventory may be a useful resource for L2 teachers and researchers who have an interest in this topic. Analysis of these studies revealed that 8 of the 9 studies were favourable to cognitive linguistic informed instruction compared to traditional approaches of vocabulary acquisition. It was also found that the comparison conditions to cognitive linguistic approaches in research remained similar to those of earlier studies. Since 1996, the comparison condition has been described as a traditional approach to instruction and teaching. Traditional approaches include providing a list of phrasal verbs with L2 definitions, metalinguistic rules and example sentences, simple correct feedback, and form-focused instructions. A change in research trends was found regarding the selection of target items. Phrasal verbs and prepositions appear to be the most frequently examined target items, whereas metaphoric words and figurative idioms were the most investigated constructs before 2010. The pedagogical implications of the findings will be discussed in detail.

**Keywords:** cognitive linguistics, cognitive inspired instruction for L2 vocabulary learning, L2 vo-

---

\*Speaker

<sup>†</sup>Corresponding author: fboers@uwo.ca

cabulary learning, systematic literature review

# The variable production of (non-)canonical structures in spoken vs. written language: German V2 vs. V3 declarative constructions

Johanna Wittner \* <sup>1</sup>, Andrea Ender \* † <sup>1</sup>

<sup>1</sup> Paris-Lodron-Universität Salzburg – Austria

Canonical German declarative sentences have a V2 position. This basically results in two possible positions for the subject: (1) preceding the finite verb in the pre-field (SVX) or (2) following the finite verb in the middle-field, which is accompanied by a subject-verb-inversion (XVS). Producing such XVS structures and renouncing from non-target like V3 structures is traditionally described as the goal of a lengthy acquisition process (e.g. Clahsen et al. 1983, Wisniewski 2020). However, in L1 language use - predominantly in the spoken language or in the informal register - V3 positions can also be found in certain declarative structures in addition to the aforementioned V2 options (e.g. Wiese et al. 2022, Bunk 2020), namely when the left outer field is filled by metacommunicative elements (DUDEN 2022). These elements known as operators or discourse markers, which precede the actual pre-field, have a projecting character, the arrangement of which has a two-part construction and provides recipient-friendly anticipatory clarification of information. L2 learners are therefore faced with the challenge of breaking up the acquired V2 sequence and also allowing V3 in specific contexts.

This article analyzes the occupation of the left outer field and the pre-field in spoken and written language use of German L2 university applicants (n = 41). The corpus used is based on language data from argumentative speaking and writing tasks. The language data show that these advanced L2 users have recognized and use the function of metacommunicative elements in the left outer field, but also use them in wider contexts compared to a smaller group of L1 first-year students (n = 8). Furthermore, the results of a grammaticality judgment task on V2 and V3 structures show that both natives and non-natives display a rather high variability in rating V3, with L2 students showing an even broader acceptability of V3. Overall, this underlines the fact that, especially in advanced German L2 acquisition, a didactic reduction to V2-only declaratives must be overcome. Variable structures and differences in oral and literal use should be explicitly addressed (Knouse & Abreu 2022), so that advanced L2 learners can acquire an appropriate degree of variation in terms of register and modality.

## References

Bunk, O. (2020): 'Aber Immer Alle Sagen Das'. The Status of V3 in German: Use, Processing and Syntactic Representation. Ph.D. Dissertation, Humboldt University Berlin.

Clahsen, H. et al. (1983): *Deutsch als Zweitsprache. Der Spracherwerb ausländischer Arbeiter*. Tübingen: Narr.

Knouse, S. M. & Abreu, L. H. (2022): Teaching Variable Structures. In: Geeslin, Kimberly L. (ed.): *The Routledge Handbook of Second Language Acquisition and Sociolinguistics*. Rout-

---

\*Speaker

†Corresponding author: andrea.ender@plus.ac.at



ledge, 72–83.

Wiese, H. et al. (2022): Heritage Speakers as Part of the Native Language Continuum. *Frontiers in Psychology* 12, 1–19.

Wisniewski, K. (2020): SLA developmental stages in the CEFR-related learner corpus MERLIN. Inversion and verb-end structures in German A2 and B1 learner texts. *International Journal of Learner Corpus Research* 6:1, 1–37.

Wöllstein, A. & Duden editorial team (Eds.) (2022): *Duden. Die Grammatik. Unentbehrlich für richtiges Deutsch*. Dudenverlag (Duden 4), 9th, completely revised and updated edition.

**Keywords:** spoken, written, variation, prefield, leftouterfield, German, register, modality, V2, V3, native, nonnative, languageuse

# Empowering multiliteracy: A case study on the use of ICT in Turkish heritage language writing in bilingual seventh graders

Till Woerfel \* <sup>1</sup>

<sup>1</sup> University of Cologne – Germany

Turkish is a widely spoken minority language in Germany. Despite the institutionalization of Turkish heritage language teaching (THLT) almost forty years ago, students' exposure to the formal register of Turkish remains limited, mainly due to monolingual norms (Gogolin, 1997) and significant challenges that THLT faces in terms of educational policy and contemporary instructional development (Küppers & Schroeder, 2017).

Consequently, while many Turkish-German bilingual students exhibit oral bilingualism, they often show monoliteracy and uncertainty in writing Turkish (e.g. Dirim, 2009). Students' hand-written Turkish texts e.g. reflect an interpretation of orthographic words influenced by German orthographic principles (Schroeder & Şimşek, 2010). Today, however, digital technologies offer various possibilities to support and develop writing processes in formal registers. However, research shows that the use of word processors only has an impact on text quality when learners are provided with additional writing support (Graham & Harris, 2018). Furthermore, text-to-speech provides the potential for revising texts, notably for L2 learners (Dahlström & Boström, 2017).

This paper presents results from a case study of three Turkish-German bilingual seventh graders who received literacy and ICT trainings as part of an intervention study (Woerfel & Michels, 2022). The intervention study aimed to find out how seventh graders' texts develop in a learning environment that includes literacy and ICT trainings. For this purpose, a control group (CG N=42) and two experimental groups (EG1 N=43; EG2 N=42) received a reading and writing strategy training. EG1 and EG2 received additional training in the use of a word processor, spell checker, and thesaurus, and EG2 received a further training in the use of text-to-speech. To assess the ability to transfer the learned strategies to the L1 Turkish, three students of EG2 completed an additional digital writing task in Turkish. The texts were annotated at various levels (orthographic, morphological, syntactic and lexical) to determine formal correctness, errors and norm deviations. The text quality was assessed by external raters. Additionally, the use of digital tools in the writing process was annotated to identify positive revisions at formal and textual levels.

Results indicate positive transfer of specific writing strategies to Turkish, such as completeness of information about the writing task, common thread, and addressing the audience. Although the students did not attend the THLT, their Turkish texts approached the quality of their digital texts in German. Furthermore, the data show that two-thirds of the formal errors in the students' drafts were reduced by the use of spell checkers and text-to-speech. Interestingly, the digital texts show specific characteristics such as in-sentence capitals, separate spelling of suffixes or alternative subordination strategies, which have been classified in previous research as an emerging new variety of "Germany Turkish" and seem to be robust to digital aids.

---

\*Speaker

The paper argues that reading and writing strategy training is an effective way to promote (ICT) literacy in L1 and L2 and serves as a gateway for multilingual students to develop their writing skills in multiple languages, fostering a sense of inclusivity and empowerment in their educational journey.

**Keywords:** digital writing, multiliteracy, ICT, literacy, digital technologies, Turkish, German, secondary students, case study, intervention study

# Un fichu caractère or un caractère fichu ? The acquisition of L2 French adjective placement by Chinese speakers with a primary language of L1

Zhanglin Xie \*<sup>1</sup>, Barbara Hemforth<sup>† 1</sup>, Maria Copot<sup>2</sup>, Claire Saillard<sup>1</sup>

<sup>1</sup> Laboratoire de linguistique formelle (LLF) – Laboratoire de Linguistique Formelle, CNRS – France

<sup>2</sup> The Ohio State University Press – United States

While Chinese requires prenominal position of attributive adjectives, French exhibits pre- and postnominal orders (Abeillé & Godard, 1999; Alexiadou, 2001; Bouchard, 1998; Thuilier, 2012, 2014), with a predominance of postnominal order in French. The acquisition of adjective placement in L2 French poses challenges for learners, irrespective of their L1 (Baihaqi, 2020; Granfeldt, 2004; Noordhuis, 2015). This study investigates the role of L1 transfer (Gass, 1988; Ionin & Montrul, 2010; Karim & Nassaji, 2013; Odlin, 1989), overgeneralization of target language rules (Al-Khresheh, 2016; Corder, 1981; Ellis & Ellis, 1994; Littlewood & William, 1984; Richards, 1974; Taylor, 1975), adjective frequency, L2 proficiency and L2 immersion for proficient learners in the acquisition of adjective ordering in L2 French. Notably, frequent adjectives prove more resistant to overgeneralization errors than infrequent ones (Ambridge, Kidd, Rowland, & Theakston, 2015). Our research is pioneering in its examination of adjective placement acquisition among L1 Chinese learners of French and represents the initial exploration of the overgeneralization, frequency hypotheses and immersion hypotheses in this context. If L1 transfer predicts the adjective placement choices of Chinese speakers learning L2 French, errors are anticipated, particularly with low-frequency postnominal adjectives, given that Chinese places adjectives before the noun. Conversely, if overgeneralization of the more general post-nominal position influences L2 production, L1 Chinese speakers will produce errors in particular for infrequent prenominal adjectives, such as *\*un caractère fichu*, *\*un menteur fieffé*, *\*un danseur piètre* etc. To investigate this, we conducted a forced-choice experiment involving three groups of Mandarin Chinese speakers learning L2 French (beginner (N = 31), intermediate (N = 29) and advanced (N = 29) levels) and included a control group of native French speakers (N = 89). We examined L2 adjective acquisition based on adjective type and frequency, aiming to test the L2 overgeneralization hypothesis against the L1 transfer hypothesis. The findings revealed a proficient acquisition of both high-frequency postnominal adjectives (*vert*, *rond* etc.) and low-frequency postnominal adjectives (*bis*, *mauve* etc.). Prenominal adjectives appearing often in textbooks (mostly high-frequency adjectives) like *beau* and *joli* exhibit a moderate level of mastery. By contrast, prenominal adjectives not explicitly covered in textbooks (mostly low-frequency adjectives like *fichu*, *piètre*, *fieffé*, *soi-disant* etc...) proved to be the most challenging for all three groups of L2 learners. We propose that this difficulty arises from overgeneralization of the default postnominal placement rule to these adjectives. The intermediate group significantly outperformed the beginner group but underperformed compared to the advanced group. Furthermore, among advanced L2 learners, we also compared the role of immersion where the outcomes of the immersion group (N = 18) did not show a significant improvement compared

---

\*Speaker

<sup>†</sup>Corresponding author: barbara.hemforth@linguist.univ-paris-diderot.fr

to the group living in China (N = 11). Our findings suggest that while L1 influence and L2 immersion are not the primary factors in explaining Second Language Acquisition of adjective placement in French, a general tendency toward overgeneralization, proficiency in French, and lexical frequency play crucial roles.

**Keywords:** L2 French acquisition, adjective placement, overgeneralization, frequency, proficiency

# Modulating motion event categorization through brief training: Meaning-focused versus form-focused instructional conditions

Yuyan Xue \*<sup>1,2</sup>, John Williams<sup>3</sup>

<sup>1</sup> University of Cambridge [UK] – United Kingdom

<sup>2</sup> University of Oxford – United Kingdom

<sup>3</sup> University of Cambridge – United Kingdom

## Modulating motion event categorization through brief training: Meaning-focused versus form-focused instructional conditions

Cognitive and psycholinguistic research has increasingly focused on whether language learning can modify cognition beyond language itself, that is, during tasks without overt language use. A growing body of evidence suggests that learning an additional language that encodes a familiar concept differently from our L1 and/or further highlights a familiar concept compared to our L1 can indeed fine-tune our cognition beyond language per se in a variety of domains including time, motion event, and number/amount of objects/substance, etc. (e.g., Athanasopoulos & Bylund, 2023). Moreover, such cognitive fine-tuning appears to be modulated by such factors as L2 proficiency, age of L2 acquisition, length of immersion in an L2 setting, etc. (Bylund & Athanasopoulos, 2014).

Here we explored the role played by language instructional conditions, a factor attracting decades of research in the SLA literature, yet which has not been systematically investigated in the context of language-cognition interactions. More specifically, we trained Mandarin L1 – English L2 participants on a linguistic system modelled on Karitiâna (an Amazonian language) involving two novel grammatical morphemes to highlight the familiar concept of (in)transitivity (*gi* being used obligatorily in intransitive sentences, *ro* in transitive sentences) under three instructional conditions: one meaning-focused (memorizing meanings of 128 exemplar sentences) and two form-focused (rule search on 128 exemplar sentences vs. direct metalinguistic explanation). We asked if participants who have learnt this target system would show higher likelihood of categorizing motion events on the basis of motion transitivity in a triads-matching task without overt language use (compared to other dimensions including motion direction and shape, colour, and size of the entit(ies) involved in the motion event), and if yes, whether this cognitive fine-tuning effect would vary with instructional conditions. By probing this research question, we also aimed to extend the focus of comparisons between form- and meaning-focused instructional conditions from language attainment to cognitive fine-tuning, a learning product probably as important as other products (e.g., development in grammar and vocabulary) yet which constitutes a great challenge to L2 learners, especially those in an instructed setting (Athanasopoulos, Burnand et al., 2015).

Results showed that participants who learnt the target system under the meaning-focused instructional condition showed a higher likelihood of categorizing based on motion transitivity immediately after training than a control group (Experiment 1); those who learnt under rule search instructional condition showed this effect only after additional grammaticality judgment

---

\*Speaker

practice (Experiment 2); whilst those who learnt through direct metalinguistic explanation did not show this effect even after such practice (Experiment 3). These differences were obtained despite the fact that all three groups achieved near-perfect performance on the grammaticality judgement task.

The findings are discussed in terms of depth of processing in instructed SLA (Leow, 2015) and interactive activation models of language-cognition interactions. This study is amongst the first to show that meaning-focused instruction can be superior to form-focused instruction, especially direct metalinguistic explanation, when measures that are sensitive to integration into the linguistic system are employed.

**Keywords:** instructional conditions, meaning, focused, form, focused, language, cognition interactions

# The processing advantage of multiword sequences: A meta-analysis

Wei Yi \*<sup>1</sup>, Yanlu Zhong \*

2

<sup>1</sup> Peking University – China

<sup>2</sup> University of California, Santa Barbara – United States

This meta-analysis synthesized 35 English studies (130 effect sizes,  $N = 1,981$ ) that employed online tasks to investigate the processing of multiword sequences (MWSs). We examined: 1) to what extent MWSs enjoy a processing advantage over novel word combinations, 2) how such a processing advantage is moderated by statistical regularities (i.e., phrasal frequency, association strength), MWS type, and explicitness of experimental tasks, and 3) whether such moderating patterns differ between L1 speakers and L2 speakers. Results confirmed the processing advantage for most subtypes of MWSs, with effect sizes ranging from small to medium. For L1 speakers and L2 speakers, the processing advantage of MWSs was found across the continuum of phrasal frequency and association strength and varied. Interestingly, task explicitness moderated the processing advantage of MWSs, but only for L2 speakers. Taken together, our results shed light on the understanding of MWSs as well as directions for future research.

**Keywords:** multiword sequences, online processing, advantage, meta, analysis

---

\*Speaker



# The effect of using gamified quizzes as a formative assessment tool on students' long-term motivation in an EFL reading course in a Chinese university

Rui Yan <sup>\*† 1</sup>, Reka Jablonkai <sup>\* ‡ 1</sup>

<sup>1</sup> University of Bath [Bath] – United Kingdom

Over recent years research has shown that gamified quizzes (e.g.: Kahoot, Quizizz) can improve student motivation in foreign language classes (e.g.: Eltahir et al., 2021; Tao & Zou, 2021) due to the use of game elements such as challenges, leaderboards, scores. However, most foreign language studies focus on the short-term effect of gamified quizzes on student motivation, normally less than 4 weeks (Boudadi & Gutiérrez-Colón, 2020), while the long-term effect has not been extensively researched (Boudadi & Gutiérrez-Colón, 2020). It remains unknown whether student motivation would continue increasing when gamified quizzes are continuously employed throughout a semester (Zarzycka-Piskorz, 2016). At the same time, relevant studies (e.g.: Kocak, 2020; Rodrigues et al., 2022) in other subject areas have shown that students' motivation may initially increase due to the novelty effect by gamification at the beginning and then decrease as students get familiar with the technology in the long run. Therefore, it is important to explore how students' long-term motivation in foreign language classes changes throughout one semester, when gamified quizzes are regularly used. The present study examined the effect of using gamified quizzes as a formative assessment tool on students' long-term motivation in an EFL reading course in a Chinese university over one semester (16 weeks). The teacher used Quizizz, a gamified quiz platform, for formative assessment purposes in each session in this EFL reading course. All 30 students in this EFL reading course with an intermediate level of English took part in this study. Quantitative data has been collected from a validated motivation questionnaire administered to the 30 students at 6 time points (Week1, Week4, Week7, Week10, Week13, Week16). Qualitative data has been collected from interviews with a sample of 11 students in 5 time points (Week4, Week7, Week10, Week13, Week16). The preliminary quantitative and qualitative analysis has shown that as students gradually get familiar with the use of Quizizz, their motivation in this EFL reading course has increased for the first 7 weeks. Findings regarding the changes of the over-time motivation throughout the whole semester will be discussed. The results will have implications for teachers who wish to use gamified quizzes to improve student motivation in EFL classes in the long run.

**Keywords:** Computer, Assisted Language Learning, EFL Motivation, Experimental study

---

\*Speaker

†Corresponding author: ry440@bath.ac.uk

‡Corresponding author: rirj20@bath.ac.uk

# The role of domain-general auditory processing in L2 speech

Silvia Yang <sup>\*</sup> <sup>1</sup>, Clara Martin <sup>1</sup>, Antje Stoehr <sup>1</sup>

<sup>1</sup> Basque Center on Cognition Brain and Language [Gipuzkoa, Espagne] – Spain

While many post-pubertal learners maintain a strong foreign accent after years of foreign language (L2) exposure, a few acquire a native-like accent in a short period of time. This difference may be attributed to two sets of domain-general abilities, auditory processing and audio-motor integration. Auditory processing is the ability to detect individual dimensions of acoustics (e.g., pitch, formant, duration); audio-motor integration is the ability to integrate acoustics into motor-function (e.g., the repetition of a melody). According to the Auditory Precision Hypothesis-L2, people with acute auditory abilities (in other words, a good ear) would successfully perceive and produce L2, while those with poor auditory abilities would fail to acquire speech proficiency even after years of immersion. Of late, the role of auditory ability in L2 speech perception has been widely explored, but less is known about its role in L2 speech production.

Although a few studies have found a tentative association between auditory abilities and L2 production, the present study aims to cover two gaps. First of all, the reliability of L2 production proficiency judgement measures recruited by prior studies remains questionable. Past research has evaluated the production performance of learners with human raters (e.g., Saito et al., 2020), which may be problematic for several reasons, including the subjectivity of judgement and the possible low auditory abilities of the raters themselves. To address this issue, the present study aims to eliminate probable bias by opting for acoustic measures of speech. Secondly, although past studies have found a correlation between auditory processing and L2 speech proficiency among advanced L2 speakers (e.g., Saito et al., 2020; Sun et al., 2021), they have not tested the hypothesis on speakers at the beginning stage of L2 proficiency. Therefore, the current study will recruit L2 beginners by using artificial words as stimuli.

We will evaluate L1 Spanish native speakers on the acquisition of lexical tones (the rising and falling tones used in Mandarin Chinese) and launch a 4-session study that includes the training and testing of 50 participants. Our prediction is that, our recruitment of acoustic analysis will offer a more fine-grained inspection of the data and reveal a stronger association between the auditory abilities and L2. Additionally, in the beginning stage of L2 acquisition, we predict more reliance on audio-motor integration than auditory processing, since the former skillset involves the mechanical memorization and repetition of speech while the latter requires a prolonged period of input before coming into effect.

As we are in the process of data collection, we expect to have the preliminary results before the time of the conference.

## References

Saito, K., Kachlicka, M., Sun, H., & Tierney, A. (2020). Domain-general auditory processing

---

\*Speaker

as an anchor of post-pubertal L2 pronunciation learning: Behavioural and neurophysiological investigations of perceptual acuity, age, experience, development, and attainment. *Journal of Memory and Language*, 115, 104168.

Sun, H., Saito, K., & Tierney, A. (2021). A longitudinal investigation of explicit and implicit auditory processing in L2 segmental and suprasegmental acquisition. *Studies in Second Language Acquisition*, 43(3), 551-573.

**Keywords:** auditory processing, audio, motor integration, production, auditory precision hypothesis, L2, SLA, lexical tones

# A Systematic Review and Meta-analysis of Storytelling Interventions in Second Language Learning

Ting Yao <sup>\*†</sup> <sup>1</sup>, Xiaotong Xi <sup>1</sup>, Pilar Prieto <sup>1,2</sup>

<sup>1</sup> Universitat Pompeu Fabra [Barcelona] – Spain

<sup>2</sup> Institució Catalana de Recerca i Estudis Avançats – Spain

The potential beneficial effects of storytelling, including both traditional (i.e., verbally telling or retelling a story) and digital formats (i.e., telling stories using multimedia resources) have gained increasing attention in second language (L2) learning. However, previous review studies of L2 storytelling intervention research have either taken on a descriptive approach, by documenting the characteristics of the studies and the research trends (e.g., Lucarevski, 2016) or focused exclusively on systematically examining digital storytelling programs (e.g., Lim et al., 2022). A meta-analytic evaluation of the effectiveness of both traditional and digital storytelling is still lacking. Therefore, the present study aims to address the characteristics of L2 storytelling research, including both formats, meta-analyze its effectiveness on language outcomes, and examine the potential factors that might mediate its efficacy.

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, a systematic search in three databases with structural inclusion procedures yielded 35 peer-reviewed publications with pretest-intervention-posttest designs. A qualitative synthesis of these studies revealed that the interventions were mostly conducted in university classroom settings and focused on improving L2 skills of adult learners. When contrasting the two storytelling formats, traditional storytelling was more frequently used and its duration ( $M = 12.11$ ,  $SD = 17.34$ ) was significantly shorter than digital storytelling ( $M = 21.60$ ,  $SD = 8.80$ ).

Based on the effect sizes and variances provided by 19 studies, a meta-analysis was performed in *metafor* R package (Viechtbauer, 2010) to compute the weighted mean effect sizes. The results indicated an overall positive effect of the interventions, with a significantly large effect on improving L2 speaking skills ( $d = 1.47$ ,  $CI (0.96, 1.98)$ ) and a medium effect on learning L2 vocabulary ( $d = 1.00$ ,  $CI (0.19, 1.82)$ ). High heterogeneity detected across studies warrants further moderator analyses. Learners' proficiency emerged as a significant moderator. Specifically, the interventions proved to be significantly more beneficial for intermediate learners ( $d = 3.36$ ,  $CI (2.67, 4.04)$ ) compared to beginners ( $d = 0.95$ ,  $CI (0.38, 1.51)$ ) and pre-intermediate learners ( $d = 0.63$ ,  $CI (0.10, 1.16)$ ). Thus, interventionists should consider learners' L2 proficiency to maximize effectiveness. By contrast, the effect was not significantly mediated by storytelling format and learners' education level, suggesting that the choice between traditional and digital storytelling, as well as the implementation across various education levels may not be critical.

All in all, the small number of studies and reduced sample sizes call for caution in the interpretation of the meta-analytic results. Further research is needed involving larger sample sizes and diverse populations, and systematically addressing the value of specific techniques and design features within narrative programs.

---

\*Speaker

†Corresponding author: ting.yao@upf.edu

## References

- Lim, N. Z. L., Zakaria, A., & Aryadoust, V. (2022). A systematic review of digital storytelling in language learning in adolescents and adults. *Education and Information Technologies*, 27(5), 6125–6155.
- Lucarevski, C. R. (2016). The role of storytelling on language learning: A literature review. *Working Papers of the Linguistics Circle*, 26(1), 24–44.
- Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36(3), 1–48.

**Keywords:** Oral narrative interventions, storytelling, systematic review, meta, analysis

# The Identity of Chinese People in a Multilingual Era: An Ethnography in a Chinese Primary School

Lijia Yu \* <sup>1</sup>

<sup>1</sup> King's College London – United Kingdom

The intersection of China's massive internal migration and language policy and planning has largely reshaped the linguistic environment of dialect-speaking regions and hence the language use and identity of people living within (Dong, 2011). While existing studies on language and identity in China have predominantly focused on mega-cities like Shanghai and Guangzhou, characterised by high levels of immigration and favourable regional language policies (Liang, 2015), the current study intends to address this research gap by looking at Ningbo, a 'tier-two' city, where migrants are more diverse and regional language policies are more in line with national language policies driven by a monoglot ideology. Setting the research field in a primary school class, the study will be an ethnography aiming to unravel the intricate linguistic identities of young Ningbonese people in the current multilingual era. While it takes a post-structuralist stance towards identity construction by seeing identity as fluid and emergent in interactions, it also draws on the notion of habitus (Bourdieu, 1977) to make the research more doable by locating identity in social structures and practices, which, in this case, are classroom discourses and students' language practices. In this way, students' identities can be viewed as the manifestations of their habitus, conditioned by classroom discourses and embodied in their language practices through their employment of particular linguistic and other semiotic resources. The research will adopt a tripartite identity model. The macro-level examines widely circulating social discourses regarding language use; the meso-level looks at the classroom discourse and explores how it draws on wider social discourses to modify students' language habitus first acquired from family socialisation; and the micro-level investigates how students employ linguistic and other semiotic resources to embody their habitus and hence identities in interactions. Using ethnographic research methods like participant observation and interviews, this study will investigate the complex interplay of the three levels that contribute to the formation of linguistic identities of Ningbonese young people in current multilingual China. The findings of the study will contribute significantly to our understanding of the language and identity issues in contemporary multilingual China, particularly in ordinary industrial cities like Ningbo, where the language environment is dramatically influenced by massive internal migration and national language policies. The findings could also offer insights into developing inclusive language policies, shaping school curricula, and enhancing teacher education programmes in these cities.

References

Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge University Press.

Dong, J. (2018). Language and Identity Construction of China's Rural-Urban Migrant Children: An Ethnographic Study in an Urban Public School. *Journal of Language, Identity & Education*, 17(5), 336–349.

---

\*Speaker

Liang, S. (2015). *Language Attitudes and Identities in Multilingual China: A Linguistic Ethnography*. Springer.

**Keywords:** identity, language policy, internal migration, ethnography, China

# A Call for Cautious Interpretation of Vocabulary Levels Tests: A focus on words with multiple meanings

Xi Yu <sup>\*† 1</sup>, Frank Boers<sup>‡ 1</sup>

<sup>1</sup> University of Western Ontario – Canada

Vocabulary Levels' Tests (VLT) were designed to gauge students' vocabulary knowledge at different frequency levels. VLT scores are often used to estimate how many words students are likely to understand in a given text, which is helpful to evaluate the appropriateness of the text. However, a potential concern with this use of VLT scores arises from the fact that the tests do not capture word knowledge beyond the target items' most common meaning, even though most words in these tests have two or more meanings. This would not be problematic if these other meanings were very rare or if knowledge of the primary meanings of words typically entailed comprehension of the secondary meanings. This study, therefore, investigated: (1) the likelihood of encountering words in their secondary meanings; and (2) the extent to which learners are familiar with the secondary meanings of the target words in a VLT.

Our first step was to determine the number of meanings of the lemmas in the two versions of the Updated Vocabulary Test (UVLT) (Webb et al., 2017), according to the Oxford Dictionary of English (*ODE*). The results revealed that 56.7% (85 words) of Version A and 58.7% (88 words) of Version B items are polysemous, with an average of 2.17 and 2.20 meanings per lemma, respectively. We then examined 100 random concordance lines for these lemmas from COCA and BNC and determined their *ODE* meanings. Interestingly, the "secondary" meaning of 6 lemmas in Version A and of 10 lemmas in Version B appeared to be more frequent in these corpora than the meaning whose knowledge is assessed by the UVLT. The corpus frequency of the secondary meanings of many other lemmas was found to be non-negligible too.

To explore the question of whether knowledge of the primary meaning of the lemmas can make one feel confident that secondary meanings will also be understood, 49 L1 Mandarin students were recruited for a counterbalanced within-participants study, which included two sessions. Their knowledge of the primary meanings was gauged by the original version UVLT. To assess their knowledge of the secondary meanings, we replaced the primary meanings of the words with a secondary meaning, provided the latter accounted for  $\geq 5\%$  of the instances in the corpora. The preliminary results revealed that learners had difficulty in identifying the words' secondary meaning although they knew the words' primary meaning. In the talk, we will present the results of the data analysis in more detail and discuss the implications of our findings for interpreting VLT test scores as a means of estimating learners' "lexical coverage" of L2 materials.

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics*, 168(1), 34-70.

---

\*Speaker

†Corresponding author: xyu369@uwo.ca

‡Corresponding author: fboers@uwo.ca



**Keywords:** Word with multiple meanings, Vocabulary Levels Tests

# Disruption of native language access by non-verbal emotional content in bilinguals

Wanyu Zhang <sup>\*</sup> <sup>1</sup>, Rafał Jonczyk <sup>2</sup>, Guillaume Thierry<sup>†</sup> <sup>3,4</sup>, Shan Gao<sup>‡</sup> <sup>5,6</sup>

<sup>1</sup> School of Foreign Languages, University of Electronic Science and Technology of China – China

<sup>2</sup> Faculty of English, Adam Mickiewicz University – Poland

<sup>3</sup> Faculty of English, Adam Mickiewicz University – Poland

<sup>4</sup> School of Psychology, Bangor University – United Kingdom

<sup>5</sup> School of Foreign Languages, University of Electronic Science and Technology of China – China

<sup>6</sup> MOE Key Laboratory for NeuroInformation, University of Electronic Science and Technology of China – China

Accumulating evidence shows that reading words in a second language (L2) automatically activates native language (L1) representations in bilinguals. Intriguingly, previous research reveals that access to such representations is filtered by negative emotional valence. When Wu and Thierry (2012) manipulated the emotional valence of prime words in a translation-priming paradigm, they found that prime words with a negative valence abolished N400 modulation by character repetition in L1 translation. Whilst priming was found for both positive and neutral primes, negative words seemed to inhibit automatic access to L1 representations during L2 word reading. Recently, Zhang et al. (2023) replicated this finding and extended it by showing that inhibition applies with a delay, that is, not for the negative word itself but for the word presented immediately afterwards.

It is unknown, however, whether such valence-dependent inhibition effects are only triggered by linguistic input or whether they extend to domain-general emotion regulation. To test this, we manipulated the emotional context in an implicit translation priming paradigm using pictures as emotional primes instead of words. In Experiment 1, Chinese-English bilinguals reported whether either of two English (L2) words presented in sequence after a picture was semantically related to the picture, whilst being unaware that some of the word pairs concealed a sound repetition when translated into Chinese (L1). Experiment 2 was a control experiment conducted in a different group of Chinese participants based on the exact same design but using Chinese words.

In Experiment 1, we found the expected reduction in N400 amplitude elicited by L1 phonological repetition when the L2 word pairs were preceded by a picture with a positive valence (Figure 1). When the picture had a negative valence, however, there was no priming, suggesting that unconscious activation of L1 translation equivalents was inhibited. In Experiment 2, we found the phonological priming effect obtained irrespective of picture valence, when words were presented in L1 Chinese and thus did not require access from L2 English.

These findings replicate, once more, the inhibition of L1 lexical access from L2 in Chinese-English bilinguals, but this time triggered by a non-verbal emotional cue. Since negative pictures can trigger the same inhibitory effect as that observed previously for negative prime words, the modulation of language non-selective lexical access in bilinguals observed here goes beyond the

---

\*Speaker

†Corresponding author: g.thierry@bangor.ac.uk

‡Corresponding author: gaoshan@uestc.edu.cn

language realm and is domain-general. We conclude that it is not only language that influences other aspects of higher cognition in humans, but also non-primarily linguistic domains of cognition, such as affect, that can modulate mechanistic aspects of language processing.

**Keywords:** bilingualism, emotion, ERPs, priming, multimodal

# Explicit instruction helps only at the beginning: Children's learning of vocabulary and grammar from cross-situational statistics.

Wensi Zhang \*<sup>1</sup>, Padraic Monaghan<sup>1</sup>, Patrick Rebuschat<sup>1,2</sup>

<sup>1</sup> Lancaster University – United Kingdom

<sup>2</sup> University of Tübingen – Germany

Explicit instruction helps children to learn language in the classroom. However, language learning also occurs in naturalistic situations. How does instruction affect learning in these other situations? Studies have shown that adults are able to learn vocabulary and grammar from cross-situational statistics (CSL); our ability to track the association of words in utterances and their referents in the environment across ambiguous situations as in naturalistic communicative situations (2,3). It is less clear that children are able to acquire both of them in the same way as adults. Also, CSL is generally conceived as involved in implicit language learning. However, little is known about how explicit information about language structure might constrain the kinds of mappings that children are able to draw from cross-situational statistics, or whether the associations formed during cross-situational learning are implicit and impervious to instruction. In this study, we will directly address this gap by investigating the extent to which children could learn both vocabulary and grammar simultaneously through cross-situational statistics and testing the effect of explicit instruction on learning under two exposure conditions, implicit and explicit.

We trained and tested fifty children, aged 8 to 9 years, on a complex artificial language comprised of 12 pseudo-words including nouns, verbs and case markers in a verb-final sentence order by means of a widely-used CSL paradigm (e.g., (2,3)). All participants in implicit and explicit instruction condition were asked to determine which of the two animated scenes the complex transitive sentences related to. Children in the explicit instruction condition were also taught the language structure (the verb-final word order), while no such explicit information was provided for children in implicit condition. Children's awareness of grammar knowledge was measured through debriefing questions after the learning of the artificial language.

Results showed that children in both conditions learned grammar but not vocabulary greater than chance, indicating that children's learning of a second language, at least of grammar, can proceed without explicit instruction or requiring consciously extracting the language structure (1). There was no evidence for the effect of instruction on grammar and vocabulary acquisition, though we did find an initial benefit of it to learning. However, explicit instruction of the language structure significantly improved awareness of grammar knowledge, yet this increased explicit awareness did not relate to grammar learning. The results have implications for theories of how L2 learning proceeds in child L2 acquisition and our understanding of whether explicit instructions can interoperate with statistical learning in supporting children's L2 acquisition. The study also has practical implications for optimising interactions between language teaching strategies and the children's learning environment. This includes, for example, determining

---

\*Speaker

what kinds of instruction influences acquisition of which language features.

#### Reference

- (1) Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). *Journal of Memory and Language*, 83, 62–78.
- (2) Monaghan, P., Ruiz, S., & Rebuschat, P. (2021). *Second Language Research*, 37(2), 261–289.
- (3) Rebuschat, P., Monaghan, P., & Schoetensack, C. (2021). *Cognition*, 206, 104475.

**Keywords:** Cross, situational statistical learning, Implicit learning, Explicit instruction, Child language acquisition

# Metaperceptions of L1 and L2 English Speakers During Interaction: The Role of Individual Differences

Chaoqun Zheng <sup>\*</sup> <sup>1</sup>, Anamaria Bodea <sup>\*</sup>

<sup>1</sup>, Pavel Trofimovich <sup>1</sup>, Kim McDonough <sup>1</sup>

<sup>1</sup> Concordia University – Canada

Both first (L1) and second (L2) language speakers underestimate how much others like them and how much others appreciate their speaking ability and their interactional skill during conversations (Boothby et al., 2018; Trofimovich et al., 2023). Speakers' beliefs about how they are perceived by others are referred to as *metaperceptions*, and these judgments play an important role in various forms of speaker decision-making. For example, if speakers believe that they are perceived unfavourably by their interlocutor, they are less likely to ask for assistance, solicit feedback, and engage in a future conversation with the same interlocutor (Mastroianni et al., 2021; Trofimovich et al., 2023). Because L2 interaction is a key component of language learning and use (Lantolf et al., 2020), it is important to identify which specific inter-individual differences predict metaperceptions, potentially explaining why speakers feel underconfident about themselves and their communication skills. The goal of this study was therefore to investigate which individual personality, cultural, affective, and social variables are related to speakers' metaperceptions during interaction and whether the role of these variables differs for L1 and L2 speakers.

Participants included 29 dyads of unacquainted L1 and L2 English students ( $N = 58$ ) from English-medium universities in Canada. The students completed a 10-minute academic discussion task for which they first read short academic texts on a familiar topic and then discussed their opinions. After the task, the students used 100-point scales to first rate each other's interpersonal liking (e.g., *I like the student*), speaking skill (e.g., *I liked how easy the student was to understand*), and interactional behavior (e.g., *I liked how well the student responded to my ideas*). The students then provided their metaperceptions of their partner's ratings for the same three dimensions (e.g., *I think the student liked me*, *I think the student liked how easy I was to understand*, *I think the student liked how well I responded to their ideas*). Finally, they completed questionnaires targeting their personality (Soto & John, 2017), acculturative stress (Benet-Martínez & Haritatos, 2005), language anxiety (MacIntyre & Gardner, 1994), and social network (Doucerain et al., 2015).

All students showed a bias to underestimate their interpersonal liking, speaking skill, and interactional behavior, where their metaperceptions were significantly lower than their interlocutors' actual ratings. In addition, metaperceptions were associated with multiple individual variables. For L1 students, personality explained metaperceptions of interpersonal liking ( $R^2 = .13$ ) and speaking skill ( $R^2 = .09$ ), where lower agreeableness scores were associated with lower metaperceptions. Interactional behavior, however, was not predicted by any variables. For L2 students, social network size predicted metaperceptions of interpersonal liking ( $R^2 = .14$ ), where greater

---

<sup>\*</sup>Speaker

L1 speaker networks were associated with higher metaperceptions; language anxiety, acculturative stress, and social network predicted metaperceptions of speaking skill ( $R^2 = .30$ ); finally, language anxiety predicted metaperceptions of interactional behavior ( $R^2 = .08$ ). In all cases, greater language anxiety and acculturative stress were associated with lower metaperceptions. We discuss the implications of these findings for the development of intercultural linguistic competence for L1 and L2 speakers.

**Keywords:** metaperception, individual differences, second language, interaction, interpersonal liking, speaking, interactional behavior

# ANALYTIC CHINESE CHARACTER RECOGNITION AMONG NONNATIVE SPEAKERS

Ying Zhou <sup>\*† 1</sup>, Nan Jiang <sup>2</sup>

<sup>1</sup> Xinjiang University – China

<sup>2</sup> University of Maryland System – United States

The recognition of Chinese characters or words may be holistic or analytic. The two processing strategies differ in terms of whether a character is treated as a single processing unit or processed sublexically, or in terms of whether the components of a character (strokes, radicals, stroke patterns) are processed in parallel or serially. The comparison of native and nonnative speakers in this regard has received limited attention in previous research and the results have been inconclusive.

The previous studies focusing on a stroke number effect have found some evidences of the difference of processing patterns among Chinese native and nonnative speakers. They found that there was a significant prolonged RT with the increasing number of strokes while reading characters among nonnative speakers, but no such effect was found in native ones (Jiang, Hou & Jiang, 2020; Jiang & Feng, 2022; Zhou & Jiang, 2023). However, to investigate this issue from the angle of the number of strokes had some limitation, for the patterns of processing could only be inferred indirectly. In the present study, we probed this issue from the radical level of characters, so that we could explore whether there is a decomposition while recognizing a character directly in the target groups.

This presentation reports the results of two experiments involving two different paradigms: the priming paradigm and a false-memory approach. Findings from both experiments demonstrated the adoption of a more analytic strategy among nonnative speakers. Specifically, nonnative speakers showed a whole-component character priming effect with prime-target pairs such as - while no such priming effect was found among native speakers. Second, it showed that nonnative speakers were more likely to produce a false memory effect than native speakers in character recognition. In this case, participants were shown a set of Chinese characters in the study phase and were asked to decide if they saw a character in the test phase. Nonnative speakers were more likely to produce an incorrect positive response to a character (e.g., ) in the test phase that was not present in the study phase but was embedded in an earlier-encountered character (e.g., ).

We will discuss a) the causes of the nonnative speakers' tendency to process Chinese characters analytically, b) the relationship between analytic processing and reading development, and c) strategies to help learners transition from analytic to holistic processing.

**Keywords:** Chinese characters, analytic processing, holistic processing, priming paradigm, false memory approach

---

\*Speaker

†Corresponding author: daisy6663@163.com



# Blocking or highlighting? The distributional effect of inflectional and adverbial cues

Liuqi Zhu <sup>\*† 1</sup>, Padraic Monaghan <sup>1,2</sup>, Patrick Rebuschat <sup>1,3</sup>

<sup>1</sup> Lancaster University – United Kingdom

<sup>2</sup> Max Planck Institute for Psycholinguistics – Netherlands

<sup>3</sup> University of Tübingen – Germany

In languages with verbal morphology, time is indicated in both lexical (e.g., "yesterday" in English) and morphological cues (e.g., "-ed"). Ellis and Sagarra (2010) observed a "blocking effect", where training on an adverbial cue inhibited learning of a subsequent morphological cue to tense. However, in naturalistic language learning, cues can occur simultaneously rather than successively, and furthermore they do not occur reliably. For example, in English, time reference can occur with just a lexical cue, a morphological cue or both. How does learning proceed when the reliability of cues varies? The present study addressed this question by using a cross-situational learning (CSL) paradigm. While previous CSL studies (e.g., Rebuschat et al., 2021) have shown that adults can learn both vocabulary and syntax by tracking cross-situational statistics, it is still unclear if and how morphological cues can be learnt in this fashion.

In CSL tasks, eighty native English speakers observed two scenes (with cartoon animals performing different actions in the past, present or future), heard an artificial language sentence that described one of the scenes, and chose which scene the sentence referred to. There were two between-subject conditions: consistent and variable. In both conditions, there were two temporal cues, namely adverbs and tense morphemes. The adverbs were always disyllabic and thus perceptually more salient than the morphemes, which were monosyllabic. In the consistent condition, the two temporal cues were consistently present, and they were thus equally reliable predictors of time. In the variable condition, each cue was present in 2/3 of the sentences, i.e. sometimes cues were both present, sometimes they occurred in isolation.

Our results showed that five minutes of exposure sufficed to observe learning of the artificial language across both conditions. Participants in the consistent condition significantly outperformed participants in the variable condition, suggesting that cue reliability boosted learning. For the consistent condition, there was no blocking effect on either of the cues, in contrast to Ellis and Sagarra (2010) where cues were learnt sequentially. For the variable condition, learning was significantly above chance for adverbial cues while morphological cues were at chance. Variable cues tuned learners' attention to the more salient adverbial cue and blocked the less salient morphological cue.

These findings extend our understanding of the blocking effect in a more naturalistic language environment where cues are variable and occur simultaneously in speech. Ellis and Sagarra (2011) suggested that native language(s) plays a role in the blocking effect for L2 learning of morphology; our study further shows that the salience difference between two variably-occurring cues turns learners' attention to the adverbial cue and contributes to the blocking effect when acquiring an additional language. When multiple cues are not variable, no blocking was observed

---

\*Speaker

†Corresponding author: l.zhu8@lancaster.ac.uk

within the language.

**Keywords:** cross, situational statistical learning, morphology, learned attention, blocking effect

# AI-enabled Lexical Glossing for Incidental L2 Word Learning

Tong Zhu <sup>\*†</sup> <sup>1</sup>, Derek Irwin <sup>1</sup>, Yanhui Zhang <sup>1</sup>, Renjie Wu <sup>1</sup>

<sup>1</sup> University of Nottingham Ningbo [China] – China

Artificial intelligence (AI) has been rapidly integrated into personalized language learning, optimizing instructional approaches and learning pace by catering to learners' individual differences in interests, needs, and abilities (Huang et al., 2023). It is often argued that vocabulary is the foundation of successful second language (L2) learning and reading is the main source of vocabulary input (Boers, 2022; Krashen, 1989; Pigada & Schmitt, 2006). AI-enabled adaptive algorithms have been applied to personalized L2 reading systems, which recommend L2 reading texts based on each individual learner's topic interest and reading ability (Ehara et al., 2013; Hsieh et al., 2012; Lee et al., 2022; Lee & Yeung, 2018; Wang, 2014). However, a limitation of these systems is that learners' unknown words in reading are not well identified and therefore glossed, which may reduce learning efficiency. To develop a more effective personalized system to facilitate L2 reading and word learning, the estimation of the likelihood of words being known plays a vital role in determining text coverage and readability, identifying in-text unknown words, and providing lexical glosses for them. Traditionally, the estimation depends on user clicks and computer-based lexical profiling. However, the underlying assumptions of lexical profiling units (i.e., word family, lemma, and word type) and criteria (i.e., word frequency in British National Corpus, the Corpus of Contemporary American English, and Nation's BNC/COCA word list) remain to be further explored. The present study aims to address these research gaps by 1) investigating the optimal lexical profiling criteria and unit, and 2) integrating learner-based word knowledge data from various empirical studies to calibrate word frequency.

This study selected a readable and popular English as an L2 (ESL) reading text to be profiled and tested to achieve the above goals. Test items were 932 content words retrieved from the sample text and were designed to assess test takers' word knowledge to a meaning-recall level. Five hundred and twenty ESL students from 33 top-tier universities in China participated in this online test. Their responses, which determined their unknown words in the sample text, were compared with lexical profiling results based on different lexical units and criteria. As a result, lemma, the lexical unit frequently used for lexical profiling of texts, was found to best represent participants' word knowledge. Results also showed statistically significant differences and moderate correlations between corpora-based word frequency and learner-based word knowledge, indicating that word frequency needs to be calibrated to better represent the likelihood of words being known. A weighting-based multi-factor assessment method was proposed to integrate learner-based word knowledge data in the present study with Schmitt et al.'s (2021) KVL list, to calibrate word frequency and better represent the likelihood of words being known.

Future research is suggested to further empower AI in language education by verifying the calibrated likelihood of words being known and extending findings to more words and learners with diversified language backgrounds.

---

\*Speaker

†Corresponding author: hvxtz4@nottingham.edu.cn

**Keywords:** AI in language education, lexical unit, word frequency, gloss, incidental L2 word learning

# Training preservice EFL teachers on nontraditional dialects of English

Germán Zárate-Sández \* <sup>1</sup>

<sup>1</sup> Western Michigan University – United States

This paper examines the effects of training preservice Brazilian teachers of English as a Foreign Language (EFL) to understand and appreciate nontraditional dialects of English. Expanding second language (L2) teachers' *dialectal competence*, defined as attitudes toward and knowledge of nontraditional dialects, is important because prioritizing a traditional or prestigious variety in the classroom invariably excludes the less prestigious varieties and the communities associated with them. (Devereaux, 2015). Teachers, in turn, can transmit these negative attitudes to their students. Brazil is not immune to these tensions. Research has shown, for example, that Brazilian universities have privileged American and British English to the detriment of other English dialects (Gimenez et al., 2016). In this context, the current project sought to answer the following research question: To what extent does a pedagogical intervention aimed at familiarizing teachers with the pronunciation of non-standard dialects of English affect their dialectal competence? The project focused on English pronunciation because dialectal variation is more readily noticeable in pronunciation than it is in grammar or vocabulary.

Participants were 20 Portuguese-speaking students in the EFL teacher-training program at a state university in southern Brazil. Twelve participants assigned to the experimental group completed four computer-delivered sessions where high variability phonetic training (HVPT) was used to train participants to discriminate and classify pronunciation features of South African and Caribbean English. HVPT rests on the principle that humans learn the sounds of a new language (including our first language) from multiple speakers (hence the *high variability* component). In L2 research, this technique has shown promise to help L2 learners improve their perception of L2 pronunciation (Barriuso & Hayes-Harb, 2018; Thomson, 2018). Pre- and posttests-designed to test the effectiveness of the training-consisted of surveys and interviews that explored participants' attitudes and beliefs, as well as comprehensibility tests that measured understanding of the dialects on which participants were trained. Eight participants also completed the tests but did not participate in the training (control group). A subset of participants from both groups was invited to participate in semi-structured interviews to probe deeper into their opinions and attitudes regarding nontraditional English dialects.

Data analysis consisted of quantitative analyses of Time (pre- and posttest) by Group (experimental and control) interactions, as well as qualitative analyses of major themes in interviews and surveys. Compared to controls, participants who received training improved sensitivity to dialect-specific phonetic features, as well as their attitudes toward these dialects, suggesting that mere exposure to the pronunciation of less commonly taught dialects can alter opinions and beliefs. Changes in individual trajectories, however, were often inconsistent and tied to participants' perceptions of themselves as current English learners and future English teachers. The discussion seeks to contribute to models of dialect familiarity effects in L2 phonetic training (e.g., Lui & Gibon, 2022; Ross et al., 2021) as well as further understand identity development among L2 teachers (e.g., Karimi & Mofidi, 2019).

---

\*Speaker

**Keywords:** teacher training, English dialects, pronunciation

# Methods and methodological challenges in the study of morphology in L2 French

Malin Ågren \* <sup>1</sup>

<sup>1</sup> Språk- och litteraturcentrum - Centre for Languages and Literature (SOL) – Lunds universitet - Box 201 - 221 00 Lund, Sweden

The acquisition of morphology is a well-studied phenomenon in SLA and has attracted attention from many linguists over the years due to its persisting difficulty in L2 learning (Ellis 2022). The present study focuses on morphology in L2 French and the objective is to give an overview of methods used in this area of research and to discuss methodological challenges emerging from the literature. The paper is part of a larger project, focusing on methodological approaches and considerations in the study of different sub-fields of French SLA (Howard, forthcoming). Focusing on different morphological phenomena, we show that gender assignment and agreement in the noun phrase and tense and aspect in the verb phrase are the most well-studied morphological domains in L2 French over the last decades. A continuum of methodological approaches is identified, ranging from qualitative to quantitative, from longitudinal to cross-sectional and from elicitation of spontaneous production to experimental designs. A developmental trend is observed from earlier studies based on conversational data or narratives, focusing on the developmental trajectories of morphological forms and error types in individual learners, to studies involving larger cohorts and experimental methods, such as grammatical judgements or event related potentials (ERPs). Furthermore, we identify limitations in the range of proficiency levels explored and in the variety of participants studied, many of which are instructed learners of French at intermediate levels speaking a Germanic first language. We conclude that future research needs to include learner groups with typologically more varied L1s, preferably in the same studies. This will help us better understand the role of the L1 when acquiring morphology in L2 French and to directly compare specific learning contexts (Frenck-Mestre et al. 2009; Ågren et al. 2021).

Turning to the methodological challenges revealed in the literature, we identify three specific domains relevant for the study of morphology in L2 French: 1) the definition and operationalization of gender assignment and agreement (Kupish, Akpınar & Stöhr 2013), 2) the transcription and analysis of homophonous verb forms in (e) (Leclercq 2020) and, finally, 3) the discrepancy between spoken and written verb morphology in French (Ågren & van de Weijer 2013). In all three areas, definitions, operationalizations and analyses of morphological concepts vary, which make results difficult to compare. In these morphological domains in particular, researchers need to clearly define concepts in relation to theory and to previous studies, to give good examples of how they categorize and analyze different morphological forms, and to continuously evaluate elicitation techniques and methods used in the literature.

We conclude that French is a complex language when it comes to morphology, and some aspects are more challenging than others in methodological terms. It is therefore of crucial importance that researchers clearly state how they handle these characteristics when studying the development and use of morphology in L2 French.

---

\*Speaker

**Keywords:** SLA, L2 French, morphology, research methods: methodology, methodological challenges



# Switching from a variable stress system to a fixed stress system: the case of L2 Hungarian

Öner Özçelik \*† <sup>1</sup>

<sup>1</sup> Indiana University – United States

Non-target-like prosody plays a crucial role in the perception of foreign accentedness in a second language (L2) (e.g. Trofimovich & Baker 2007). Little research has, however, been done to understand L2 prosody, and the focus has been almost exclusively on L2 English (e.g. Archibald 1992, 1993, Pater 1997, Tremblay 2007). L2 acquisition of stress/prominence in fixed-stress languages such as Hungarian and Turkish, which have fixed word-initial/final prominence respectively, has almost never been investigated. One reason for this trend is the assumption that, in these languages, acquisition should be easy, i.e. error free, so that it will not provide much insight into the abstract generalizations made by L2 learners. In this study, focusing on L2 Hungarian, which has left-to-right, non-iterative, weight-insensitive trochees (e.g. ((kr)kdlm) ‘commerce’), I demonstrate that the acquisition task in these languages is far from simple. I hypothesize, along with the Prosodic Acquisition Path Hypothesis (Özçelik 2018), that prosodic parameters with *Yes/No* values, as with Weight-Sensitivity and Iterativity, are more difficult to reset from *Yes* to *No* than from *No* to *Yes*, given certain input-related considerations (e.g. availability and quality of positive evidence). I further hypothesize that no such dichotomy exists for prosodic parameters with *Left/Right* values, as with Headedness (Trochaic vs. Iambic) and Directionality (Left-to-Right vs. Right-to-Left footing), for these values are equal in terms of availability and quality of input. Accordingly, I predict that the acquisition process will be far more difficult for L1 English-speaking learners of L2 Hungarian than for L1 Hungarian-speaking learners of L2 English, as the task involves, for the former, switching from an iterative and weight-sensitive grammar (e.g. (a(méri)) vs. ((àt)(lán))) to a non-iterative and weight-insensitive grammar. Although the two languages also differ in directionality of footing, with English being right-to-left and Hungarian left-to-right, no difficulties are predicted in either population with resetting Directionality.

In order to investigate these predictions, a bidirectional semi-controlled production experiment was conducted with English-speaking learners of L2 Hungarian (n=6) and Hungarian-speaking learners of L2 English (n=8), of various proficiency levels. The stimuli, composed of 60 words, of various lengths and syllable structure profiles, were first uttered in isolation, and then in a carrier sentence. Only the latter were transcribed and analyzed for acoustic measures. The results of the experiments confirm that English-speaking learners of L2 Hungarian had much greater difficulties than Hungarian-speaking learners of English. That is, it is much more difficult to reset Weight Sensitivity and Iterativity from *Yes* to *No* than *No* to *Yes*. Although the former group was able to make various UG-constrained changes to their grammar, such as resetting Directionality from *Right-to-Left* to *Left-to-Right* in order to produce consistently left edge stress, they were unable to produce words with unstressed closed syllables (weight insensitivity) or (longer) words without secondary stress (non-iterativity). That is, although the task

---

\*Speaker

†Corresponding author: oozcelik@indiana.edu

for L2 Hungarian learners seems much simpler cognitively (i.e. ‘simply stress the initial syllable of every word’), it is harder linguistically, suggesting that L2 acquisition is domain-specific.

**Keywords:** stress, Universal Grammar, acquisition of prosody, weight sensitivity, iterativity, fixed vs. variable stress, acquisition of word stress, difficulty vs. ease in acquisition of phonology

# The relationship between auditory alertness, mental transformation, and previous EFL learning in young learners' foreign language aptitude for a novel language

Jasenka Čengić <sup>\*† 1</sup>, Agnieszka Kaldonek-Crnjaković <sup>\*</sup>

<sup>2</sup>, Tanja Angelovska <sup>\*</sup>

3

<sup>1</sup> University of Zagreb – Croatia

<sup>2</sup> University of Warsaw – Poland

<sup>3</sup> University of Salzburg – Austria

Recent definitions of foreign language (FL) aptitude describe it as a set of cognitive abilities that play a major role in both second and FL learning (Kiss & Nikolov, 2005; Li, 2022). It is considered that for adults, FL aptitude gives incremental predictive validity (Curic et al., 2019; Doughty, 2019). Aptitude research focusing on young learners (YLS) considers aptitude to be dynamic (Biedroń, 2022) in its nature and composed of multiple independent components (Alexiou, 2005). Previous studies in FL aptitude of YLS confirmed the predictivity of both domain-specific measures, such as auditory alertness (Muñoz, 2014) and language analytic ability (Roehr-Brackin & Tellier, 2019), as well as domain-general measures, such as spatial ability (Alexiou, 2005). Therefore, the focus of this study was to investigate the incremental predictive validity of auditory alertness and language analytic ability, working memory (Massonie et al., 2022), non-verbal reasoning (Ehrlich et al., 2006) and type and frequency of previous English as a FL (EFL) learning for YLS aptitude.

The study involved first language (L1) Croatian YLS of English as a second language (N=209; mean age=7.81; SD = 0.47) at the beginning of their formal schooling. A new aptitude measure was designed following the miniature natural language learning paradigm proposed by Kempe and Brooks (2016). We used a natural language (Hungarian), which was a fully novel language to the participants. The final version of the aptitude test consisted of two measures: an auditory alertness measure and a language analytic ability measure consisting of two subtasks (auditory morphophonological pattern recognition and auditory perceptual acuity). Non-verbal reasoning was measured via a mental transformation task (Ehrlich *et al.*, 2006), whereas verbal working memory was assessed via a backward digit span (Massonie et al., 2022). YLS language learning achievement at the end of grade one was measured via a listening comprehension task designed by Enever et al. (2011). Previous language learning was taken as an independent variable composed of information provided by the parents about: (1) whether the children previously attended formal English instruction, (2) the type of instruction (private group course or individual lessons) and (3) the frequency of the instruction (number of lessons per week).

---

\*Speaker

†Corresponding author: jasenka.cengic@gmail.com

The results point towards auditory alertness being strongly positively correlated with both verbal working memory ( $r(201) = 0.19$ ) and mental transformation ( $r(201) = 0.23, p < 0.001$ ). Finally, the incremental validity of FL aptitude was confirmed by auditory alertness explaining 4% of variance of overall success on the listening comprehension measure. Overall, 28% of variance of success at the end of the first year of FL learning was explained with the new model of YLs FL aptitude construct comprising auditory alertness, mental transformation, and prior EFL learning. This result is in line with Rabagliati et al. (2016) who could not conclude that prediction serves children rapid learning.

The findings highlight the contribution of previous exposure to EFL for the development of met-alinguistic awareness of children which presumably leads to the development of the FL aptitude and thus, helps the subsequent acquisition of other FLs during lifetime.

**Keywords:** language aptitude, young learners, English as a foreign language

# Surprisal effects in L2 structural priming of English dative constructions

Duygu F. Şafak <sup>\*† 1</sup>, Holger Hopp <sup>1</sup>

<sup>1</sup> TU Braunschweig – Germany

In English, most ditransitive verbs can take both double-object dative (DO) and prepositional-object dative (PO). This optionality is constrained by gradient selectional restrictions on the type of their complements: DO-bias verbs, e.g., *pay*, tend to occur more frequently with DO, while PO-bias verbs, e.g., *send*, show a probabilistic tendency to prefer PO (e.g., Şafak & Hopp, 2023). Research on priming in L1 comprehension indicates that the magnitude of structural priming increases when the prime structure goes against its verb’s bias (e.g., Chen et al., 2022; Fine & Jaeger, 2013). Such surprisal effects of verb bias among adult L1 speakers and child L1 learners are in line with implicit learning models, claiming that prediction error is a key mechanism underlying structural priming (e.g., Dell & Chang, 2014). However, it is an open question whether prediction-error-driven implicit learning accounts of priming extend to L2 learners. Against this backdrop, we probe whether prediction error constitutes a mechanism of structural priming also in a late-learned L2. In two experiments combining priming and visual world eye-tracking, we first test whether prediction error, occasioned by verb bias, affects structural priming in L2 comprehension. Second, we examine whether structural priming occurs in the absence of prediction error.

In Experiment 1, adult L1 German intermediate to advanced L2 learners of English ( $n = 48$ ) first read aloud prime sentences crossing Verb Bias (DO-bias/PO-bias) and Structure Type (DO/PO). Subsequently, they listened to target sentences with non-biased verbs (e.g., *show*) while viewing visual scenes with an agent referent (e.g., *the tailor*), a theme referent (e.g., *the dress*), and a recipient referent (e.g., *the model*). Cluster-based permutation analyses revealed PO-priming, as evidenced by more looks to the theme than to the recipient during the post-verbal temporal region in target sentences following PO (vs. DO) prime sentences. Further, PO-priming was modulated by surprisal effects of verb bias, as priming was larger when the prime structure mismatched the bias of prime verb, i.e., after PO primes with DO-bias verbs. These effects show that L2 learners adapt to the structure of the recently processed prime sentence through learning from their prediction errors.

In Experiment 2 with a comparable group of L1-German–L2-English learners ( $n = 48$ ), the materials were the same as in Experiment 1, except that prime sentences were preceded by a context sentence introducing the themes and recipients. Moreover, the full noun-phrase themes and recipients in PO prime sentences were replaced by their pronominal counterparts. DO-bias prime verbs thus elicited no prediction error, given that PO is the only possible structure for any ditransitive verb with two pronominal complements. As a consequence, PO-priming did not reach significance in Experiment 2. This suggests that, when there is no prediction error to learn from, structural priming may not be strong enough to affect L2 learners’ processing of the target sentences.

The findings from Experiments 1 and 2 demonstrate that priming in L2 comprehension is driven

---

\*Speaker

†Corresponding author: d.safak@tu-braunschweig.de

by prediction error like in L1 comprehension, and further point to an error-based prediction mechanism underpinning L2 structural priming.

**Keywords:** L2 sentence processing, structural priming, prediction error, dative alternation, verb bias

# Author Index

- Abe, Nozomi, 6  
Abel, Niklas, 241  
Addo, Fredina, 8  
Agebjörn, Anders, 10, 365  
Akef, Soroosh, 12  
AKINCI, Mehmet, 2  
Al-Maani, Alaa, 308  
Alazard-Guiu, Charlotte, 15  
Alger, Mari, 14  
Ali, Saandia, 126  
Almeida, Letícia, 15  
Alraddadi, Hadil, 17  
Andorno, Cecilia, 19  
Andringa, Sible, 359  
Angelovska, Tanja, 467  
Argyri, Froso (Effrosyni), 21  
Arnaus Gil, Laia, 161, 258  
Attal, Anna, 111  
Avello, Daniela, 23  
AWADA, Layana, 4  
Aykurt Buchwalter, Sülün, 25  
Azpilicueta-Martínez, Raúl, 27
- Baills, Florence, 31  
Baker, Lewis, 33  
Baqué, Lorraine, 31  
Baranowska, Karolina, 316  
Bayram, Fatih, 385  
Bear, Elizabeth, 35  
Beauvoir-Hourdel, Pauline, 37  
BELIA, Eftychia, 29  
Benazzo, Sandra, 19  
Benigno, Giuseppe Dario, 39, 41  
Bennett, Sophie, 43  
Bernolet, Sarah, 346  
Berthele, Raphael, 92, 398  
Beyer, Maike, 45  
Bibauw, Serge, 46  
Bienati, Arianna, 53  
Bodea, Anamaria, 454  
Boers, Frank, 48, 219, 430, 448  
Bokander, Lars, 318  
Boone, Griet, 50  
Borro, Ilaria, 78
- Bottini, Raffaella, 201, 272  
Bowden, Harriet, 222  
Brakovec, Hunter, 51, 82  
Brasolin, Paolo, 53  
Braunewell, Aylin, 55  
Bril, Marco, 57  
Bromberek-Dyzman, Katarzyna, 169  
Bruhn, Ann-Christin, 59, 183  
Bsharat-Maalouf, Dana, 334
- Cabrelli, Jennifer, 197  
Caltabellotta, Eva, 65  
Carbonara, Valentina, 67  
Castle, Chloe, 69, 426  
Cerezo, Lourdes, 233  
CHABANAL, Damien, 165  
Chandy, Radha, 71  
Chaouch-Orozco, Adel, 73  
CHEN, ZHUOHAN, 61  
Chung, Yoojin, 326  
Cilibrasi, Luca, 75  
Clenton, Jon, 77  
Copin, Mireille, 193  
Copot, Maria, 436  
Correia, Susana, 43, 310, 373  
Cremers, Mieke, 57  
Criado, Raquel, 405  
CRISTINA, ALIAGA-GARCIA, 63  
Cruz Enríquez, Maura, 284
- D'Angelo, Larissa, 78  
Dabrowska, Ewa, 181  
Daems, Joke, 115  
Darcy, Isabelle, 82  
Daskalaki, Evangelia, 357  
Daveluy, Rachel, 84  
De Cat, Cecile, 385  
De Cristofaro, Elisa, 86  
De Fino, Verdiana, 80  
De Haes, Hanna, 88  
De Paolis, Bianca Maria, 90  
De Wilde, Vanessa, 50  
Degani, Tamar, 334  
Degrave, Pauline, 217

Deme, Andrea, 373  
 DETEY, Sylvain, 80  
 Dherbey Chapuis, Nathalie, 92  
 Dick, Fred, 173  
 Dimroth, Christine, 19  
 Dinis Fernandes, João, 43  
 Divjak, Dagmar, 220  
 Dosi, Ifigeneia, 94  
 Drackert, Anastasia, 96  
 Du, Chenxi, 98  
 Du, Yueming, 99  
 Duan, Xuewen, 101  
 Dvorina, Natasha, 136  
 Długosz, Kamil, 103  
  
 EDMONDS, Amanda, 37, 105  
 Egger, Katharina, 107  
 El Ayari, Sarra, 109  
 El-Dakhs, Dina, 355  
 Ender, Andrea, 432  
 ERCETIN, Gülcan, 2  
 Esteve Gibert, Núria, 31  
 Eyckmans, June, 14  
  
 Faez, Farahnaz, 123  
 Faretta-Stutenberg, Mandy, 222  
 Fernandes, João Dinis, 310  
 Fernandez Cuenca, Sara, 111  
 Fernández Santos, Sara, 181  
 Ferraresi, Adriano, 209  
 Finger-Bou, Rebeca, 113  
 Fontan, Lionel, 80  
 Fonteyne, Margot, 115  
 Fraser, Simon, 77  
 Fu, Xingyi, 117  
 Fukuta, Junya, 119  
 Fullana, Natalia, 63  
 Füreder, Birgit, 121  
  
 Gagen, Tomlin, 123  
 Galvin, Tesni, 318  
 Gamper, Jana, 55  
 Gao, Shan, 450  
 Garcia, Guilherme, 428  
 García Alcaraz, Estela, 39, 41  
 García-Guerrero, Elena, 124  
 Garnier, Marie, 126  
 Gastmann, Freya, 146  
 Gavriilidou, Zoe, 94  
 Ge, Yuxin, 128, 310, 420  
 Gerbier, Emilie, 105  
 Gil, Kook-Hee, 226  
  
 Gilabert, Roger, 175, 253, 389  
 Goad, Heather, 428  
 Godfroid, Aline, 149  
 Golegos, Angelika, 130  
 González-Fernández, Beatriz, 132, 367  
 Gorba, Celia, 134, 211  
 Gordishevsky, Galina, 136  
 Granfeldt, Jonas, 138, 314  
 Granget, Cyrille, 140, 247, 312  
 Guijarro-Fuentes, Pedro, 39, 41  
 Gullberg, Marianne, 361  
 Guo, Ziwei, 191  
 Guzzo, Natália, 428  
  
 Haiden, Martin, 140  
 Han, Zhixing, 142  
 Hartsuiker, Rob, 346  
 Hemforth, Barbara, 436  
 Henry, Alastair, 138, 314  
 Hervé, Coralie, 144  
 Hiligsmann, Philippe, 217  
 Hill, Virginia, 288  
 Hopp, Holger, 146, 228, 406, 469  
 Hu, Shanshan, 148  
 Hui, Bronson, 149  
  
 Ibáñez, Socorro, 151  
 Imbert, David, 140  
 Irwin, Derek, 459  
 Issa, Bernard, 222  
  
 Jablonkai, Reka, 441  
 Jackson, Carrie N., 228  
 Jacopo, Saturno, 78, 155  
 Jadouille, Pauline, 381  
 Jakub, Przybył, 157  
 Jakupčević, Eva, 159  
 Jankowiak, Katarzyna, 169  
 Jasor, Ingrid, 160  
 Jedamski, Lynn, 183  
 Jeong, Seyoung, 187  
 Jiang, Chaofan, 324  
 Jiang, Nan, 456  
 Jiménez-Gaspar, Amelia, 161  
 Jin, Can, 163  
 Jin, Yishan, 165  
 Jochum-Critchley, Thomas, 228  
 Jonczyk, Rafał, 450  
 Jorméus, Victor, 167  
 JOUANNAUD, Marie-Pierre, 153  
 Jończyk, Rafał, 169  
 Järnefelt, Pia, 171



Kachlicka, Magdalena, 128, 173  
 Kaderoglu, Kadir, 175  
 Kanli, Eleni, 177  
 Karadimoula, Evgenia, 94  
 Karawani, Hanin, 334  
 Kałdonek-Crnjaković, Agnieszka, 467  
 Ke, Huimin, 179  
 Keijzer, Merel, 292  
 Kenanidis, Panagiotis, 181  
 Kennedy, Sara, 205  
 Kersten, Kristin, 59, 183  
 Khaled, Barkaoui, 185  
 Kimura, Takayuki, 119, 187  
 Kohama, Runa, 187  
 Kolb, Nadine, 189  
 Koronkiewicz, Bryan, 222  
 Korpala, Paweł, 169  
 Kourtali, Nektaria, 191  
 Krulatz, Anna, 213  
 Kubota, Maki, 197  
 Kupisch, Tanja, 197  
 Köpke, Barbara, 4, 193  
 Köylü, Yılmaz, 195  
  
 Laméris, Tim, 197  
 Laufer, Batia, 199  
 Lauwers, Peter, 88  
 Le Foll, Elen, 201  
 Leclercq, Pascale, 37  
 Leung, Shue Sum, 99  
 Leśniewska, Justyna, 284  
 Li, Jiayi, 324  
 Liang, Meiqiong, 203  
 Lindberg, Rachael, 205  
 Liu, Christina, 207  
 Liu, Siyuan, 209  
 Llompart, Miquel, 134, 181, 211  
 Lorenz, Eliane, 213  
 Lorette, Pernelle, 215, 379  
 Loubonia, Klisilda, 94  
 Louis, Laetitia, 217  
 Lozano, Cristóbal, 124, 302  
 Luo, Liting, 219  
 Luo, Yuhan, 220  
 Luque, Alicia, 222  
 Lázaro Ibarrola, Amparo, 322  
  
 Macis, Marijana, 300  
 Macken, Lieve, 115  
 Majdoub, Hedi, 224  
 Mangana, Anastassia, 226  
 MANOILOV, Pascale, 153  
  
 Marinis, Theodoros, 130  
 Marsden, Emma, 228, 276  
 Marsden, Heather, 228, 308  
 Martin, Clara, 442  
 Martin, Mary Elisabeth, 326  
 Martín-Villena, Fernando, 73  
 Marx, Nicole, 230  
 Mathysen, Jolien, 232  
 Matthews, John, 119  
 McBride, Sophie, 233  
 McDonough, Kim, 454  
 Meara, Paul, 318  
 Meir, Natalia, 136  
 Menares, Carla, 235  
 Mendes, Amália, 12  
 Menke-Bazhutkina, Iryna, 241  
 Meriläinen, Lea, 237  
 Merlo, Camila, 239  
 Meurers, Detmar, 12, 286  
 Michel, Marije, 142, 241, 326  
 Mifka-Profozic, Nadia, 243  
 Milicevic, Maja, 209  
 Milin, Petar, 220  
 Minor, Serge, 189, 245  
 Miralpeix, Imma, 268, 290  
 Mireille, Copin, 247  
 Mitrofanova, Natalia, 189, 245, 426  
 Mitterer, Holger, 414  
 Moden, Rebecca, 249  
 Molway, Laura, 418  
 Monaghan, Padraic, 43, 128, 324, 420, 452, 457  
 Monte, Ludovica, 19  
 Montero Perez, Maribel, 115  
 Mora-Plaza, Ingrid, 63, 251  
 Moskvina, Natalia, 253  
 Mouquet, Marine, 255  
 Mukhopadhyay, Lina, 387  
 Muñoz, Carmen, 23, 113, 298  
 Métral, Lucie, 257  
 Möller, Franziska, 96  
 Müller, Natascha, 258  
  
 Naets, Hubert, 278  
 Naranowicz, Marcin, 169  
 Nicolás-Conesa, Florentina, 233  
 Nixon, Jessie S., 260  
 Noreillie, Ann-Sophie, 65  
 Nosedá, Valentina, 155  
 Nottbeck, Emilia, 262  
  
 O'Reilly, David, 243, 264  
 O'Reilly, Jelena, 264, 266

Okada, Rune, 330  
 Olender, Paulina, 268  
 Olszewska, Anna, 103, 270  
 Ortega Duran, Mireia, 63  
 Osawa, Ryusho, 187  
 Otsuka, Hikari, 187  
 Ott, Christopher, 272  
  
 Palasis, Katerina, 105  
 Pan, Junlan, 276  
 Paquot, Magali, 278, 381, 408  
 Paradis, Johanne, 357  
 Parrish, Kyle, 280  
 Patil-Ingle, Vrishali, 387  
 Paubel, Pierre-vincent, 140  
 Pavlekovic, Renato, 316  
 Pellicer-Sanchez, Ana, 326  
 Pellicer-Sánchez, Ana, 71  
 Peters, Elke, 65, 282, 294  
 PFENNINGER, Simone, 274  
 Pichette, Francois, 284  
 Pili-Moss, Diana, 286  
 Pirvulescu, Mihaela, 288  
 Poarch, Gregory, 146  
 Popova, Margarita, 290  
 Porkert, Joanna, 292  
 Possidónio, Margarida, 15  
 Prieto, Pilar, 134, 211, 444  
 Prophète, Elie, 294  
 Puimege, Eva, 296  
 Pujadas, Georgia, 298  
 Pulido, Manuel, 300  
  
 Quesada, Teresa, 302  
  
 Radetzky, Stefanie, 306  
 Rami, Youssef, 308  
 Rast, Rebekah, 224  
 Rato, Anabela, 310  
 Ray, Rhea, 318  
 Rebuschat, Patrick, 12, 43, 128, 272, 310, 324,  
 420, 452, 457  
 Repiso, Isabel, 140, 312  
 Rivera Zurita, Marta, 39, 41  
 Roberts, Leah, 117, 264, 266  
 Rocher Hahlin, Céline, 138, 314  
 Rodina, Yulia, 385  
 Roehr-Brackin, Karen, 316  
 Rogers, Brian, 318  
 Rogers, Vivienne, 228, 318  
 ROJAS, Minerva, 111, 304  
 Romano, Anna, 320  
  
 Roothoof, Hanne, 322  
 Rothman, Jason, 197  
 Ruan, Yaoyao, 173  
 Ruiz, Simón, 324  
 Ryan, Oisín, 359  
 Révész, Andrea, 326  
 Rützi-Joy, Olivia, 328  
  
 Saddour, Ines, 193, 247  
 Saillard, Claire, 436  
 Saito, Kazuya, 128, 173  
 Sasaki, Miho, 330  
 Savory, Clare, 418  
 Scheffler, Pawel, 316  
 Schimke, Sarah, 146, 377  
 Schlauch, Julia, 55  
 Schlenter, Judith, 332  
 Schmidt, Torben, 286  
 Schmidtke, Jens, 334  
 Schroyens, Sarah, 346  
 Segura, Marta, 336  
 Serfaty, Jonathan, 338  
 Serrano, Raquel, 71  
 Seynhaeve, Shauny, 340  
 Shepperd, Louise, 342  
 Shimoyama, Urara, 187  
 Shue Sum, Leung, 344  
 Sijyeniyo, Edwige, 346  
 Sikkens, Fred, 348  
 Simon, Ellen, 88, 414  
 Situ, Liwen, 350  
 Siyanova-Chanturia, Anna, 292, 355  
 Skalba, Anna, 69  
 Skočilová, Tina, 75  
 Smallwood, Jye, 352  
 Smeets, Liz, 428  
 Smith, Tugce, 353  
 Snape, Neal, 197  
 Sneed German, Elisa, 37  
 Snoder, Per, 199  
 Sonbul, Suhad, 300, 355  
 Soto-Corominas, Adriana, 336, 357  
 Spit, Sybren, 359  
 Splendido, Frida, 361  
 Springer, Helene, 361  
 Stahnke, Johanna, 258  
 Stainer, Matthew, 326  
 Steinkrauss, Rasmus, 142, 179  
 Stoehr, Antje, 442  
 Strobl, Carola, 241  
 Su, Jiajia, 428  
 Suet-sin, Cheung, 363

Suhonen, Lari-Valtteri, 365  
 Sundqvist, Pia, 282, 294, 401  
 Suprapas, Iwarin, 367  
 Suzukida, Yui, 251, 369  
 Suárez, Maria del Mar, 175  
 Symons, Ashley, 173  
  
 Takizawa, Kotaro, 371  
 Tavares, Gabriela, 373  
 Tekin, Oguzhan, 205  
 Terada, Megumi, 375  
 Terlaak, Clara, 377  
 TERRIER, Linda, 126  
 Thierry, Guillaume, 169, 450  
 Thoma, Dieter, 306, 379  
 Thwaites, Peter, 381, 408  
 Tierney, Adam, 173  
 Timukova, Anna, 96  
 Toda Cosi, Mireia, 383  
 Tomic, Aleksandra, 385  
 Torgersen, Eivind N., 213  
 Torregrossa, Jacopo, 67, 387  
 Trager, Vera, 389  
 Treffers-Daller, Jeanine, 17  
 Trenchs-Parera, Mireia, 151  
 Trofimovich, Pavel, 205, 454  
 Trévisiol, Pascale, 224  
 Tsikulina, Alina, 391  
 Tsimpli, Ianthi, 387  
 Turner, James, 33  
 Tuzcu, Ayşen, 393  
 Twente, Leonie Regina, 395  
  
 Uchihara, Takumi, 77, 296, 369  
 Udes, Richard, 397  
 Udry, Isabelle, 398  
 Urbanek, Lukas, 400  
 Uztosun, Mehmet Sercan, 401  
  
 Valera-Garcia, Arturo, 405  
 Van de Weijer, Joost, 138, 314  
 Van Dijk, Chantal, 406  
 Van Eijdsden, Sanne, 326  
 Van Osch, Brechje, 222  
 Van Rij, Jacolien, 260  
 Van Steendam, Elke, 65  
 Vandeweerd, Nathan, 408  
 Vanek, Norbert, 117  
 Vanessa, De Wilde, 410  
 Vedder, Ineke, 412  
 Verbeke, Gil, 414  
 Verspoor, Marjolijn, 179  
  
 VILLE, Stella, 403  
 Viry, Mallaury, 416  
  
 Wakabayashi, Shigenori, 119  
 Walldén, Robert, 10  
 Wang, Xia, 420  
 Wang, Xiaofei, 48  
 Wang, Xiaoyun, 422  
 WATOREK, Marzena, 109, 224  
 Webb, Stuart, 298  
 Weissbecker, Kristina, 424  
 Wendebourg, Katharina, 286  
 Westergaard, Marit, 69, 189, 245, 332, 426  
 White, Lydia, 428  
 Whyte, Shona, 105  
 Wi, Injung, 430  
 Williams, John, 119, 438  
 Wirtz, Mason, 274  
 Wittner, Johanna, 432  
 Wiśniewska, Danuta, 157  
 Woerfel, Till, 434  
 Wolf, Johanna, 377  
 WOORE, ROBERT, 418  
 Wu, May, 179  
 Wu, Renjie, 459  
 Wu, Xiaoli, 6  
  
 Xi, Xiaotong, 444  
 Xie, Yuanhua, 203  
 XIE, Zhanglin, 436  
 Xue, Yuyan, 119, 438  
  
 Yalcin, Sebnem, 353  
 Yamane, Riku, 330  
 Yan, Rui, 441  
 Yang, Silvia, 442  
 Yang, Siqi, 324  
 Yao, Ting, 444  
 YI, WEI, 440  
 Yu, Lijia, 446  
 Yu, Xi, 48, 448  
 Yuan, Boping, 119  
  
 Zhang, Wanyu, 450  
 Zhang, Wensi, 324, 452  
 Zhang, Yanhui, 459  
 Zheng, Chaoqun, 454  
 Zhong, Yanlu, 440  
 Zhou, Ying, 456  
 Zhu, Liuqi, 457  
 Zhu, Tong, 459  
 Zárate-Sández, Germán, 461

Ågren, Malin, 463  
Öwerdieck, David, 146  
Özçelik, Öner, 465

Čengić, Jasenka, 467

Şafak, Duygu F., 469