



Lexical interference and prediction in sentence processing among Russian heritage speakers: an individual differences approach

Yanina Prystauka · Jiuzhou Hao · Reinaldo Cabrera Perez · Jason Rothman

Received: 29 December 2023 / Revised: 22 August 2024 / Accepted: 26 August 2024
© The Author(s) 2024

Abstract In line with emerging trends examining sentence processing in heritage language (HL) bilingualism and adopting an individual differences approach, the present study investigated the role of prediction and lexical interference in 68 Russian HL bilinguals in North America. The former was investigated by examining how Russian HL bilinguals processed sentences with semantically constraining verbs, whereas the latter was assessed examining how contextually globally irrelevant but locally coherent color word information is processed. The findings revealed that, as a group, participants demonstrated predictive processing and were subject to lexical interference. To understand inter-participant variation, we examined the roles prediction and/or interference play as a function of individual differences in cognitive

control (as measured by a Flanker task), performance on vocabulary and gender knowledge tasks as well as participants' individual language history and use (LHQ3, Li et al., *Biling Lang Cogn* 23:938–944, 2020). Indeed, these analyses revealed that individual language knowledge and domain-general cognitive control play crucial roles in how prediction and/or interference play out in HL processing.

Keywords Heritage language (HL) bilingualism · Sentence processing · Online eye tracking · Individual differences

Introduction

Decades of research have demonstrated that prediction—pre-activation of expected input given (sentential) context (e.g., Altmann & Kamide, 1999)—is an integral aspect of language processing (Ferreira & Chantavarin, 2018; Kaan & Grüter, 2021). Given that language is rule-governed and systematic, anticipating¹ what might come next in speech enhances processing efficiency.

¹ Here, we are using the terms 'prediction' and 'anticipation' interchangeably. However, for a more nuanced understanding of the distinction between these two terms, particularly in the context of computational models, refer to Altmann and Mirković (2009).

Y. Prystauka (✉)
University of Bergen, Bergen, Norway
e-mail: yanina.prystauka@uib.no

J. Hao · J. Rothman
UiT The Arctic University of Norway, Tromsø, Norway

R. Cabrera Perez
University of California, Irvine, USA

J. Rothman
Lancaster University, Lancaster, UK

J. Rothman
Centro de Investigación Nebrija en Cognición (CINC),
University Nebrija, Madrid, Spain

Prediction takes place in a top-down way, meaning that our prior knowledge at different scales—general knowledge about the world, the context of the conversation, sentence meaning built so far, the quality of our linguistic representations, the consistency of provided grammatical cues in the current input, etc.—affects the integration of incoming information. The other side of the coin is bottom-up processing whereby incoming lower-level information (e.g., auditory or written words) is taken as an input which modifies higher-level information. While such bottom-up information is influenced by contextual constraints (Altmann & Kamide, 1999), research suggests that the language system also processes transient linguistic representations that operate independently of these constraints (Kukona et al., 2014; Swinney, 1979; Tanenhaus et al., 1979). Thus, language comprehension necessarily and inherently relies on both bottom-up and top-down information. Moreover, as work on prediction suggests, the interplay between the two is modulated by a number of factors such as cue reliability, language users' goals and awareness, task demands and individual-difference factors such as comprehension ability, working memory, age, engagement with and proficiency in their language(s), etc. (Kaan & Grüter, 2021). Thus, in recent years, the focus in studying prediction has shifted away from the debate of whether it is an essential or necessary component of language processing to viewing it as a function of its expected utility (Kuperberg & Jaeger, 2016). Under this account, the degree to which comprehenders engage in prediction depends on their goals, confidence in their prior knowledge as well as their estimate of reliability of the input. Some of the pressing questions that are currently being explored in this domain ponder the circumstances under which prediction is most beneficial, the cues which are being used for it and how individual differences affect it.

In the context of this theoretical landscape, examining predictive processing in heritage language (HL) bilinguals, henceforth heritage speakers (HSs), could be especially illuminating. HSs are native speakers of a minority language acquired naturalistically in a context where there is a distinct majority language (Montrul, 2016; Polinsky & Scontras, 2020; Rothman, 2009; Rothman & Treffers-Daller, 2014). HSs most often (but not exclusively, see Kupisch & Rothman, 2018) wind up being dominant in the societal majority language, especially as age increases. Thus, HSs'

goals of language comprehension and ability to estimate reliability of input cues can vary considerably, not least because their exposure to input and/or opportunities for engaging with the HL are reduced compared to other types of native speakers (e.g. monolingual, L1-dominant bilinguals). While such variation exists also for monolinguals and L1-dominant bilinguals, the degree of this variation is much more pronounced across individual HSs. Amongst other co-morbid factors, this HS reality has been argued to be a source of grammatical representational and processing differences between HS aggregates and monolinguals and/or between individual HSs (see e.g. Montrul, 2022; Polinsky & Scontras, 2020; Kupisch & Rothman, 2018). In principle, reduced input, degree of variation in available input affecting cue validity, and/or lower opportunities for language use should have correlational and explanatory validity in the domain of predictive processing. And yet, comparatively few studies have examined the factors that affect predictive processing in HSs, much less those focusing on the individual level. The present study seeks to begin to fill this gap, showing how doing so both aids in better understanding HS processing, on the one hand, and, language processing, in particular prediction, more generally, on the other.

We approach this by examining both top-down and bottom-up processing in the HL of a group of Russian HSs in North America. Using a webcam-based eye tracking task, we investigate two domains: (i) prediction based on verb semantic constraint and (ii) lexical interference caused by incoming lexical information which is locally but not contextually (globally) coherent. Targeting the verb semantic constraint effect allows us to examine whether HSs of Russian use semantic information of a verb in a top-down manner, i.e., to generate predictions about the upcoming linguistic input. Exploring the lexical interference effect allows us to examine the interaction between top-down semantic and morphosyntactic (grammatical gender) cues and bottom-up lexical information. As we will explain in greater detail below, by juxtaposing these effects specifically with HSs we are able to test the relative contribution of various factors for predictive processing in general while simultaneously contributing to a deeper knowledge of individual variation in HS processing.

The effects of interest

Verb semantic constraint effect

Language processing occurs incrementally with comprehenders using incoming linguistic information to continuously update the set of potential continuations to include referents that meet the accumulating constraints (Altmann & Kamide, 1999). Arguably, the most significant type of constraints imposed by individual words are those linked to verbs which provide both semantic and syntactic restrictions on the kinds of complements or arguments they can take (Trueswell et al., 1993). Here, we focus on the effect of verb semantic constraints, i.e. how the meaning of a verb narrows down the set of possible arguments. For example, consider the verb *knit*, which implies an action of creating something by manipulating yarn (or other materials) with needles. Therefore, the types of objects that can appropriately follow the verb *knit* are those which can be made through knitting (e.g. sweater, scarf, mittens, etc.). It is still a rather large set of objects but the verb *knit* is more constraining in this way than the verb *lose*, for example, which can include all of the knittable objects and many more. Previous work shows that comprehenders do use this semantic information to anticipate what comes next in a sentence. This has been shown with native L1-dominant and L2 speakers (Dijkgraaf et al., 2017; Hopp, 2015; Kamide et al., 2003a, 2003b; Koehne & Crocker, 2015, among many others), however much less work in this domain has been done with HSs. The present study thus contributes to this sparse literature while taking a closer look at the role of individual factors in HSs' linguistic and cognitive abilities in predictive processing based on verb semantic constraints.

Lexical interference effect

While comprehenders use top-down information to restrict a set of potential continuations, incoming information which does not satisfy the sentential constraints is still considered, at least to some degree. Some previous work shows that comprehenders activate the contextually irrelevant meanings of ambiguous words (Swinney, 1979), consider locally but not globally consistent arguments for verbs (Peters et al., 2018), process referents for color words despite their

incongruence with the context (Kukona et al., 2014), to give a few examples. The degree of activation of such locally coherent but globally inconsistent information has been shown to vary as a function of language experience, among other factors. For example, Kukona et al. (2014) used a self-organizing neural network to simulate the color word effect whereby subjects considered a white car as a referent for the color word *white* presented as a part of a sentence *The boy will eat the white cake*. The results of their simulation suggested that such bottom-up activation is a residue of the learning process. First, the system learns to operate on bottom-up representations, and only once such representations are acquired, context-dependent effects emerge. Thus when studying prediction especially in non-dominant speakers of a language it becomes important to consider not only the degree of activation of the predicted entities, but also the degree of activation/inhibition of competing representations and factors that affect it. For example, does locally coherent but globally inconsistent information interfere less with top-down predictions as a function of the number and kind of cues used for prediction? And does this interference effect change as a function of language and cognitive abilities of comprehenders? Here, taking advantage of the grammatical gender system of the Russian language, we build on the work by Kukona et al. (2014) by introducing a morphosyntactic—grammatical gender—feature into the *eat this white cake* context (as expressed both at the demonstrative pronoun and the adjective) to examine whether this additional cue leads to the fading of the interference effect. Testing this effect in a population of speakers with varying degrees of attainment of grammatical gender also allows us to investigate it as a function of individual differences.

Interim summary

Previous literature strongly suggests that comprehenders use verb semantic constraint cues to identify a set of plausible arguments. Both of our experiments tap into this ability. Additionally, the lexical interference paradigm also taps into the ability to ignore incoming color word information when it applies to objects outside this set by relying on semantic and morphosyntactic (e.g., grammatical gender) cues. The lexical interference effect thus involves not only forming top-down predictions or understanding

bottom-up input, but rather the complex interaction between these processes. Juxtaposing these two effects makes HSs an ideal group for exploring how individual differences in language use and experience influence predictive processing. Specifically, we hypothesize that individual differences should not significantly impact the verb semantic constraint effect or do so significantly less as compared to the lexical interference effect, provided the subjects know the lexical meaning of the verbs themselves. Because such information is (more or less) conceptually universal, the reliability of this cue should be less vulnerable to input and engagement effects. In contrast, the interaction between top-down and bottom-up processing tested by the lexical interference effect is likely to be more sensitive to individual differences than simpler more isolated processes, like making predictions based solely on verb semantics. Specifically, speakers with higher levels of HL proficiency and maintenance are more likely to more successfully balance different types of linguistic information and inhibit locally coherent but globally inconsistent input (Kukona et al., 2014; Peters et al., 2018). Additionally, previous work on grammatical gender in Russian HSs in North America has shown it to be a particularly vulnerable domain, or at least one where there is considerable interspeaker variability (Polinsky, 2006, 2008). Such variability would be predicted to affect at least some HSs' determination of the reliability of grammatical gender as a viable predictive cue—e.g., those who have a simplified system and/or those whose exposure to others is characterized by significant degrees of variation in grammatical gender assignment/agreement.

Predictive processing in HSs

Verb semantic constraints

An emerging body of work has examined predictive processing in HLs across various domains. However, research specifically focusing on prediction based on verb semantic cues remains highly limited. To our knowledge, only Ito et al. (2023) have tested the verb semantic constraints in HSs. Using Visual World Paradigm (VWP), they tested anticipation in HSs in Germany by manipulating verb semantic constraints (choosing verbs that had the same or different constraints in Vietnamese and German) and classifier

constraints (specific classifiers are associated with certain prototypical semantic classes, functioning as a semantic cue). The study showed that HSs were using both verb semantics and classifier information to predict the target. When comparing them to the group of L1 Vietnamese-L2 German late bilinguals, the authors did find that HSs were slower in attending to the target when verb semantic constraints were different in Vietnamese and German than the late bilinguals, suggesting an influence from the societal dominant language German. Overall, this study suggests that HSs do make use of semantic information being encoded in verbs or classifiers, albeit with a delayed anticipation (compared to L1-dominant users) that may be caused by the influence of the societal dominant language.

While not focusing specifically on verbs, Hao et al. (2024) have also shown that HSs use semantic information predictively. In this study, Mandarin-speaking HSs were tested using the VWP to examine how they make use of Mandarin sortal classifiers to predict the upcoming nouns. Different from Vietnamese, Mandarin classifiers function both as a grammatical form class cue and a semantic cue. The results suggested that HSs, at the aggregated group level, relied primarily on grammatical information provided by classifiers to identify the upcoming target. However, an analysis of individual differences revealed that not all participants rely on grammatical form class cue. Instead, HSs with higher Mandarin Social Use scores (based primarily on questions related to informal literacy) and those having attended Mandarin language programs were more sensitive to semantic cues. The authors suggest that the effects of Mandarin social use and language program may be due to their impact on metalinguistic awareness, particularly given that the semantic association between classifiers and nouns is generally more robust than item-specific grammatical form class associations. The Hao et al.'s study highlights the importance of an individual differences approach, and the current study extends this approach to verb-based semantics prediction in HSs.

Another study which did not examine the verb semantic constraint per se but is still relevant for our purposes here is the one by Parshina et al. (2022) who examined lexical and morphosyntactic prediction in Russian HSs in the US and focused on individual differences. Targeting various types of Russian

grammatical structures (various word orders, active and passive voice, various types of relative clauses) in a reading paradigm, they found evidence for anticipation of specific lexical items and that this effect was modulated by the Russian literacy experience and English reading fluency. As for the morphosyntactic cues, the authors examined six different features: word class, noun gender, case and number, as well as verb tense and number. HSs showed evidence of anticipation of the word class as well as noun and verb number. Both the lexical and morphosyntactic effects were modulated by Russian literacy experience and reading fluency in English.

The findings from the studies reviewed above are relevant to the present research as they indicate that HSs do utilize semantic information for prediction, although how much they rely on it varies depending on their language background. Therefore, in the population tested herein we should expect there to be a discernible effect of the semantic properties of the verb, i.e., evidence of the verb semantic constraints. Moreover, our approach does not rely solely on aggregated trends, but rather seeks to unpack anticipated individual differences in HL processing. While inspired by previous work on HS individual differences, our selection of factors related to language use remains somewhat exploratory. These factors will be summarized below, after we introduce our second effect of interest.

Lexical interference effect

To our knowledge, no study has examined the lexical interference effect in HSs. However, previous research with other speaker populations offers valuable insights into the individual background factors that can predict participants' behavior. For example, Nozari et al. (2016) show that the effect of bottom-up interference varies as a function of domain-general inhibitory control measured in a Flanker task (Nozari et al., 2016), i.e. the more people are distracted by the incongruent markers in the Flanker task, the more they are susceptible to locally but not globally coherent information in a linguistic task. Thus, the non-linguistic Flanker task (see Eriksen & Eriksen, 1974; Eriksen, 1995; Hübner & Töbel, 2019) and the linguistic lexical interference task seem to tap into the same underlying construct of competition suppression. Along similar lines, Peters et al. (2018)

looked at how bilinguals activate representations which are compatible either with both the local and the global context or only with a local context. Using eye tracking, they showed that when hearing a sentence *The pirate chases the ship* and observing a display with a bone, a ship, a treasure box and a cat, participants with smaller vocabularies looked more at the locally but not globally coherent items (e.g. a cat who is chaseable but is unexpected in the context of a pirate). Summarizing across these studies, we predict that the interference effect will vary as a function of language experience, specifically, vocabulary size and usage patterns as well as inhibitory control ability in our sample of Russian HSs.

Other relevant work

It should be noted that in the context of our study, grammatical gender cues may function as additional markers, strengthening participants' commitment to their predictions and directing their attention away from distractors that are inconsistent with the sentential context. A few studies have examined whether HSs use grammatical gender as a cue for anticipating what comes next as language unfolds. For example, Sekerina (2015) presented their participants with 4 panels depicting a flying event with different objects and played a sentence describing it, e.g. *По небу летела серебристая ракета* -In sky was flying_{FEM-SG} silver_{FEM-SG-NOM} rocket_{FEM-SG-NOM}. The gender and the number of the target object could be identified upon hearing the verb in the past tense and the adjective preceding the object name. The authors manipulated the competitors which either matched or not the target object in grammatical gender/number. They found that participants used grammatical number and feminine (but not masculine) gender predictively.

Using eye tracking, Fuchs (2022a) investigated facilitative use of grammatical gender in on-line processing in HSs of Spanish. She recorded participants' eye movements while they were observing a visual scene with two objects with either matched or mismatched grammatical gender and answering the question *¿Dónde está [ARTICLE] [NOUN]?* (Where is [article] [noun])? In Spanish, grammatical gender is marked on the article and thus it is possible to identify the target in the gender mismatch condition already at the article. Fuchs found that when taking into account

only trials with the canonical gender assignment by HSs, their processing is qualitatively similar to that of dominant L1 Spanish users and they do identify the target object earlier in the mismatch relative to the match condition. Some subsequent work with HSs of Polish (Fuchs, 2022b) showed predictive use of grammatical gender information in this population of speakers as well.

Other psycholinguistic work with HSs has investigated whether they use other grammatical cues predictively. For example, both Karaca et al. (2024) and Özsoy et al. (2023) looked at the predictive use of grammatical case in HSs of Turkish in the Netherlands. The results are not straightforward. Karaca et al. (2024) found that their participants used grammatical case only in the verb-medial position, when grammatical cues were scaffolded by verb semantics, and not in the verb-final position when only morphosyntactic cues were available. Moreover, exposure to and use of Turkish as well as engagement with literacy² in both Turkish and the majority language (Dutch) predicted the strength of this effect. In Özsoy et al. (2023) two different approaches to statistical analyses resulted in two opposite patterns of results. But even the approach which supported the predictive use of case in this population did so only in the group of participants who took part in the lab-based experiment and not in the online experiment. A closer look at individuals across both testing modalities (in-lab and online) showed that they could be divided in three groups: those who reliably use case predictively, those who do not, and those who use case predictively at a chance level.

² Their findings support earlier work showing that literacy in general has an effect on spoken predictive processing skills (Huettig & Pickering, 2019; Mani & Huettig, 2014; Mishra et al., 2012). The cited work tested the effect of literacy (on language processing) focusing on adults with different levels of literacy skill and children with developing literacy skills in their native language, probing literacy skills according to the canonical definition of literacy—“an individual’s capacity to understand, use and reflect on written texts, in order to achieve one’s goals, to develop one’s knowledge and potential and to participate in society (Organization for Economic Co-operation and Development, 2006, p. 46, as cited in Keefe & Copeland, 2011). Psycholinguistic experimental work on HL bilingualism that has considered so-called literacy as a potential individual differences factor, however, has operationalized literacy in a different sense (e.g., Bayram et al., 2019; Karaca et al., 2024) whereby it refers primarily to degree of engagement with informal activities requiring readings and writing (e.g. texting; emailing) as well as amount of formal education in the HL.

While Özsoy et al. (2023) did not have any further language background or cognitive details about their subjects, they do acknowledge that further work should look into what factors underlie such variability.

In contrast to the studies reviewed above, Sekerina and Laurinavichyute (2020) report evidence that HSs of Russian failed to use case information predictively. They conducted an eye-tracking study to explore both real time processing and offline comprehension of two types of questions with different word orders in dominant L1 users and HSs of Russian: SOV subject Wh-questions with a scrambled word order (absent in English, e.g. WhoNOM girlACC kissed at school?) and OSV object Wh-questions (similar to English, e.g. WhoACC boyNOM kissed at school), following a context sentence disambiguating who did what to whom. Dominant L1 Russian users were able to predict the target already at the verb (kissed) while HSs avoided commitment to the target reference for both types of questions. Offline comprehension of these questions was equally good in both groups. Sekerina and Laurinavichyute (2020) speculated that such online behavior may be attributable either to insufficient online resources while calculating a short syntactic dependency or low confidence in processing in the HL. These speculations encourage follow-up work looking at individual variation in these domains and how it affects reliance on grammatical cues for predictive processing.

Summary of research questions

Following the emerging trend in the (heritage) bilingualism literature to investigate bilingualism as a spectrum of experiences (Rothman et al., 2023) rather than a categorical variable, the present study is asking the following research question: do prediction and lexical interference in HSs vary as a function of their individual differences in language experience, proficiency and cognitive skills? And what does the answer to this tell us about predictive processing in general? Inspired by the above-reviewed studies, the factors this work examines are vocabulary, cognitive control, knowledge of grammatical gender, engagement with literary-based activities in the HL and a multilingual engagement score. These factors are further discussed in “The receptive vocabulary task”—“The language history questionnaire” sections.

Methods

Participants

The resulting final dataset consisted of 68 participants. Their demographic information is reported in Table 1. We recruited HS participants either through personal communication or via contacting US or Canadian universities with heritage Russian programs, asking their instructors to share the study details with students. Given our main interests lie in describing and unpacking individual differences between HS bilinguals and in following with argumentation that this can be best done without direct comparison to L1 dominant users as an assumed baseline, our participant recruitment focused uniquely on HSs (e.g., De Houwer, 2023; Rothman et al., 2023). We defined HS inclusion for this study as individuals who (i) spoke Russian because it was a/the language spoken in their home, (ii) were born in the US or Canada or immigrated there before school age (iii) also had native fluency in English and were primarily

educated in English and (iv) were at least 18 years old. Seventy-three participants entered the study on Gorilla Experiment Builder (Anwyl-Irvine et al., 2019, 2021). The data from four people were excluded during the preprocessing stage due to low sampling rate (< 5 Hz, see section *Data preprocessing and analysis* for more details). The data from one additional participant were excluded because they did not specify Russian as one of their languages in the language history questionnaire.

Tasks

The eye tracking task

We used the same materials (and setup) previously reported in Prystauka et al. (2024), who successfully used web-based eye tracking to test anticipation and lexical interference in just over 200 Russian L1 dominant users. That study contained a discussion of methodological details related to online eye tracking, which we do not repeat here, that couples together

Table 1 Participants' demographic information

Age	M = 22.34, SD = 4.56, range—18 to 39
Gender	50 female, 17 male, 1 non-binary
Onset of exposure to Russian ^a	M = 0.71, SD = 1.29, range—0 to 7
Onset of exposure to English	M = 2.49, SD = 2.13, range—0 to 8
Country of origin	Belarus (N = 2) Ukraine (N = 6) Russia (N = 10) Georgia (N = 1) Uzbekistan (N = 1) Canada (N = 1) USA (N = 47)
Country of residence	USA (N = 62) USA & Canada (N = 1) Norway (N = 1) (born and raised in the US, moved to Europe at 24) The Netherlands (N = 1) (was born in the US, grew up speaking Russian and American English, started learning Dutch at 20) Poland (N = 1) (moved to Canada from Russia at 4, moved to Poland at 34) Ukraine (N = 1) (born and raised in the US, moved to Ukraine very recently) Israel (N = 1) (early exposure to Hebrew, 7 y.o.)

^aOnset of exposure to a language (Russian and English, in this case) was operationalized as the age at which participants first started learning the language in terms of listening

with an emerging methodological literature on web-based eye tracking (Slim & Hartsuiker, 2022; Vos et al., 2022) to serve as the basis for our confidence in the method to address the present research questions. The experimental procedure comprised two sets of materials (intermixed), both presented to the same participants within the same study. The materials for testing the semantic constraint effect consisted of sixteen stimulus sets. Each set included a quadrant-based visual display featuring four images (Fig. 1) and two minimal sentence pairs. These sentence pairs differed only in the choice of verb: constraining or non-constraining, determined by the verb's meaning in relation to the visual scene.

In the constraining condition, verbs allowed reference to only one of the four objects in the visual scene post-verbally. For example, *Женщина польёт растение* (The woman will water the plant). In the non-constraining condition, all four objects could be referred to post-verbally. For instance, *Женщина подвинет растение/вилку/пылесос/диван* (The woman will move the plant/fork/vacuum cleaner/couch). However, the target object was always the same as in the constraining condition.

To counterbalance, we created two stimulus lists, each containing eight sentences in each condition. Participants were presented with only one sentence from a minimal pair during the experiment.



Fig. 1 Example scene used to test the effect of verb semantic constraint. *Note.* Participants heard *Женщина польёт/подвинет растение* (The woman will water/move the plant)

Materials for testing the lexical interference effect comprised thirty-two sets of stimuli. Each set included a quadrant-based visual scene displaying four images (Fig. 2) and two minimal sentence pairs. These pairs differed only in the color adjective used. Visual scenes were composed of two pairs featuring distinct object types, with objects within each pair differing only in color—for instance, a scene might include both a black and a brown pipe and a black and a brown hat. In half of the stimuli pairs, the object types matched in grammatical gender, and in the other half they mismatched (this manipulation is explained in more details below).

All sentences in this set had constraining verbs, meaning only two of the four objects in a scene could appear in a post-verbal position. The target object could be identified upon hearing the color adjective. Each object in a scene represented one of four conditions in a 2 by 2 design, incorporating factors of verb consistency and color consistency. For example, in a scenario involving pipes and hats, a participant might hear the sentence *Дедушка выкурит эту черную трубку* (The grandfather will smoke this black pipe). In this context, the black pipe is consistent with both the verb's selectional restrictions and the color adjective, the brown pipe aligns only with the verb and not the adjective, the black hat aligns

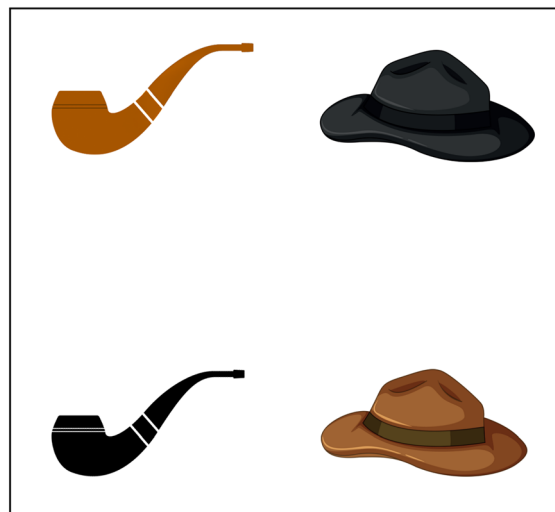


Fig. 2 Example scene used to test the effect of lexical interference. *Note.* Participants heard *Дедушка выкурит эту черную трубку* (The grandfather will smoke this black pipe) or *Дедушка выкурит эту коричневую трубку* (The grandfather will smoke this brown pipe) (color figure online)

with the adjective but not the verb, and the brown hat does not align with either the adjective or the verb.

To address potential issues related to color saliency, half of the participants were presented with sentences featuring one color as the target, while the other half heard sentences featuring the alternative color as the target.

There was also an additional factor of grammatical gender in this subset of stimuli, such that the target and the distractor either matched ($N = 16$) or not ($N = 16$) in grammatical gender. There are three grammatical genders in Russian: masculine (46% of the nominal lexicon), feminine (41% of the nominal lexicon) and neuter (13% of the nominal lexicon). Variation in the resulting gender systems of Russian in HS contexts has been reported, especially in the US. Overall, Russian HSs in general use phonological cues, leading to a straightforward categorization of nouns based on their endings: consonant endings are associated with the masculine gender for all, while vowel endings can be categorized in two ways, resulting in a two-gender system for some (feminine being assigned to all nouns ending in a vowel). Still other HSs distinguish between neuter nouns (those ending in a stressed $-o$, thus easily differentiable from the rest of the vowel-final nouns) and all other nouns, which are interpreted as feminine (Polinsky, 2008). With this in mind, we only used masculine nouns ending in consonants and feminine nouns ending in vowels (not with stressed $-o$). Within each condition (i.e. match and mismatch), an equal number of targets ($N = 8$) were of masculine or feminine grammatical gender. Grammatical gender in Russian is expressed on parts of speech agreeing with the noun, which in this design—not least as Russian does not otherwise have overt articles—included a demonstrative pronoun *this* ($\text{ЭТО}_{\text{masc}} / \text{ЭТА}_{\text{fem}}$) and an adjective. The prediction this study is testing is whether in the mismatch condition subjects will use an additional cue (grammatical gender) available already at the demonstrative pronoun to disregard the contextually irrelevant distractor despite its consistency with the color word, resulting in fewer looks to this distractor.

A female native speaker of Russian (a professional voice actress) was instructed to produce the sentences naturally. The stimuli were recorded in a sound-attenuated booth and edited using the Audacity software. Visual stimuli were created with the images

sourced from the ClipArt collection (clipart.com). All materials are provided in [Appendix](#).

The receptive vocabulary task

To assess participants' receptive vocabulary in Russian we used an adapted version of the KORABLIK test—Clinical Assessment of the Development of Basic Linguistic Competencies (Lopukhina et al., 2019). Specifically, we used a subset of the test designed to assess comprehension of nouns and verbs. On each trial, participants were presented with 4 pictures and an auditory word. Their task was to choose the matching picture. There were 4 practice trials and 48 experimental trials, evenly split between nouns and verbs. Given previous work, we hypothesized that vocabulary score would be an important factor for both the verb semantic constraint and lexical interference effects (Kukona et al., 2014; Peters et al., 2018).

The Flanker task

To assess inhibitory control, we used the Flanker task (Eriksen & Eriksen, 1974). Participants were presented with sets of five arrows or an arrow surrounded by two dashes on each side. Their task was to specify the direction of the arrow in the middle. Surrounding arrows (flankers) were pointing either in the same direction (congruent condition, <<<<<<) or in a different direction (incongruent condition, <<<<<>). Trials with an arrow and dashes represented the neutral condition. Incongruent trials require participants to inhibit the conflicting information coming from the surrounding arrows, thus leading to greater recruitment of the cognitive control system. The task consisted of 10 practice and 144 experimental trials evenly split into three conditions.

We primarily included this task to examine the link between inhibitory control and the lexical interference effect (Nozari et al., 2016). However, previous literature suggests a relationship between predictive behavior in language processing and different aspects of cognitive control (Noh & Lee, 2017; Zirnstein et al., 2018). That is why we also explored the role of cognitive control for our verb semantic constraint effect.

The gender knowledge task

The gender knowledge task consisted of the subset ($N = 32$) of items from the lexical interference condition. Participants were presented with an image of an object such as the black pipe, and four auditory descriptions of the target object. These descriptions varied in the adjective which was manipulated based on its gender agreement and color consistency with the noun. Using the black pipe as an example, participants heard *черная трубка* (black_{fem} pipe), *белая трубка* (white_{fem} pipe), *черный трубка* (black_{masc} pipe), *белый трубка* (white_{masc} pipe). Their task was to click on the corresponding description.

We included this task to explore whether participants' knowledge of gender assignments for the items in the study interacts with their ability to use grammatical gender predictively, which would provide them with an additional cue to inhibit contextually irrelevant items.

The language history questionnaire

We used the LHQ3 version of the Language History Questionnaire (Li et al., 2020) which includes a comprehensive set of questions for assessing the linguistic background of multilingual speakers. It also allows for the calculation of multiple aggregated scores such as proficiency, dominance, immersion and multilingual language diversity (MLD). The proficiency score is derived from self-rated proficiency across listening, speaking, reading, and writing components. Immersion quantifies the amount of time speakers were immersed in a particular language. Dominance integrates self-reported proficiency and the time spent on various language components. Lastly, MLD estimates the balance in the use of up to four languages, considering both dominance and frequency (and is based on the language entropy construct discussed in Gullifer & Titone, 2020). Moreover, during the calculation of the aggregated scores, researchers can assign different weights to the modules, for example it is possible to estimate proficiency in writing and reading (i.e. skills that require literacy activities) separately from speaking and listening. To test our research questions, we focused on self-rated proficiency in reading and writing in Russian as a proxy for engagement with literacy in the HL, as well as the MLD score.

Regarding proficiency, we adjusted the calculation of this score to include an estimation of engagement with reading and writing. This modification was made as HSs tend to exhibit significant variation in this domain, and previous research has demonstrated its predictive value in language attainment (Bayram et al., 2019) and predictive processing (Mani & Huettig, 2014; Mishra et al., 2012). Regarding MLD, higher scores on this scale indicate more balanced—integrated—use of languages whereby all languages are used across different contexts and language mixing is a common phenomenon. Lower scores indicate more compartmentalized language use whereby one language is associated with one context. MLD calculation is based on language dominance which is in turn based on self-assessed proficiency and language use estimations in different contexts. It is thus a comprehensive measure reflective of individual differences in current language experience.

Procedure

The experiment was programmed in the Gorilla Experiment Platform (Anwyl-Irvine et al., 2019). Personal computers were set as the only allowed device type. The study started with a video instruction explaining the purpose and the general procedure of the experiment, following which the participants were directed to the consent form. Upon providing their consent, participants were redirected to the eye tracking task, which started with more specific video instructions, example trials and the first webcam calibration procedure. Participants who successfully finished calibration proceeded to the eye tracking experiment which consisted of 48 trials (16 for the semantic constraint effect, 32 for the lexical interference effect, intermixed), split into three blocks, divided by two additional calibration routines (i.e., a new calibration occurred after every 16 trials). Each trial started with a fixation cross and progressed to a visual display once the participants clicked on it. After a preview time of 1000 ms, the audio was played (the actual audio onset time varied somewhat among participants, which is further discussed in the Data Preprocessing section). Participants were instructed to press on the object mentioned in the sentence after the sentence offset, with the button press activated after the audio offset. Following the eye tracking task, participants also performed a vocabulary task,

presented in one block, the Flanker task across three blocks and the grammatical gender task presented in a single block. Afterwards, they were redirected to Qualtrics (Qualtrics 2015) where they filled out the Language History Questionnaire. As the study targeted HSs whose reading and writing abilities may vary significantly, all instructions were auditory and all responses were programmed as button clicks (rather than typing). All task instructions were in Russian. The Language History Questionnaire required written responses and was in English.

Data preprocessing

We followed the same preprocessing as in Prystauka et al. (2024). Gorilla provides two output data quality metrics for the eye tracking data. One of them is the mean convergence value (“convergence”) for fitting a facial model. This represents the model’s confidence in finding a face (and accurately predicting eye-movements), with values varying from 0 to 1. Values below 0.5 suggest a likely convergence of the model. Another metric is “face_conf”, which represents the Support Vector Machine (SVM) classifier score for the face model fit. This score indicates the degree to which the image under the model resembles a face, with values ranging from 0 to 1. Numbers exceeding 0.5 indicate a well-fitted model. In our sample, all “convergence” values were 0, and approximately 0.06% of data points with “face_conf” values below 0.5 were excluded.

The number of incorrect responses (failure to click on the right object) was 1.5%. They were excluded from the data.

Sampling rate of the original sample ($N = 73$) varied from 1.7 to 29.9 Hz (see Fig. 3 for the distribution of sampling rate in our sample). We excluded participants with less than 5 samples per second, which resulted in the exclusion of 4 individuals and an additional participant based on their responses on the LHQ3. The mean sampling rate in the resulting group was 23.2 Hz ($SD = 4.85$ Hz, range 10.9–29.9 Hz).

Additionally, the experiment design included an image preview time of 1000 ms, although the *actual* onset time of the sentence varied between participants due to the properties of their hardware and connection speed. Gorilla offers the option to download additional metrics on audio events, including the timing of when

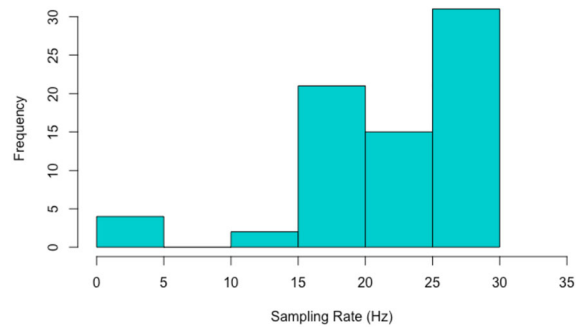


Fig. 3 The distribution of participants’ sampling rate

the audio event actually started (as opposed to when it was requested). The range of actual onset times spanned between 1001 and 3977 ms with only five trials exhibiting onset times above 2000 ms across different participants. The mean onset was 1055 ms, with a standard deviation of 77 ms. The distribution of onset times is illustrated in Fig. 4. Our analysis accounted for these variations in the onset of the audio.

Analysis of the eye tracking task

We defined our regions of interest (ROI) as quadrants, with the critical ROI being the one containing the target image (i. e., the object mentioned in the sentence) for the segment of the study testing the verb semantic constraint effect. For the lexical interference effect the critical ROIs were the ones containing the distractor items (i.e. items of the other semantic type than the target).

To analyze our data, we employed (generalized) linear mixed effects modeling, given its ability to test for condition effects while accounting for random

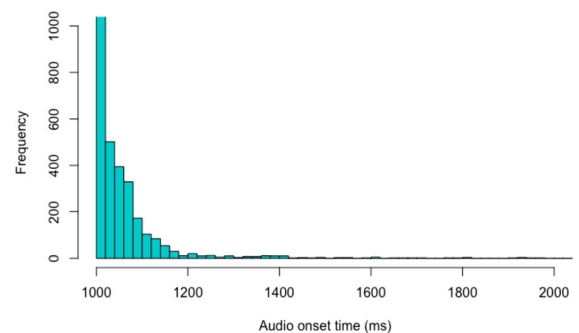


Fig. 4 The distribution of audio onset times among participants

effects such as trials and/or subjects. Different dependent variables were specified for the two effects of interest.

Verb semantic constraint effect (Altmann & Kamide, 1999)

First, we defined our time window of interest from 200 ms after the verb onset and until 200 ms after the noun onset (since it takes roughly 200 ms to plan and perform a saccade, Matin et al., 1993). We coded looks to the target binomially (1 = fixated if there was at least one look to the target, 0 = not fixated) for 100 ms time bins within our temporal window of interest and ran a mixed-effects logistic regression model (GLMM) using the `glmer` function from the `lme4` package (Bates et al., 2015), which is commonly used for binary outcomes or proportions. The following fixed factors were used: condition (constraining vs non-constraining verb), mean vocabulary score (i.e. mean accuracy on the vocabulary task), Flanker effect size (defined as the reaction time difference between the incongruent and congruent trials), self-rated proficiency in writing and reading (a proxy for engagement with literacy) in Russian and MLD (based on the LHQ3 questionnaire). Our primary hypotheses revolved around the interaction of individual background variables with the verb condition, for this reason we included the interaction term between these variables and the condition variable. Specifically, when the verb is constraining and enables predictions about the upcoming target, various language background factors are expected to predict the likelihood of a speaker to act on this information, i.e., make predictions. The continuous variables were centered and scaled. VIF values were calculated for the predictor variables, ranging from 1 to 1.4. These low VIF values indicate that multicollinearity is not a significant concern in the model. Participants and trials were entered as random factors. The model included the following fixed and random effects:

Lexical interference effect (Kukona et al., 2014)

Given that both distractors are present in the picture, the looks to one entail not looking at the other. That is why we computed the advantage score defined as the log ratio of the proportion of looks to the distractor of the same color to the proportion of looks to the distractor of a different color in our time window of interest including 0.5 in both the numerator and the denominator to prevent computation errors in cases where the denominator is zero (Ito & Knoeferle, 2022). As previous studies (Kukona et al., 2014; Prystauka et al., 2024) as well as the data from this experiment suggest, the interference effect is most pronounced on the noun, that is why we defined our window of interest as 1000 ms post noun onset. Using a linear mixed-effects model, we then examined the advantage score as a function of our factors of interest. The following fixed factors were used: condition (gender match/gender mismatch), Flanker effect size (defined as the reaction time difference between the congruent and incongruent trials), mean accuracy score from the gender judgment task, mean accuracy score from the vocabulary task, self-rated proficiency in writing and reading in Russian and MLD, as well as the interaction between gender (mis)match and gender score. Our dependent variable, the advantage score, accounted for the difference in looks to the distractors of different colors, thus detecting an effect of any individual difference variable would entail the relationship between that variable and the difference in looks to the distractors of different colors. We incorporated only one interaction factor, specifically the interplay between gender manipulation and performance on the Gender Knowledge task. This decision was based on our prediction that participants who performed better on a gender task would be more inclined to use gender information as an additional cue to suppress contextually irrelevant distractors. We chose not to include interaction terms between the gender manipulation and other individual difference

*Model < – glmer(Looks to the Target ~ Condition * (Mean Vocabulary Score + Flanker Difference RTs + Self – Rated Proficiency in Writing and Reading in Russian + Multilingual Language Diversity) + (1 | Trial) + (1 | Participant))*

variables due to a lack of strong theoretical grounds to anticipate a significant relationship between these variables and different levels of a gender condition.

The continuous variables were centered and scaled. VIF values were calculated for the predictor variables, ranging from 1 to 1.9. These low VIF values indicate that multicollinearity is not a significant concern in the model. Participants and trials were used as random factors. The resulting model was as follows:

$$\begin{aligned} \text{Model} < - \text{lmer}(\text{Log Ratio} \sim \text{Gender (Mis)Match} * \text{Gender Score} + \text{Vocabulary Score} \\ + \text{Self} - \text{Rated Proficiency in Writing and Reading in Russian} + \text{Multilingual Language Diversity} \\ + \text{Flanker Difference RTs} + (1|\text{Trial}) + (1|\text{Participant})) \end{aligned}$$

Results

Gender knowledge and vocabulary tasks

In the gender knowledge task, accuracy varied from 66 to 100%, with an average of 97% (SD = 6.7%). Accuracy on the vocabulary task varied from 62.5 to 100%, with an average of 91% (SD = 8.8%).

Flanker task

Four people did not provide their responses on the Flanker task, thus the data below is reported for the 64 participants who did. Accuracy in the congruent condition varied from 86 to 100%, with an average of 99% (SD = 2.2%). Accuracy in the incongruent condition varied from 46 to 100%, with an average of 96% (SD = 8.1%). Reaction times in the congruent condition varied from 337.3 ms to 574.7 ms, with an average of 444.6 ms (SD = 56.4 ms). Reaction times in the incongruent condition varied from 385.4 ms to 619.9 ms, with an average of 488 ms (SD = 56.6 ms).

To address the issue of missing data,³ we utilized the Multiple Imputation by Chained Equations (MICE) method in R (van Buuren & Groothuis-Oudshoorn, 2011). The imputation process was carried out on a part of a dataset that contained average values of other

predictor variables (accuracy on the Vocabulary and Gender knowledge tasks, Proficiency in Reading and Writing in Russian and MLD) per participant. The imputed values followed the distribution of the observed data, suggesting that they could be plausible measurements had they not been missing.

LHQ3

Summary of the relevant LHQ3 based composite scores is provided in Table 2.

Verb semantic constraint effect (Altmann & Kamide, 1999)

The time course of the proportions of looks to the target is illustrated in Fig. 5.

The analysis revealed a significant main effect of condition (with more looks to the target following constraining verbs) and significant interaction between condition and Flanker effect. The results are summarized in Table 3 and plotted in Fig. 6. The interaction suggests that the influence of the Flanker effect on the allocation of visual attention to the target is moderated by the verb constraint. Specifically, a constraining verb context seems to lead to a decreased likelihood of looking at the target with an increasing Flanker effect, while a non-constraining verb context shows the opposite pattern.

Lexical interference effect (Kukona et al., 2014)

The time course of the proportions of looks to all items in a scene is illustrated in Fig. 7.

The analysis revealed a main effect of Vocabulary score (Table 4 and Fig. 8), whereby participants with a higher Vocabulary score exhibited a smaller difference between looks to the distractors of the same and

³ We also re-ran our models on the 64 participants with complete datasets and obtained the same pattern of results.

Table 2 Summary of LHQ3 based composite scores

	Self-rated proficiency in reading and writing in Russian	MLD (varies from 0 to 2)
Mean (SD)	0.58 (0.19)	1.37 (0.31)
Min	0.21	0.78
Max	1	1.97

different colors indicating reduced susceptibility to lexical interference.

Discussion

This study investigated the interplay between language processing, visual attention, inhibitory control and language use in HSS of Russian in North America. We looked at two specific phenomena: (i) the

Fig. 5 The average proportions of looks to the target across time. *Note.* A. Error ribbons represent standard error. Vertical lines indicate the time window between 200 ms after the verb onset and 200 ms after the noun onset

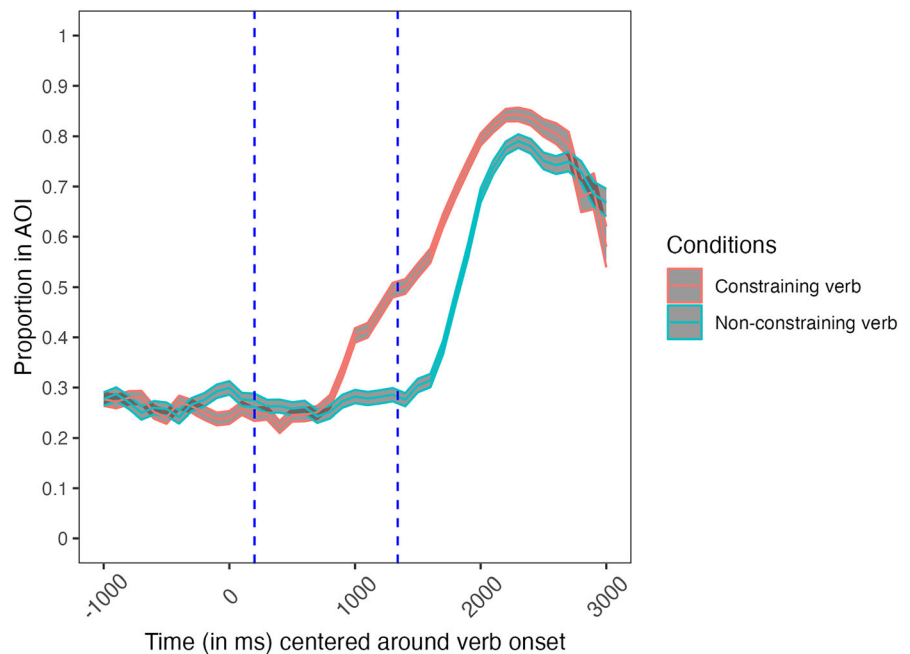


Table 3 Summary of the model to test the verb semantic constraint effect

Predictors	Fix target		
	Odds Ratios	CI	<i>p</i>
(Intercept)	0.38	0.32–0.46	< 0.001
Verb [constraining]	1.28	1.19–1.39	< 0.001
Vocabulary	0.96	0.88–1.04	0.322
FlankerRTdiff	1.08	0.99–1.18	0.073
Proficiency Rus lit	1.01	0.93–1.09	0.886
MLD	1.04	0.96–1.13	0.324
Verb [constraining] * Vocabulary	1.06	0.98–1.16	0.158
Verb [constraining] * FlankerRTdiff	0.89	0.81–0.96	0.005
Verb [constraining] * Proficiency Rus lit	1.08	0.99–1.17	0.071
Verb [constraining] * MLD	0.94	0.87–1.02	0.122

p-values less than 0.05 were considered statistically significant

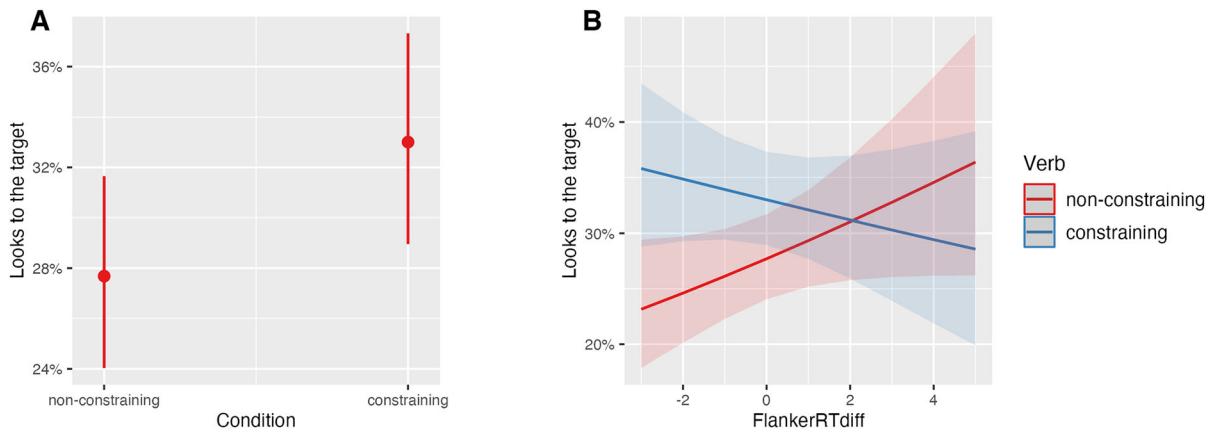


Fig. 6 Predicted values illustrating the **a** main effect of Condition, **b** significant interaction between Condition and Flanker effect (RT difference between the incongruent and congruent conditions). *Note.* The post-hoc tests on the Flanker

effect indicate that the significant interaction is primarily driven by participants with a smaller Flanker effect, who exhibited a greater difference in the proportion of predictive looks to the target between the constraining and non-constraining conditions

anticipation driven by verb semantic constraints and (ii) the presence of lexical interference. Our approach focused on modeling these as potential functions of individual differences, including vocabulary size, knowledge of grammatical gender, cognitive control, self-rated proficiency in reading and writing in Russian, and language entropy (MLD). Crucially, we consider how doing so informs current discussion on predictive processing more generally.

The basic findings of this study suggest that the present Russian HSs utilized verb semantics to anticipate what comes next; they also experienced lexical interference from adjectives that were color word-consistent but verb-inconsistent, similar to the results observed in the original studies with monolinguals (Altmann & Kamide, 1999; Kukona et al., 2014) and adult L2 learners (Dijkgraaf et al., 2017). Detecting the basic effects of our experimental manipulations in this population of HSs is a valuable finding in itself. While a sizeable literature investigating HSs' grammars in different language combinations exists (see Montrul, 2022; Polinsky, 2018; Montrul & Polinsky, 2021; Kupisch & Rothman, 2018 for review), the majority of available data come from offline behavioral methods. Recently, there have been calls to expand the experimental methods used with HSs to include the same variety of psycholinguistic experimentation used with L1-dominant speakers and L2 learners, especially online ones (Bayram et al., 2021; Polinsky & Scontras, 2020). Indeed, online (real time) psycholinguistic measures provide a more

automatic way of estimating individuals' linguistic knowledge. As such, their increased use might help overcome issues related to the unfamiliarity of HSs with the use of their HL in a formal context, which, in turn, might affect their metalinguistic judgements (Bayram et al., 2021; Montrul et al., 2014; Montrul, 2022). Our findings thus contribute to a growing literature on language processing in HSs and underscore the importance of utilizing a broader array of experimental techniques to fully capture the nuances of their linguistic abilities. Moreover, the fact that these real time processing effects were detected using online (i.e. web-based, as opposed to lab-based) eye tracking adds even more to the literature, highlighting the efficacy and utility of this methodology in general and with HSs in particular (Degen et al., 2021; Slim & Hartsuiker, 2022; Vos et al., 2022; Prystauka et al., 2024). This approach is particularly advantageous for studying populations such as HSs, who are often geographically dispersed and difficult to access for testing in a single physical location.

Perhaps a more significant contribution of the present paper concerns the individual differences results. The effect of anticipation based on verb semantic constraint was modulated by participants' performance on a Flanker task. The measure we chose for the Flanker effect is the difference in reaction times between the incongruent and congruent trials which measures the additional time required to process and respond to incongruent trials relative to congruent ones. This measure is a commonly used indicator of

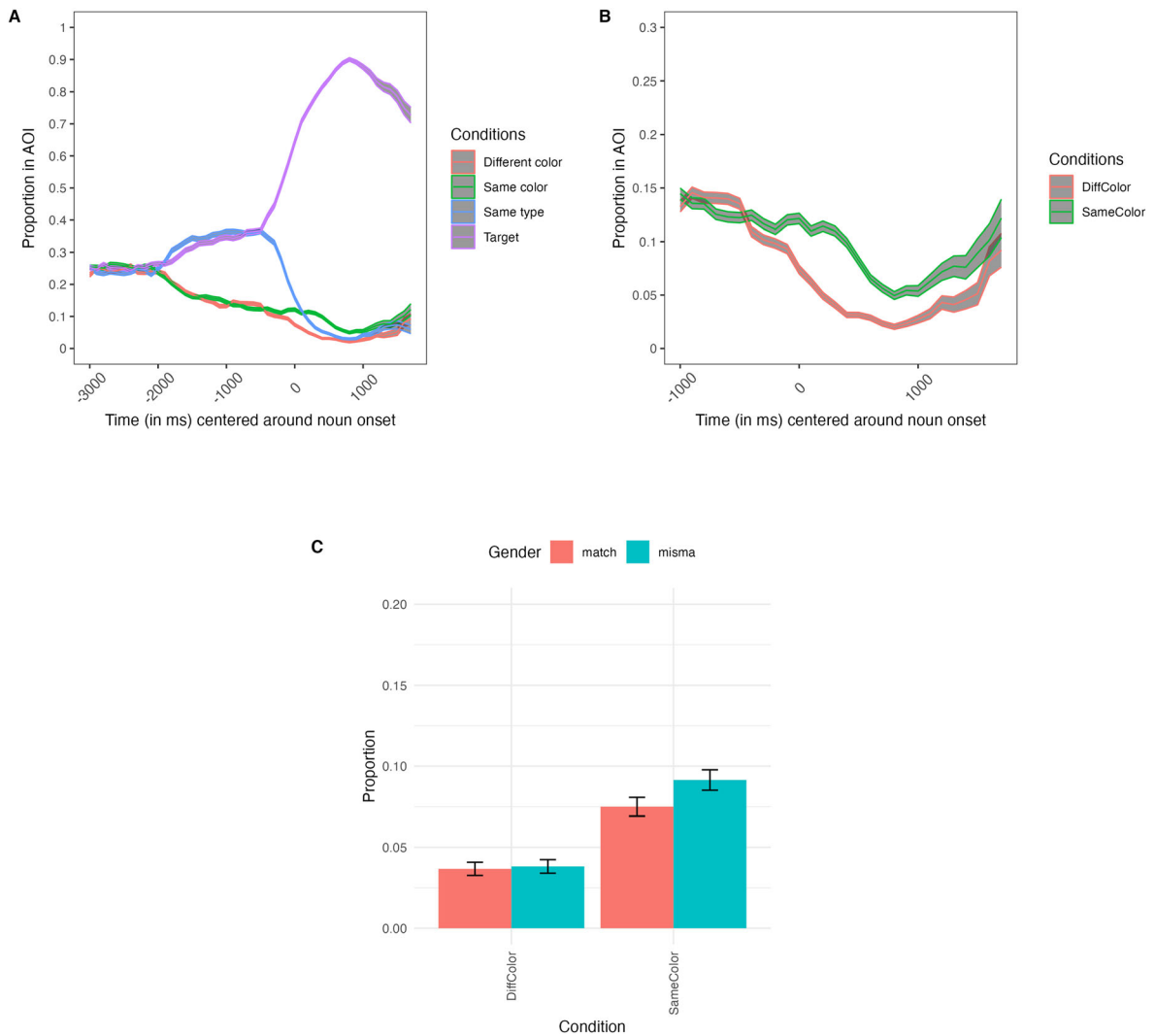


Fig. 7 The average proportions of looks to all objects in a display (a); to the distractors in a zoomed-in window (b) and to the distractors broken down by the gender manipulation (c)

Table 4 Summary of the model to test the lexical interference effect

Predictors	Log prop noun		
	Estimates	CI	<i>p</i>
(Intercept)	0.0119	0.0044–0.0194	0.002
GenderCond [mismatch]	0.0041	– 0.0064 to 0.0145	0.447
GenderTask	0.0004	– 0.0034 to 0.0042	0.847
Vocabulary	– 0.0052	– 0.0081 to – 0.0022	0.001
Proficiency Rus lit	– 0.0022	– 0.0050 to 0.0007	0.134
MLD	– 0.0002	– 0.0030 to 0.0027	0.898
FlankerRTdiff	0.0016	– 0.0013 to 0.0046	0.285
GenderCond [mismatch] * GenderTask	– 0.0007	– 0.0056 to 0.0042	0.768

p-values less than 0.05 were considered statistically significant

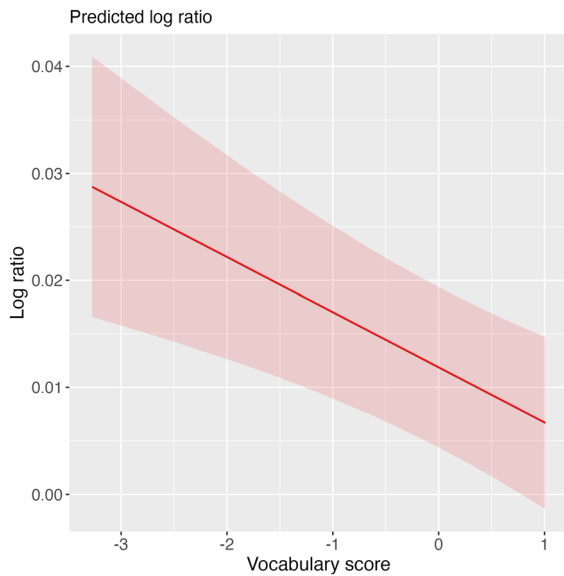


Fig. 8 Predicted values illustrating the main effect of Vocabulary Score and Flanker effect size

the efficiency of cognitive processes involved in handling interference. Our results suggest that in the constraining condition, looks to the target increased with a decreasing Flanker effect while in the non-constraining condition, looks to the target decreased as a function of a decreased Flanker effect. In other words, those who managed interference in the Flanker task more adeptly (indicated by a lower reaction time cost) were also better at leveraging semantic constraint cues, ignoring distractions, and focusing on the target. Assuming that larger Flanker effect size reflects reduced inhibitory control, our findings suggest that attending to the target is not just a function of linguistic context but also relates to an individual participant's ability to effectively and strategically allocate attention in situations which afford that. There is a sizeable body of work looking at the relationship between language processing and domain-general cognitive control, particularly in the bilingualism field where researchers investigate control processes involved in managing multiple languages (Abutalebi & Green, 2007; Bialystok, 2024; Bialystok & Craik, 2022; Luk et al., 2012) or in studies with monolinguals investigating linguistic stimuli with some sort of conflict (e.g. ambiguity or violations at different levels, Hsu et al., 2021; Brown-Schmidt, 2009; Novick et al., 2014; for a review, see Ness et al., 2023; Patra et al., 2023).

The literature on the role of domain-general cognitive abilities for predictive processing is more sparse, yet there is previous work highlighting different aspects of this relationship in different types of speakers. Noh and Lee (2017) tested the effect of verb semantic constraint utilizing a VWP design (similar to the one we report here) in a group of Korean L2 learners of English. They found that accuracy on a Flanker task correlated with anticipatory processing—a result very similar to the one reported here, yet we detected this relationship using an RT measure. Zirnstein and colleagues conducted a series of EEG experiments exploring a relationship between cognitive control and predictive processing in native and L2 speakers (Zirnstein et al., 2018), in older mono- and bilingual adults (Zirnstein et al., 2019) and in heritage and L2 speakers under noise conditions (Fricke & Zirnstein, 2022). Our findings extend this (and other Covey et al., 2024; Dave et al., 2021) work by providing evidence that cognitive control modulates predictive processing in a group of HSs more generally. Our participants listened to sentences which were neither ambiguous nor contained violations (as opposed to more complex/nuanced designs utilized in the above mentioned studies), and yet their inhibitory control abilities still predicted how much they rely on the semantic information of the verb to anticipate how the language will unfold constrained by the four images in a visual scene. It could be that this relationship was detected in our study because processing HLs is (often) more demanding and, thus, requires increased reliance on domain general control even when there is no violation/ambiguity. If on the right track, this would underline a crucial role of cognitive control not only in navigating linguistic complexity per se, but also in enhancing efficiency of language processing more generally. This perhaps is most readily detectable in populations like HSs because they are a type of native, naturalistic speaker who has had significantly less exposure to and opportunity to use the (heritage) language.

Of course, HSs vary at the individual level in terms of the quantity and quality of their language exposure and usage. It, thus, follows that the magnitude of such an effect would calibrate to an individuals' measurements that proxy for these relevant differences. In other words, it is not to be taken for granted that (compensatory) processing manifestations of reduced input/intake and usage patterns will manifest to the

same degree for all HSs, if at all for some. This bears out in the individual differences approach to the present data set and is in line with an emerging body of work focusing on the added value of individual differences approaches to bilingualism research (Rothman et al., 2023). As such, the present data set provides novel insights into the processing of HLs, on the one hand, and the relationship between domain-general cognition and language processing more generally, on the other (Luk & Rothman, 2022).

Our finding from testing the interference effect suggested that participants with higher vocabulary scores were less subject to lexical interference. This vocabulary finding is in line with our predictions, previous computational modeling, and experimental work (Kukona et al., 2014; Peters et al., 2018). This suggests that participants with larger lexicons are more likely to efficiently extract relevant information, i.e. from the constraining verb in our experiment, and inhibit the contextually irrelevant distractor despite it being locally coherent with the incoming lexical information. Vocabulary has long been used as a proxy for language proficiency (Treffers-Daller & Silva-Corvalán, 2016; Lemhöfer & Broersma, 2012), and our results add to the literature showing how knowing more words can affect sentence processing strategies and allocation of attention. Given that the domain we are investigating is highly dependent on lexically based processing, it stands to reason that vocabulary size would be a primary predictive factor for individual differences. As such, it should not be interpreted that vocabulary size is a holy grail of sorts to explain HS individual differences *par excellence*, but rather that it is a good candidate as one such variable for processing that is highly based on lexical constraints.

Returning to our predictions regarding the juxtaposition of the two effects, as anticipated, the factor that emerged as crucial for the verb semantics effect was not tied to language use and exposure, but instead was linked to participants' domain-general cognitive abilities, specifically cognitive control. This finding underscores the idea that verb semantic constraints, being conceptually universal, are robust across varying levels of language experience and are less susceptible to the effects of reduced input or engagement. On the other hand, the factor that significantly influenced the interference effect was indeed related to participants' language abilities, reinforcing the notion that this effect is more sensitive to individual

differences in language use and experience. These results contribute to the broader literature on predictive processing by highlighting the differential impact of cognitive and linguistic factors on the use of predictive cues (Huettig & Janse, 2016; Kaan & Grüter, 2021; Kuperberg & Jaeger, 2016).

One rather surprising finding was that our gender manipulation did not have any effect on participants' attention. There can be two explanations for this. First, it could be that HSs (or at least participants in our sample) do not use gender information predictively as an additional cue to rule out contextually irrelevant information. Second, there could be a methodological caveat which prevented us from detecting this relationship: the lexical interference effect is rather small (around 5%), and thus does not leave much room for the exploration of individual differences. Follow-up work utilizing more sensitive designs with expected larger effect sizes should further investigate this relationship in HSs. For example, one line of follow-up work could look at whether HSs of Russian use gender predictively and, if so, under what specific linguistic conditions and/or under what individual differences pertaining to their personal language experiences. Doing so, would help to align this and related work that capitalizes on processing of grammatical gender as a potential secondary cue with the literature that focuses primarily on gender acquisition/processing itself in HLs in general and Russian specifically (Fuchs et al., 2022a; 2022b; Di Pisa et al., 2022; Mitrofanova et al., 2018; Mitrofanova et al., 2022; Polinsky, 2008; Sekerina et al., 2006; Sekerina, 2012, 2015).

Conclusions

In summary, this study examining 68 HSs of Russian in North America, showed that participants exhibited anticipatory processing and were susceptible to lexical interference. The analysis of individual differences underscored the non-uniform nature of these effects. Vocabulary and domain-general cognitive control emerged as important factors influencing lexical interference and prediction, respectively, among Russian HSs. These findings contribute to a deeper understanding of the interplay between cognitive mechanisms and language processing in HSs,

shedding light on the multifaceted nature of bilingual experience.

Author contributions YP and JR were primarily responsible for writing the main manuscript text, while RCP played a key role in participant recruitment and JH contributed to the statistical analysis. All authors actively participated in the manuscript review process.

Funding Open access funding provided by University of Bergen (incl Haukeland University Hospital). Yanina Prystauka, Jiuzhou Hao and Jason Rothman were funded by the AcqVA Auora Center grant and the Tromsø Forskningsstiftelse Grant No. A43484: the Heritage-bilingual Linguistic Proficiency in their Native Grammar (HeLPiNG) (2019–2023).

Data availability The data, visual materials, and analysis scripts are available at <https://osf.io/rceqf/>. The study was not preregistered.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

This appendix contains materials used in the eye tracking experiment.

Stimuli for testing the verb semantic constraint effect.

1. Девочка почистит/возьмет банан.
The girl will peel/take the banana.
2. Женщина выжмет/купит лимон.
The woman will squeeze/buy the lemon.
3. Парень погасит/нарисует сигарету.
The young man will put out/draw the cigarette.

4. Доктор пропишет/подготовит таблетки.
The doctor will prescribe/prepare the pills.
5. Мужчина расчешет/рассмотрит бороду.
The man will comb/inspect the beard.
6. Женщина растопит/отдаст шоколадку.
The woman will melt/give away the chocolate bar.
7. Путешественник разведет/сфотографирует костер.
The hiker will start/photograph the fire.
8. Женщина польет/подвинет растение.
The woman will water/move the plant.
9. Девушка застелит/обойдет кровать.
The young woman will make/walk around the bed.
10. Парень расстегнет/продаст пальто.
The guy will unzip/sell the coat.
11. Бабушка распутает/спрячет клубок.
The grandmother will untangle the ball of yarn.
12. Мужчина пролистает/одолжит газету.
The man will flip through/borrow the newspaper.
13. Девочка подпишет/выберет открытку.
The girl will sign/choose the postcard.
14. Мужчина срубит/увидит дерево.
The man will cut/see the tree.
15. Мужчина построит/потрогает стену.
The man will build/touch the wall.
16. Женщина пришьет/выбросит пуговицу.
The woman will sew on/throw away the button.

Stimuli for testing the Lexical Interference effect.

1. Мальчик словит этот белый/розовый мяч.
The boy will catch this white/pink ball.
2. Студент соберет этот красный/желтый рюкзак.
The student will pack this red/yellow backpack.
3. Женщина выпьет этот желтый/белый напиток.
The woman will drink this yellow/white beverage.
4. Женщина порежет этот зеленый/оранжевый перец.
The woman will chop this green/orange bell pepper.
5. Бабушка свяжет этот красный/серый свитер.
The grandmother will knit this red/grey sweater.

6. Девочка развернет этот зеленый/красный подарок.
The girl will unwrap this green/red gift.
7. Мужчина зашнурует этот коричневый/серый ботинок.
The man will lace this brown/grey boot.
8. Женщина сорвет этот розовый/желтый цветок.
The woman will pluck this pink/yellow flower.
9. Женщина выгуляет эту белую/коричневую собаку.
The woman will walk this white/brown dog.
10. Мужчина заведет эту красную/синюю лодку.
The man will start this red/blue boat.
11. Женщина прочитает эту синюю/коричневую книгу.
The woman will read this blue/brown book.
12. Мужчина пожарит эту серую/красную рыбу.
The man will fry this grey/red fish.
13. Дедушка выкурит эту черную/коричневую трубку.
The grandfather will smoke this black/brown pipe.
14. Мальчик наклеит эту розовую/оранжевую марку.
The boy will stick this pink/orange stamp.
15. Женщина застегнет эту желтую/зеленую куртку.
The woman will zip this yellow/green jacket.
16. Женщина наполнит эту голубую/красную миску.
The woman will fill this blue/red bowl.
17. Мальчик съест этот коричневый/белый торт.
The boy will eat this brown/white cake.
18. Мужчина откроет этот серый/желтый замок.
The man will open this grey/yellow lock.
19. Мальчик надует этот оранжевый/зеленый шарик.
The boy will inflate this orange/green balloon.
20. Женщина стирает этот розовый/коричневый шарф.
The woman will wash this pink/brown scarf.
21. Мужчина завяжет этот синий/красный галстук.
The man will tie this blue/red tie.
22. Девочка заточит этот зеленый/желтый карандаш.
The girl will sharpen this green/yellow pencil.
23. Мужчина запечатает этот синий/зеленый конверт.
The man will seal this blue/green envelope.
24. Женщина пропылесосит этот желтый/голубой ковер.
The woman will vacuum this yellow/blue carpet.
25. Женщина наденет эту коричневую/белую юбку.
The woman will put on this brown/white skirt.
26. Мужчина погладит эту белую/черную рубашку.
The man will iron this white/black shirt.
27. Девушка зажжет эту желтую/розовую свечу.
The girl will light up this yellow/pink candle.
28. Женщина подоит эту коричневую/белую корову.
The woman will milk this brown/white cow.
29. Мальчик разрушит эту оранжевую/зеленую башню.
The boy will ruin this orange/green tower.
30. Девочка обнимет эту белую/розовую куклу.
The girl will hug this white/pink doll.
31. Женщина нажмет эту зеленую/синюю кнопку.
The woman will push this green/blue button.
32. Мужчина настроит эту красную/синюю гитару.
The man will tune this red/blue guitar.

References

- Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20(3), 242–275. <https://doi.org/10.1016/j.jneuroling.2006.10.003>
- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264. [https://doi.org/10.1016/S0010-0277\(99\)00059-1](https://doi.org/10.1016/S0010-0277(99)00059-1)
- Altmann, G. T., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, 33(4), 583–609. <https://doi.org/10.1111/j.1551-6709.2009.01022.x>
- Anwyl-Irvine, A. L., Massoníé, J., Flitton, A., Kirkham, N. Z., & Evershed, J. K. (2019). Gorilla in our midst: An online behavioural experiment builder. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-019-01237-x>
- Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online

- experiment platforms, web browsers, and devices. *Behavior Research Methods*, 53(4), 1407–1425. <https://doi.org/10.3758/s13428-020-01501-5>
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Bolker, M. B. (2015). Package ‘lme4’. *Convergence*, 12(1), 2
- Bayram, F., Rothman, J., Iverson, M., Kupisch, T., Miller, D., Puig-Mayenco, E., & Westergaard, M. (2019). Differences in use without deficiencies in competence: Passives in the Turkish and German of Turkish heritage speakers in Germany. *International Journal of Bilingual Education and Bilingualism*. <https://doi.org/10.1080/13670050.2017.1324403>
- Bayram, F., Pisa, G. D., Rothman, J., & Slabakova, R. (2021). Current Trends and Emerging Methodologies in Charting Heritage Language Grammars. In S. Montrul & M. Polinsky (Eds.), *The Cambridge Handbook of Heritage Languages and Linguistics* (pp. 545–578). chapter, Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108766340.025>
- Bialystok, E. (2024). Bilingualism modifies cognition through adaptation not transfer. *Trends in Cognitive Science*. <https://doi.org/10.1016/j.tics.2024.07.012>
- Bialystok, E., & Craik, F. I. (2022). How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic Bulletin & Review*, 29(4), 1246–1269.
- Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, 16, 893–900. <https://doi.org/10.3758/PBR.16.5.893>
- Covey, L., Fiorentino, R., & Gabriele, A. (2024). Island sensitivity in L2 learners: Evidence from acceptability judgments and event-related potentials. *Second Language Research*, 40(1), 19–50. <https://doi.org/10.1177/02676583221116039>
- Dave, S., Brothers, T., Hoversten, L. J., Traxler, M. J., & Swaab, T. Y. (2021). Cognitive control mediates age-related changes in flexible anticipatory processing during listening comprehension. *Brain Research*, 1768, 147573. <https://doi.org/10.1016/j.brainres.2021.147573>
- Degen, J., Kursat, L., & Leigh, D. D. (2021). Seeing is believing: Testing an explicit linking assumption for visual world eye-tracking in psycholinguistics. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43, No. 43)
- De Houwer, A. (2023). The danger of bilingual–monolingual comparisons in applied psycholinguistic research. *Applied Psycholinguistics*, 44(3), 343–357. <https://doi.org/10.1017/S014271642200042X>
- Di Pisa, G., Kubota, M., Rothman, J., & Marinis, T. (2022). Effects of markedness in gender processing in Italian as a heritage language: A speed accuracy tradeoff. *Frontiers in Psychology*, 13, 965885. <https://doi.org/10.3389/fpsyg.2022.965885>
- Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, 20(5), 917–930. <https://doi.org/10.1017/S1366728916000547>
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143–149. <https://doi.org/10.3758/BF03203267>
- Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, 2(2–3), 101–118. <https://doi.org/10.1080/13506289508401726>
- Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, 27(6), 443–448. <https://doi.org/10.1177/09637214187944>
- Fricke, M., & Zirnstein, M. (2022). Predictive processing and inhibitory control drive semantic enhancements for non-dominant language word recognition in noise. *Languages*, 7(3), 239. <https://doi.org/10.3390/languages7030239>
- Fuchs, Z. (2022a). Facilitative use of grammatical gender in Heritage Spanish. *Linguistic Approaches to Bilingualism*, 12(6), 845–871. <https://doi.org/10.1075/lab.20024.fuc>
- Fuchs, Z. (2022b). Eyetracking evidence for heritage speakers’ access to abstract syntactic agreement features in real-time processing. *Frontiers in Psychology*, 13, 960376. <https://doi.org/10.3389/fpsyg.2022.960376>
- Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, 23(2), 283–294. <https://doi.org/10.1017/S1366728919000026>
- Hao, J., Kubota, M., Bayram, F., González Alonso, J., Li, M., & Rothman, J. (2024). Schooling and home language usage matter in heritage bilingual processing: Sortal classifiers in Mandarin. *Second Language Research*. <https://doi.org/10.1177/02676583241270900>
- Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306. <https://doi.org/10.1515/iral-2015-0014>
- Hsu, N. S., Kuchinsky, S. E., & Novick, J. M. (2021). Direct impact of cognitive control on sentence processing and comprehension. *Language, Cognition and Neuroscience*, 36(2), 211–239. <https://doi.org/10.1080/23273798.2020.1836379>
- Hübner, R., & Töbel, L. (2019). Conflict resolution in the Eriksen flanker task: Similarities and differences to the Simon task. *PLoS ONE*, 14(3), e0214203. <https://doi.org/10.1371/journal.pone.0214203>
- Huetting, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80–93. <https://doi.org/10.1080/23273798.2015.1047459>
- Huetting, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: Prediction of spoken language. *Trends in Cognitive Sciences*, 23(6), 464–475. <https://doi.org/10.1016/j.tics.2019.03.008>
- Ito, A., & Knoeferle, P. (2022). Analysing data from the psycholinguistic visual-world paradigm: Comparison of different analysis methods. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-022-01969-3>
- Ito, A., Nguyen, H. T. T., & Knoeferle, P. (2023). German-dominant Vietnamese heritage speakers use semantic constraints of German for anticipation during comprehension in Vietnamese. *Bilingualism Language and Cognition*. <https://doi.org/10.1017/S136672892300041X>

- Kaan, E., & Grüter, T. (2021). *Prediction in second language processing and learning*. John Benjamins Publishing Company.
- Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003a). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156. [https://doi.org/10.1016/S0749-596X\(03\)00023-8](https://doi.org/10.1016/S0749-596X(03)00023-8)
- Kamide, Y., Scheepers, C., & Altmann, G. T. (2003b). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32, 37–55. <https://doi.org/10.1023/A:1021933015362>
- Karaca, F., Brouwer, S., Unsworth, S., & Huettig, F. (2024). Morphosyntactic predictive processing in adult heritage speakers: Effects of cue availability and spoken and written language experience. *Language, Cognition and Neuroscience*, 39(1), 118–135. <https://doi.org/10.1080/23273798.2023.2254424>
- Keefe, E. B., & Copeland, S. R. (2011). What is literacy? The power of a definition. *Research and Practice for Persons with Severe Disabilities*, 36(3–4), 92–99. <https://doi.org/10.2511/027494811800824>
- Koehne, J., & Crocker, M. W. (2015). The interplay of cross-situational word learning and sentence-level constraints. *Cognitive Science*, 39(5), 849–889. <https://doi.org/10.1111/cogs.12178>
- Kukona, A., Cho, P. W., Magnuson, J. S., & Tabor, W. (2014). Lexical interference effects in sentence processing: Evidence from the visual world paradigm and self-organizing models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(2), 326. <https://doi.org/10.1037/a0034903>
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. <https://doi.org/10.1080/23273798.2015.1102299>
- Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22(5), 564–582. <https://doi.org/10.1177/136700691665435>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44, 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language history questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, 23(5), 938–944. <https://doi.org/10.1017/S1366728918001153>
- Lopukhina, A., Chrabaszcz, A., Khudyakova, M., Korkina, I., Yurchenko, A., & Dragoy, O. (2019). Test for assessment of language development in Russian “KORABLIK”. In *Proceedings of the Satellite of AMLaP Conference “Typical and Atypical Language Development Symposium”*. [s. l.] (p. 30).
- Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2012). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes*, 27(10), 1479–1488. <https://doi.org/10.1080/01690965.2011.613209>
- Luk, G., & Rothman, J. (2022). Experience-based individual differences modulate language, mind and brain outcomes in multilinguals. *Brain and Language*. <https://doi.org/10.1016/j.bandl.2022.105107>
- Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology*, 126, 264–279.
- Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, 53, 372–380. <https://doi.org/10.3758/BF03206780>
- Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, 5(1), 1–10.
- Mitrofanova, N., Rodina, Y., Urek, O., & Westergaard, M. (2018). Bilinguals’ sensitivity to grammatical gender cues in Russian: The role of cumulative input, proficiency, and dominance. *Frontiers in Psychology*, 9, 395647. <https://doi.org/10.3389/fpsyg.2018.01894>
- Mitrofanova, N., Urek, O., Rodina, Y., & Westergaard, M. (2022). Sensitivity to microvariation in bilingual acquisition: Morphophonological gender cues in Russian heritage language. *Applied Psycholinguistics*, 43(1), 41–79. <https://doi.org/10.1017/S0142716421000382>
- Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139030502>
- Montrul, S. (2022). The psycholinguistics of heritage language acquisition. In *The Routledge handbook of second language acquisition and psycholinguistics* (pp. 73–84). <https://doi.org/10.4324/9781003018872>
- Montrul, S., Davidson, J., De La Fuente, I., & Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism: Language and Cognition*, 17(1), 118–138. <https://doi.org/10.1017/S1366728913000114>
- Montrul, S., & Polinsky, M. (Eds.) (2021). *The Cambridge handbook of heritage languages and linguistics*. Cambridge University Press. <https://doi.org/10.1017/9781108766340>
- Ness, T., Langlois, V. J., Kim, A. E., & Novick, J. M. (2023). The state of cognitive control in language processing. *Perspectives on Psychological Science*. <https://doi.org/10.1177/17456916231197122>
- Noh, Y., & Lee, M. (2017). The impact of inhibitory controls on anticipatory sentence processing in L2. *Journal of Cognitive Science*, 18(1).
- Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. F. (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language, Cognition and Neuroscience*, 29(2), 186–217. <https://doi.org/10.1080/01690965.2012.758297>
- Nozari, N., Trueswell, J. C., & Thompson-Schill, S. L. (2016). The interplay of local attraction, context and domain-general cognitive control in activation and suppression of semantic distractors during sentence comprehension.

- Psychonomic Bulletin & Review*, 23, 1942–1953. <https://doi.org/10.3758/s13423-016-1068-8>
- Özsoy, O., Çiçek, B., Özal, Z., Gagarina, N., & Sekerina, I. A. (2023). Turkish-German heritage speakers' predictive use of case: webcam-based vs. in-lab eye-tracking. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2023.1155585>
- Parshina, O., Lopukhina, A., & Sekerina, I. A. (2022). Can heritage speakers predict lexical and morphosyntactic information in reading? *Languages*, 7(1), 60. <https://doi.org/10.3390/languages7010060>
- Patra, A., Kirkwood, J., Middleton, E. L., & Thothathiri, M. (2023). Variation in how cognitive control modulates sentence processing. *Royal Society Open Science*, 10(4), 211969. <https://doi.org/10.1098/rsos.211969>
- Peters, R. E., Grüter, T., & Borovsky, A. (2018). Vocabulary size and native speaker self-identification influence flexibility in linguistic prediction among adult bilinguals. *Applied Psycholinguistics*, 39(6), 1439–1469. <https://doi.org/10.1017/S0142716418000383>
- Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics*, 14(2), 191–262.
- Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal*, 6(1), 40–71.
- Polinsky, M. (2018). *Heritage languages and their speakers* (Vol. 159). Cambridge University Press
- Polinsky, M., & Scontras, G. (2020). A roadmap for heritage language research. *Bilingualism: Language and Cognition*, 23(1), 50–55.
- Prystauka, Y., Altmann, G. T., & Rothman, J. (2024). Online eye tracking and real-time sentence processing: On opportunities and efficacy for capturing psycholinguistic effects of different magnitudes and diversity. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-023-02176-4>
- Qualtrics (2005). Provo, UT, USA. <https://www.qualtrics.com>. Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-020-01501-5>
- Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, 13(2), 155–163. <https://doi.org/10.1177/1367006909339814>
- Rothman, J., & Treffers-Daller, J. (2014). A prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Applied Linguistics*, 35(1), 93–98. <https://doi.org/10.1093/applin/amt049>
- Rothman, J., Bayram, F., DeLuca, V., Di-Pisa, G., Dunabeitia, J. A., Gharibi, K., Hao, J., Kolb, N., Kubota, M., Kupisch, T., & Laméris, T. (2023). Monolingual comparative normativity in bilingualism research is out of “control”: Arguments and alternatives. *Applied Psycholinguistics*, 44(3), 316–329. <https://doi.org/10.1017/S0142716422000315>
- Sekerina, I. A. (2012). The effects of grammatical gender in Russian spoken-word recognition. In V. Makarova (Ed.), *Russian language studies in North America. New perspectives from theoretical and applied linguistics* (pp. 107–132). Anthem Press. <https://doi.org/10.7135/upo9780857286505.006>
- Sekerina, I. A. (2015). Predictions, fast and slow. *Linguistic Approaches to Bilingualism*, 5(4), 532–536.
- Sekerina, I. A., & Laurinavichyute, A. K. (2020). Heritage speakers can actively shape not only their grammar but also their processing. *Bilingualism: Language and Cognition*, 23(1), 43–45. <https://doi.org/10.1017/S1366728919000440>
- Sekerina, I. A., Brooks, P. J., & Kempe, V. (2006). Gender transparency facilitates noun selection in Russian. In *The Proceedings of the 14th Annual Workshop on Formal Approaches to Slavic Linguistics. The Princeton Meeting*.
- Slim, M. S., & Hartsuiker, R. J. (2022). Moving visual world experiments online? A web-based replication of Dijkgraaf, Hartsuiker, and Duyck (2017) using PCIbex and WebGazer.js. *Journal of Behavior Research Methods*. <https://doi.org/10.3758/s13428-022-01989-z>
- Swinney, D. A. (1979). Lexical access during sentence comprehension:(Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18(6), 645–659. [https://doi.org/10.1016/S0022-5371\(79\)90355-4](https://doi.org/10.1016/S0022-5371(79)90355-4)
- Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, 18(4), 427–440. [https://doi.org/10.1016/S0022-5371\(79\)90237-8](https://doi.org/10.1016/S0022-5371(79)90237-8)
- Treffers-Daller, J., & Silva-Corvalán, C. (Eds.) (2016). *Language dominance in bilinguals: Issues of measurement and operationalization*. Cambridge University Press.
- Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(3), 528. <https://doi.org/10.1037/0278-7393.19.3.528>
- Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45, 1–67. <https://doi.org/10.18637/jss.v045.i03>
- Vos, M., Minor, S., & Ramchand, G. (2022). Comparing infrared and webcam eye tracking in the visual world paradigm. *Glossa Psycholinguistics*. <https://doi.org/10.5070/G6011131>
- Zirnstien, M., van Hell, J. G., & Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition*, 176, 87–106. <https://doi.org/10.1016/j.cognition.2018.03.001>
- Zirnstien, M., van Hell, J. G., & Kroll, J. F. (2019). Cognitive control and language ability contribute to online reading comprehension: Implications for older adult bilinguals. *International Journal of Bilingualism*, 23(5), 971–985. <https://doi.org/10.1177/1367006918763143>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.