

Lexical Associations in a Native and Non-Native Language

Affect Retrieval-Induced Forgetting

Madalina Vlasceanu (mov@princeton.edu)*

Department of Psychology, Princeton University
Princeton, NJ 08544 USA

Karina Tachihara (tachihara@princeton.edu)*

Department of Psychology, Princeton University
Princeton, NJ 08544 USA

Adele Goldberg (adele@princeton.edu)

Department of Psychology, Princeton University
Princeton, NJ 08544 USA

Alin Coman (acoman@princeton.edu)

Department of Psychology, Princeton University
Princeton, NJ 08544 USA

Abstract

Recent work suggests that speakers' lexical networks in their native and secondary languages are organized somewhat differently, with native languages showing greater systematicity. We here test this claim in a new way, by making use of the "Retrieval-induced forgetting" effect (RIF). Specifically, practicing previously encoded information through rehearsal is expected to result in better memory for that information, regardless of which language the information is encoded. The RIF effect involves the suppression of information that is *associated* with the practiced information but is itself unpracticed. Since RIF is understood to rely on the association between the practiced and unpracticed memories, we predict it will be weaker when applied in a language with weaker or less systematically organized lexical associations. Results confirm that while the expected practice effect was evident in participants' native and second languages, the RIF effect was only significant in participants' native language. We discuss the relevance and implications of this finding for second language speakers.

Keywords: retrieval-induced forgetting, second language, memory

Introduction

We live in a world more globalized than ever before, with very large numbers of immigrants, and high levels of social connectivity. This gives rise to increasing levels of multilingual experiences, especially for those

navigating life in a new language. For instance, in 2017, 67 million people in the US, or half of the residents of the 5 largest US cities, reported speaking a language other than English at home (U.S. Census Bureau, 2018). Do people process a language learned later in life the same way they process their native language(s)?

Multilingual individuals often face the complex task of switching between languages in their day-to-day functioning. Recent work has found that second language (L2) speakers are less likely to make online predictions during listening than native speakers are, even when the L2 speakers demonstrate the requisite knowledge in off-line tasks (Grüter et al., 2014; Kaan, 2014; Lew-Williams & Fernald, 2010; Ito et al., 2017; Kaan et al., 2010, 2016; Martin et al., 2013). For example, in a study that measured participants' eye gaze, Lew-Williams & Fernald (2010) found that non-native speakers of Spanish did not make use of the gendered determiner to predict an upcoming noun, despite being reasonably fluent and aware of the noun's grammatical gender. There is also recent evidence that competition between different lexical items (or *constructions*) may not have the same impact in a second language (Robenalt & Goldberg, 2016; Tachihara & Goldberg, 2019). Of course, individual differences, degree of proficiency, the relation between L1 and L2, and task demands have been found to play a role in the extent to which predictions about upcoming forms are made by L2 comprehenders (e.g., Borovsky,

Elman & Fernald 2012; Kaan 2014; Hopp 2013; DeLong et al. 2012; Martin et al. 2013; Rossi et al., 2017; Havik et al. 2009; Foucart et al. 2014; Linck et al. 2009; Wilson 2009).

Intriguingly, a semantic network analysis on free recall data suggests that L2 speakers' linguistic representations are less systematically organized than those in a native language (Borodkin, Kennet, Faust, & Mashal 2016). The memorably titled paper, "When pumpkin is closer to onion than to squash," reported that highly proficient L2 speakers showed less evidence of standard lexical associations found in L1 speakers (e.g., *pumpkin - squash*) (see also Wilks & Meara 2002). A difference in the degree of association strength may be due to less experience with specific words and categories, or due to age of acquisition effects in lexical retrieval (Ellis & Morrison 1998). Relatedly, Ning, Bartolotti, & Marian (2019) showed that Spanish-English bilinguals judged the semantic relatedness between non-obviously related concepts (*cloud-present*) to be higher than English monolinguals did. This is consistent with the idea that lexical associations in L2 may be less strong or less systematically organized.

While the free-recall task of as many instances of a category as possible, used by Borodkin et al. (2016), is a useful way to tap into the structure of broad knowledge, the possibility of including low frequency items may affect results. By definition, L2 speakers learned their second language as adults, so certain words are likely to be less familiar to them. For instance, it is possible to be entirely unfamiliar with the less common vegetable, *squash*. Moreover, one may have a different interpretation of *pumpkin* if one did not grow up carving them. More generally, to test the idea that the lexical networks of L1 and L2 speakers are structured differently, it would be especially compelling to find differences in how L2 speakers respond to relatively high frequency associations such as *apple - orange*.

The current paper tests this idea using the Retrieval-Induced Forgetting (RIF) effect (Anderson, Bjork & Bjork, 1994; Murayama, Miyatsu, Buchli, & Storm, 2014), with stimuli presented in speakers' native or second language. The classic RIF paradigm involves three phases. Participants first study a list of category-exemplar pairs (e.g., *Fruit-Apple*, *Fruit-Orange*, *Profession-Nurse*, *Profession-Teacher*). They then perform a retrieval task on a subset of studied pairs, which provides practice on just those pairs. During this phase, participants are provided with a cue such as "Fruit-*Ap*_" and are asked to supply the second word (*Apple*). Finally, participants are asked to recall as many of the pairs that had been presented during the initial study phase as possible.

Unsurprisingly, the items that were practiced during the partial recall task (Retrieval Practice Plus [RP+] items: e.g., *Apple*) are remembered best. Of interest, the unpracticed items that are associated with the practiced items (Retrieval Practice Minus [RP-] items: *Orange*)

are recalled less often than the other unpracticed items (baseline items, labeled No Retrieval Practice [NRP] items: e.g., *Nurse*, *Teacher*). The idea is that, during the selective recall task, the cue provided *Fruit-Ap_* partially activates other associates such as *Orange*. Because *Orange* is incompatible with the cue (*Ap*_), it must be suppressed, resulting in its continued suppression or "forgetting" during the recall task.

According to a widely supported account of retrieval-induced forgetting (Murayama et al., 2014), the effect occurs when, during the practice phase, related items compete for activation because of their semantic associations. This competition is resolved by the suppression of semantically associated but unpracticed items (Anderson, 2003; Levy & Anderson, 2002; Norman, Newman, Detre, 2007). Thus, if words are less strongly or less systematically associated with one another in a second language, and the retrieval-induced forgetting effect depends on this association to trigger response competition during the practice phase, then we would expect to see reduced RIF in a second language.

Thus, our interest lies in whether the RIF effect is evident to the same extent in L2 as in L1. All of the items used involve relatively high frequency words as described below in order to investigate aspects of the lexical network that are central and highly likely to be familiar to L2 speakers. RIF is known to be sensitive to factors relevant to lexical network structure (e.g., prototypicality: *apple-orange* leads to suppression but *guava-orange* does not) (Baüml, 1998). Therefore, because the retrieval of items in the RIF paradigm is dependent on the lexical relatedness of items, it allows us to investigate differences in the lexical networks of L1 vs. L2.

Given that lexical networks are part of our semantic memory, the current experiment aims to probe them within the context of a memory task. There is evidence that memory retrieval affects the structure of memory itself (McDonald, 2013; Karpicke, 2012; Lewis-Peacock & Norman, 2014). Thus, testing the L1 and L2 difference within a memory task may suggest an answer to the question as to *why* L1 and L2 lexical networks are different in the first place.

We recruited a sample of 100 native English speakers who speak Spanish as a second language, and randomly assigned them to either perform the experiment in English or in Spanish. We hypothesized that participants in both conditions would display the rehearsal (practice) effect (*hypothesis 1*), but that only participants in the native speaker (English) condition would display the retrieval-induced forgetting effect (*hypothesis 2*).

Method

Participants

To detect a moderate effect size of 0.4 for paired-sample comparisons with 0.80 power, we

collected data from 100 participants ($M_{age}=31.46$; $SD_{age}=9.19$; 51% women), as preregistered on *aspredicted.com*. Participants were recruited via Cloud Research (Litman, Robinson, & Abberbock, 2016) with an advertisement that requested native English speakers who spoke Spanish. They were compensated at the platform's standard rate. The pre-registered stopping rule for data collection was that we collected data until we reached 50 participants in each of the two between-subject conditions. All participants confirmed that they were native English speakers. We ensured that participants could speak Spanish by asking them five 4-answer multiple choice questions that required reasonably strong knowledge of Spanish grammar; only participants who answered at least 3 out of 5 questions correctly were included.

The study protocol was approved by the Princeton University Institutional Review Board.

Stimulus materials

We used eight categories (e.g., *fruits, mammals, professions, mammals, furniture, emotions, foods, electronics*) with six representative exemplars in each (e.g., *apple, orange*). For the L2 condition, Spanish category and exemplar names were used. Each exemplar was chosen such that its usage frequency was higher than 4000 (number of occurrences in text corpus, from <https://www.english-corpora.org/coca/>, <https://www.wordfrequency.info/spanish.asp>) in both languages, and such that its first two letters were unique within a category in both languages.

Design and Procedure

Participants were told they would be participating in an experiment about categorization and were directed to the survey on the Qualtrics platform. After giving informed consent, all participants were given the Spanish vocabulary test which took the form of five multiple choice questions (chance level performance = 25%) (e.g., “En mi barrio no hay *ningún/algún/ninguno/ alguno* parque”). Then, participants were randomly assigned to one of two between-subjects conditions: L1 Condition or L2 Condition. The two conditions were equivalent in all aspects other than the language in which the experiment was conducted: the L1 Condition was administered in English, whereas the L2 Condition was administered entirely in Spanish.

Participants in both conditions took part in all three phases of a standard RIF paradigm: study (encoding), selective cued-retrieval (practice), and recall (test) (Anderson, Bjork, Bjork, 1994). During the *encoding*

phase, participants were instructed to pay attention to a series of category-exemplars pairs (one on each page). The eight categories with six exemplars in each were presented in a randomized block design, with the items within each category also randomized. Each category-exemplar pair (e.g., *Fruit-Orange*) was on the screen for exactly 5 seconds, thus, the entire encoding phase lasted 4 minutes. Then, participants went through a *practice* phase, in which they were given a category name and the first two letters of an exemplar name (e.g., *Fruit: Or__*), and they had to write the rest of the exemplar letters in a text box (one on each page). They were allowed to take as long as 10 seconds for each practiced exemplar. In total, participants practiced half of the exemplars from half of the categories, both randomly chosen.

Finally, in the *test* phase, participants were asked to remember all the initially encoded exemplars, in a cued recall task in which they were given the category name and they had to write as many exemplars as they could remember. The category names were presented in a random order, one on each page, and participants were allowed to spend up to 30 seconds on each category, as this response window was found to be ample in prior work. Participants were then asked a series of demographic questions, after which they were debriefed. All procedures and analyses were preregistered.

Results

The pre-registered analysis was a repeated-measures ANOVA with retrieval type as the within-subject variable: retrieval practice items (RP+), items from related category but without retrieval practice (RP-), and items from an unrelated category without retrieval practice (baseline/NRP). Condition (English, Spanish) was a between-subjects variable, and proportion of items recalled was the dependent variable. We found a significant main effect of retrieval type, $F(2, 196)=114.884$, $p<0.001$, $\eta^2=0.54$, a significant main effect of condition, $F(1, 98)=15.012$, $p<0.001$, $\eta^2=0.133$, and a significant interaction $F(2, 196)=4.001$, $p<0.02$, $\eta^2=0.39$.

As anticipated, the first hypothesis was confirmed: the expected rehearsal effect was found in both L1 and L2 groups: the recall rate of RP+ items was significantly higher than the recall rate of baseline/NRP items. In L1: ($M=0.73$, $SD=0.19$ vs. $M=0.50$, $SD=0.25$), $t(49)=8.19$, $p<0.001$, $d=1.15$, $CI[0.17, 0.28]$, and L2 Condition ($M=0.52$, $SD=0.22$ vs. $M=0.34$, $SD=0.21$), $t(49)=7.98$, $p<0.001$, $d=1.12$, $CI[0.13, 0.22]$ (Figure 1).

Results also support the second hypothesis. Specifically, the retrieval-induced forgetting effect was significant in the L1 condition, but not in the L2 condition: in L1, RP- items ($M=0.41$, $SD=0.29$) were recalled significantly less than baseline/NRP items ($M=0.50$, $SD=0.25$), $t(49)=4.11$, $p<0.001$, $d=0.58$, $CI[0.04, 0.13]$. In L2, RP- items ($M=0.31$, $SD=0.21$) were not significantly less well recalled than

baseline/NRP items ($M=0.34$, $SD=0.21$), $t(49)=1.71$, $p=0.093$, $d=0.24$, $CI[-0.075, 0.006]$ (Figure 1). Because recall in the RP- condition was numerically lower than the baseline/NRP condition, we additionally ran a power analysis and found that even if we doubled the sample size, the RIF effect in L2 would not be significant.

In order to include items as well as subjects as random variables, we additionally ran a generalized mixed effects model (using R version 3.1_0, with lme4 version 1.1_21, logistic link function). To predict recall, we included condition (L1, L2) and retrieval type (RP+, RP, baseline/NRP) as fixed effects, as well as by-participant and by-item random intercepts. We find a significant effect of type (practiced: $\beta=1.24$, $SE=0.18$, $z=6.94$, $p<0.001$; unpracticed: $\beta=-0.45$, $SE=0.17$, $z=-2.63$, $p=0.008$) and condition ($\beta=-0.88$, $SE=0.24$, $z=-3.65$, $p=0.003$). We also find more items were recalled in the L1 condition than in L2. This may be due to across-the-board increased difficulty of lexical retrieval in L2 (Antón-Méndez & Gollan, 2010; Sandoval, Gollan, Ferreira, & Salmon, 2010; Taler, Johns, Young, Sheppard, & Jones, 2013; cf. Roberts & Dorze, 1997; Rosselli et al., 2000). There is a significant interaction of condition and type for practiced items ($\beta=-0.33$, $SE=0.17$, $z=-2.0$, $p=0.046$), but not for unpracticed items ($\beta=0.27$, $SE=1.65$, $z=1.64$, $p=0.1$).

Because the ANOVA analysis showed a significant interaction of condition and retrieval-type, we also looked at the L1 and L2 conditions separately, using models that were the same except that condition was not included. This confirms the expected rehearsal effect in both English ($\beta=1.12$, $SE=0.17$, $z=6.42$, $p<0.001$) and Spanish ($\beta=0.91$, $SE=0.16$, $z=5.52$, $p<0.001$). It also confirms evidence of the RIF effect only in English ($\beta=-0.47$, $SE=0.17$, $z=-2.67$, $p=0.007$), not in Spanish ($\beta=-0.19$, $SE=0.17$, $z=-1.09$, $p=0.274$).

Finally, we performed an exploratory analysis of the effect of L2 proficiency on recall. We ran a generalized mixed model on the L2 dataset with recall as our outcome and retrieval type and proficiency as interacting fixed effects with by-participant and by-item random intercepts. We find no interaction effect for practice items ($\beta=0.009$, $SE=0.005$, $z=1.78$, $p=0.076$) and unpracticed items ($\beta=0.005$, $SE=0.005$, $z=0.9$, $p=0.36$). We caution against inferring that proficiency does not impact RIF effects, since self-reported proficiency measures in online studies are known to be noisy. Additionally, we had limited our L2 sample to those who passed the vocabulary test but did not consider themselves to be 100% proficient in L2. A larger sample with a more variable proficiency distribution would be needed to better assess the effect of proficiency on the RIF effect.

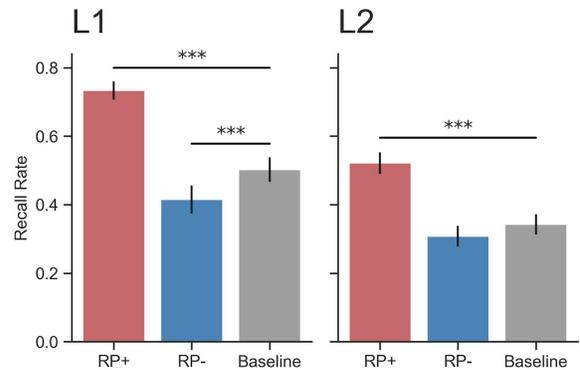


Figure 1: Recall proportions in the L1 Condition (left) and in L2 Condition (right). RP+ items represented in red, RP- items represented in blue, Baseline/NRP items represented in grey; Error bars represent ± 1 standard errors of the mean.

Discussion

The current study compared the degree of retrieval-induced forgetting in a first language condition (English) and a second language condition (Spanish) among a group of monolingual English speakers who were reasonably proficient in Spanish as a second language. We find that while participants in both conditions benefited from the rehearsal effect -- practiced items were remembered better than unrelated, unpracticed items -- only participants in the first language condition revealed the retrieval-induced forgetting effect in which semantically-related unpracticed items were remembered less well than unrelated unpracticed items.

These findings extend and complement prior work in various ways. First, the first language condition results replicate the retrieval-induced forgetting effect (Anderson, Bjork, Bjork, 1994). The lack of evidence of the RIF effect in a second language is consistent with the idea that the network of semantic associations is less strong or less systematically organized in a second language than in a first language (Borodkin, et al., 2016). That is, to the extent that word pairs associated in L1 are less associated in L2, the current results are expected: there should be no difference between unpracticed related and unrelated items, if neither is strongly associated with the practiced items. In other words, for our native English-speaking participants, “apple” appears to be more strongly associated with “orange” in English than in Spanish. Critically, the fact that current evidence comes from relatively high frequency words further suggests that even central aspects of the lexical network may be structured differently in L2.

The absence of the retrieval-induced forgetting in the second language condition could be due to several other factors. It is possible that participants lacked sufficient proficiency in Spanish, their second language, to be able to recall unpracticed items at strong rates. Participants were recruited online and only

needed to pass the brief grammaticality test in order to qualify for the experiment. The overall main effect of better recall in L1 is consistent with this possibility. At the same time, participants in the L2 condition did show the rehearsal effect, demonstrating that on the basis of only partial cues (e.g., *Fruit: Ap__*) they were able to rehearse the items enough to result in better recall of those items (e.g., *Apple*). Therefore, while lower proficiency may explain the main effect of lower recall in L2, it does not directly explain the lack of RIF effect: the fact that related unpracticed items failed to show evidence of suppression.

Another possibility for the absence of the RIF effect in the L2 condition may involve differences in the retrieval process rather than differences in representation. It is possible that the need to suppress L1 when using one's L2 may result in less effective competition processes in L2. For instance, it may be that competition between words from two languages (e.g., *manzana* vs. *apple*) may lead to weaker competition between words in L2 during the practice phase within (e.g., *manzana* vs. *naranja*). This may then reduce the suppression of the semantically related but unpracticed item in the second language (*naranja*), leading to a lack of retrieval-induced forgetting in the second language condition. A third possibility is that the representation and the retrieval process interact with one another. If the retrieval process in L2 is affected by the need to inhibit L1, it could impact the lexical network of L2 over the course of learning, which may in turn impact the retrieval process and so on. As noted in the introduction, this was one of the reasons why we used a process-oriented approach, testing RIF, in order to investigate the lexical structures in L1 and L2. These different possible mechanisms need to be teased apart empirically in future work.

Limitations, future work, and implications

The current study requires replication on a larger sample of participants with a better calibrated measure of L2 proficiency in order to confirm the finding and investigate whether it is dependent on proficiency in L2. Furthermore, investigating other systematic differences among L2 speakers, such as age of acquisition and amount of exposure might be of interest. We know that individual experiences shape memory structures and retrieval processes (Bilson, Yoshida, Tran, Woods, & Hills, 2015). Thus it is plausible that differences in linguistic experiences among L2 speakers' would shape their memory structures in differing ways.

To better understand the mechanisms involved, the current findings would also benefit from work investigating the neural substrates producing these behavioral results. Neural mechanisms that have already been proposed to explain the RIF effect as a frontally-mediated executive-control process (Levy & Anderson, 2002), or a medial temporal lobe inhibitory

process (Norman et al., 2007) provide a natural starting point in such an investigation.

The findings reported here might have important implications for the collective memory of second language speakers. The retrieval-induced forgetting effect has been shown to extend to social interactions, in which speakers can trigger response competition and thus induced forgetting in both themselves *and* in their listeners, an effect known as socially shared retrieval-induced forgetting (Cuc, Koppel, Hirst, 2007; Coman, Manier, Hirst, 2009). Occurring through conversation, the effect has been shown to propagate through social networks, and thus impact the collective memory of communities (Coman, Momennejad, Drach, Geana, 2016; Vlasceanu, Enz, Coman, 2018) and even the collective beliefs of a community (Vlasceanu, Morais, Duker, Coman, 2020). If collective memories discussed in a second language are less amenable to the influence of the socially-shared retrieval-induced forgetting effect, it may have potentially significant implications related to the cohesiveness of immigrant communities who use a second language to communicate with one another or for the potential of immigrants using a second language to socially cohere with the larger community.

Acknowledgements

We thank the reviewers for helpful feedback and suggestions. This research project was funded by the Graduate Student Research Grant awarded by the Program in Cognitive Science at Princeton University.

References

- Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of memory and language*, 49(4), 415-445.
- Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1063.
- Antón-Méndez, I., & Gollan, T. H. (2010). Not just semantics: Strong frequency and weak cognate effects on semantic association in bilinguals. *Memory & cognition*, 38(6), 723-739.
- Bäuml, K. (1998). Strong items get suppressed, weak items do not. *Psychonomic Bulletin and Review*, 5(3), 459-463.
- Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., & Hills, T. T. (2015). Semantic facilitation in bilingual first language acquisition. *Cognition*, 140, 122-134.

- Borodkin, K., Kenett, Y. N., Faust, M., & Mashal, N. (2016). When pumpkin is closer to onion than to squash: The structure of the second language lexicon. *Cognition*, 156, 60-70.
- Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of experimental child psychology*, 112(4), 417-436.
- Coman, A., Manier, D., & Hirst, W. (2009). Forgetting the unforgettable through conversation: Socially shared retrieval-induced forgetting of September 11 memories. *Psychological Science*, 20(5), 627-633.
- Coman, A., Momennejad, I., Drach, R. D., & Geana, A. (2016). Mnemonic convergence in social networks: The emergent properties of cognition at a collective level. *Proceedings of the National Academy of Sciences*, 113(29), 8171-8176.
- Cuc, A., Koppel, J., & Hirst, W. (2007). Silence is not golden: A case for socially shared retrieval-induced forgetting. *Psychological Science*, 18(8), 727-733.
- DeLong, K. A., Groppe, D. M., Urbach, T. P., & Kutas, M. (2012). Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language*, 121, 226-239.
- Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(2), 515.
- Foucart, A., Martin, C. D., Moreno, E., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 1461-1469.
- Futrell, R., & Gibson, E. (2017). L2 processing as noisy channel language comprehension. *Bilingualism: Language and Cognition*, 20(4), 683-684.
- Grüter, T., Hurtado, N., Marchman, V. A., & Fernald, A. (2014). Language exposure and online processing efficiency in bilingual development. *Input and experience in bilingual development*, 13, 15-36.
- Havik, E., Roberts, L., van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning*, 59, 73-112.
- Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33-56.
- Ito, A., Martin, A., & Nieuwland, M. (2017). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69, 574-588.
- Kaan, E., Dallas, A., & Wijnen, F. (2010). Syntactic predictions in second-language sentence processing. *Linguistics Today*, 164, 207-214.
- Kaan, E., Kirkham, J., & Wijnen, F. (2016). Prediction and integration in native and second-language processing of elliptical structures. *Bilingualism: Language and Cognition*, 19, 1-18.
- Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, 4(2), 257-282.
- Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in cognitive sciences*, 6(7), 299-305.
- Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of memory and language*, 63(4), 447-464.
- Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological science*, 20(12), 1507-1515.
- Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior research methods*, 49(2), 433-442.
- Martin, C. D., Thierry, G., Kuipers, J., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69, 574-588.
- Miller, A. K. (2019). Growing Anticipatory Connections during Online Processing: The Use of Grammatical Gender Cues in L2 French. In *2017 Second Language Research Forum* (pp. 117-126). Cascadilla Proceedings Project.
- Murayama, K., Miyatsu, T., Buchli, D., & Storm, B. C. (2014). Forgetting as a consequence of retrieval: A meta-analytic review of retrieval-induced forgetting. *Psychological bulletin*, 140(5), 1383.
- Ning, S., Bartolotti, J., & Marian, V. (2019). On Language and Thought: How Bilingualism Affects Conceptual Associations. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*.
- Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological review*, 114(4), 887.
- Robenalt, C., & Goldberg, A. E. (2015). Judgment evidence for statistical preemption: It is relatively better to vanish than to disappear a rabbit, but a lifeguard can equally well backstroke or swim

- children to shore. *Cognitive Linguistics*, 26(3), 467-503.
- Robenalt, C., & Goldberg, A. E. (2016). Nonnative speakers do not take competing alternative expressions into account the way native speakers do. *Language Learning*, 66(1), 60-93.
- Roberts, P. M., & Le Dorze, G. (1997). Semantic organization, strategy use, and productivity in bilingual semantic verbal fluency. *Brain and language*, 59(3), 412-449.
- Rosselli, M., Ardila, A., 2002. A cross-linguistic comparison of verbal fluency tests. *Second Language Research*, 18 (4), 303–324.
- Rossi, E., Diaz, M., Kroll, J. F., & Dussias, P. E. (2017). Late bilinguals are sensitive to unique aspects of second language processing: Evidence from Clitic pronouns word-order. *Frontiers in Psychology*, 8, 1–13.
- Sandoval, T.C., Gollan, T.H., Ferreira, V.S., Salmon, D.P., 2010. What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition* 13 (02), 231–252.
- Tachihara, K., & Goldberg, A. E. (2020). Reduced Competition Effects and Noisier Representations in a Second Language. *Language Learning*, 70(1), 219-265.
- Taler, V., Johns, B. T., Young, K., Sheppard, C., & Jones, M. N. (2013). A computational analysis of semantic structure in bilingual verbal fluency performance. *Journal of Memory and Language*, 69(4), 607-618.
- Vlasceanu, M., Enz, K., & Coman, A. (2018). Cognition in a social context: a social-interactionist approach to emergent phenomena. *Current Directions in Psychological Science*, 27(5), 369-377.
- Vlasceanu, M., Morais, M., Duker, A., & Coman, A. (2020). The synchronization of collective beliefs: From dyadic interactions to network convergence. *Journal of Experimental Psychology: Applied*. Advance online publication. <https://doi.org/10.1037/xap0000265>
- Wilks, C., & Meara, P. (2002). Untangling word webs: Graph theory and the notion of density in second language word association networks. *Second Language Research*, 18(4), 303-324.
- Wilson, F., & Chalton, F. F. L. W. (2009). Processing at the syntax-discourse interface in second language acquisition.
- Jones, M. N., Hills, T. T., & Todd, P. M., Hidden processes in structural representations: A reply to abbott, austerweil, and griffiths (2015). (2015).