

Language learners' unacceptability judgments improve with repeated exposure to acceptable sentences

Karina Tachihara (tachihara@princeton.edu)

Department of Psychology, Princeton University
Princeton, NJ 08540 USA

Adele E. Goldberg (adele@princeton.edu)

Department of Psychology, Princeton University
Princeton, NJ 08540 USA

Abstract

Recent work has raised a question about whether adult language learners take advantage of indirect negative evidence (here, *statistical preemption*) while learning a new language. Statistical preemption predicts that exposure to conventional formulations results in better recognition that unconventional formulations are unacceptable. In a preregistered study, 61 undergraduates enrolled in Spanish classes were exposed to instances of *conventional* constructions in Spanish for 3 days to determine whether the exposure would bring their responses to *unconventional* formulations into closer alignment with those of native Spanish speakers. Judgment data confirms an effect of statistical preemption: students showed an increased recognition of the fact that unconventional (unwitnessed) formulations were unacceptable. These results are consistent with the idea that learning a new language is, in large part, learning which formulations to avoid: learning what *not* to say.

Keywords: second language, learning, competition

Introduction

Learning a new language as an adult is difficult and slow for most people. Learners often produce unconventional utterances that are judged unacceptable by “native” speakers, which is often viewed as an indication that the learner lacks language proficiency, regardless of communicative success. Unconventional language may be used because learners are unaware of a conventional way to express their intended message, or they may be aware of a conventional alternative without fully appreciating that the unconventional formulation is considered unacceptable. Thus, a language learner needs to learn both what to say and what not to say, in order to be recognized as a proficient speaker of a new language.

Results from several studies demonstrate that language learners are markedly more tolerant of unconventional sentences than native speakers are (Ambridge & Brandt, 2013; Robenalt & Goldberg, 2016; Tachihara & Goldberg, 2020). For instance,

adult learners of English rate sentences like, *Dan forced that Helen plays tennis*, as markedly more acceptable than native English speakers do (Tachihara & Goldberg, 2020). In fact, in multiple studies, the difference between language learners' and native speakers' judgments on unacceptable formulations is significantly larger than on conventional formulations of the same complexity (*Dan forced Helen to play tennis*). This means that language learners' judgments differ from native speakers' more on unconventional than conventional language (Robenalt & Goldberg, 2016; Tachihara & Goldberg, 2020).

How does one learn that certain formulations are unconventional and unacceptable? Child learners of a language inevitably learn what is conventional and what is unconventional in their native language through their experience with the language. Presumably, they come to avoid unconventional formulations because they repeatedly witness and learn to strongly prefer alternatives over the course of many years of daily exposure. The conventional formulations come to *suppress* alternative ways of expressing the identical message. This process, called *statistical preemption* relies on competition between the unconventional and conventional formulations (Goldberg, 1995; see also Chouinard & Clark, 2003 for a related discussion of *recasts*). Statistical preemption occurs when a learner predicts one formulation but repeatedly witnesses another. For instance, a learner may predict that a speaker will say *Dan forced that Helen plays tennis*. But because the listener is far more likely to instead hear, *Dan forced Helen to play tennis*, to express the identical intended message, the latter conventional formulation is reinforced, and the unwitnessed formulation becomes suppressed (Boyd & Goldberg, 2011; Robenalt & Goldberg, 2015; Perek & Goldberg, 2017).

The current paper asks whether adult language learners – particularly classroom students learning Spanish – can benefit from negative indirect evidence in the form of statistical preemption in order to learn

what is unacceptable. We, therefore, picked four constructions that were verified by Spanish instructors as common mistakes amongst Spanish learners. This issue of learning unacceptability is relevant because in prior work, we found that learners' judgments on an unconventional formulation were unaffected by a single exposure to a conventional paraphrase immediately before judgments on the unconventional formulation was elicited (Tachihara & Goldberg 2020). Here we hypothesize that statistical preemption (in language learners) may simply require repeated exposure, and possibly, sleep to show an effect.

The idea is motivated from prior work in the lexical domain, which has shown that competition between linguistic formulations requires long-term memory. In a novel word learning study by Gaskell & Dumay (2003), participants showed signs of high recognition accuracy but no lexical competition immediately after exposure. Yet after the 4th day of training lexical competition observed. This, along with other studies comparing immediate and delayed testing, suggest that lexical competition requires a period of consolidation through sleep (Dumay & Gaskell 2007; Gais, Lucas, & Born, 2006; Lindsay & Gaskell 2010; Mattys & Clark 2002). While lexical competition in word recognition tasks differs from that of sentence-level competition in a judgment task, we hypothesize that the general processes are the same.

Therefore, in the current study, we ask whether *three days* of repeated exposure to conventional formulations leads learners' judgments on unconventional paraphrases to align closer to those of native speakers of Spanish. Specifically, we look at whether English-speaking students learning Spanish become better able to recognize unconventional sentences as unacceptable after the three days of exposure to acceptable paraphrases. We specifically predict that learners at the intermediate level would benefit the most from statistical preemption, since early learners may not activate competitive forms sufficiently, and highly proficient learners already behave similarly to native speakers (Robenalt & Goldberg, 2016; Tachihara & Goldberg, 2020). In addition to the judgment task, which measured explicit awareness of acceptability, we included a self-paced reading task as an implicit, online measure of sentence processing.

Method

Preregistration. The preregistered experimental design and analyses were followed unless otherwise specified (anonymous preregistration: https://aspredicted.org/CJB_BSN).

Participants

We planned to test 100 participants enrolled in any level of Spanish class at Princeton University. Their course level provided an objective measure of proficiency. 128 participants took part in the study during 2 semesters of recruiting; however, only 73 completed the critical final assessment. Two participants were excluded because they indicated that their native language was Spanish and they rated their Spanish proficiency to be at ceiling. All other participants specified that they were native English speakers. Ten additional participants were excluded for not passing the preregistered 75% threshold on comprehension questions. Out of the 61 total participants, there was roughly an equal number of participants in the beginner level, intermediate level, and advanced level.

Stimulus materials

The stimuli consisted of 4 construction types: copula choice, adjective position, grammatical gender, and the clausal complement construction (Table 1), all recognized to be challenging for Spanish learners to master. The initial assessment consisted of 16 unconventional sentences, 16 conventional sentences, and 40 conventional filler sentences. The final assessment included additional unconventional

Construction type	Unconventional formulation/ Conventional formulation
1. <i>ser</i> vs. <i>estar</i>	?La estación de tren es en esta calle. La estación de tren está en esta calle. “The train station is on this street.”
2. prenominal vs. postnominal adjectives	?El viejo hermano de Lola es guapo. El hermano viejo de Lola es guapo. “Lola’s older brother is handsome.”
3. <i>el</i> vs. <i>la</i>	?Usamos la mapa para encontrar la casa. Usamos el mapa para encontrar la casa. “We use the map to find the house.”
4. <i>que</i> vs. <i>a</i>	?Rafael obligó que ellos irían al cine. Rafael obligó a ellos a ir al cine. “Rafael forced them to go to the movie.”

Table 1: Four Construction Types used as stimuli: which types participants were exposed to was counterbalanced across groups.

sentences in order to see whether any effect of exposure would generalize beyond the particulars of the sentences witnessed. Thus, the final assessment consisted of 32 unconventional sentences, 16 conventional sentences, and 64 conventional filler sentence.¹ On each of the 3 exposure days, participants witnessed only conventional sentences: 8 conventional target sentences and 8 conventional filler sentences.

Design and Procedure

The experiment was administered over the course of one week. At pretest, participants responded to a questionnaire about their language backgrounds. At the initial assessment and the final assessment on day 2 and 6, respectively, participants read conventional and unconventional sentences in a self-paced reading task and completed a judgment task on the same set of sentences. During the intervening days of exposure (days 3-4-5), participants read only conventional sentences in a self-paced reading task. The conventional sentences on each of the exposure days were unique instances of the same constructions (Table 2). All sessions except the pretest questionnaire included comprehension questions, which were used to encourage and assess participants' attention. Comprehension questions served as the preregistered exclusion criterion (75% accuracy required). In an additional effort to engage participants, we included 16 non-linguistic encouragement gifs (ex: Jennifer Lopez clapping) which appeared at random intervals throughout the tasks.

1	Pretest	
2	First assessment	?La estación de tren es en esta calle. La estación de tren está en esta calle.
3	Exposure	La estación de tren está en esta calle.
4	Exposure	La tienda que le gusta a Daria está en ese bloque.
5	Exposure	El taxi está en la calle incorrecta.
6	Final assessment	?La estación de tren es en esta calle. La estación de tren está en esta calle.

Table 2: Example stimuli for each day of the experiment.

We divided participants into two subgroups as follows. On days 3-4-5, one group of participants was exposed to the two construction types: *ser* vs. *estar* and prenominal vs. postnominal adjectives. The other

group was exposed to two other construction types: *el* vs. *la* and *que* vs. *a*. This design allows us to compare the effect of exposure on particular constructions between subgroups, while controlling for the delay and the initial and final assessments.

In the judgment task on Day 2 and 6, participants rated the acceptability of each sentence on a gradient scale between 0-100 (100 being fully acceptable). Two examples were provided to clarify the task: one unacceptable sentence (*A mí me gusto la película* was assigned low rating) and one acceptable (*Yo vivo aquí*, was assigned a high rating).

Participants read all the sentences using a cumulative self-paced reading task on Inquisit 6 (Just, Carpenter & Woolley, 1982). Words appeared one at a time and remained visible until participants pressed the space button to see the following word. When each sentence ended, participants clicked a button that appeared on the bottom right corner of the screen. Unconventional sentences made up only 25% of the stimuli, in an effort to mitigate participants' expectations of reading unconventional sentences. To familiarize participants with self-paced reading, the first 25% of sentences were filler sentences. For the rest of the self-paced reading task and for the entirety of the judgment task, the order of the sentences was randomized for each participant.

Results

We first tested whether learners were able to distinguish conventional from unconventional sentences at the initial assessment. As expected, they were, replicating prior work on learners of English (Robenalt & Goldberg, 2016; Tachihara & Goldberg, 2020). That is, a linear mixed model confirms that conventionality predicted acceptability judgments, with random intercepts for subjects and items included, for the learners of Spanish at the initial assessment ($\beta = 24.15$, $t = 7.41$, $p < 0.0001$). Thus, participants knew Spanish well enough to assign higher acceptability ratings to conventional than to unconventional sentences.

In order to examine the effect of proficiency, we consider acceptability scores at initial assessment as a function of class level. We ran a mixed effects model with conventionality and class as interacting fixed effects and random intercepts for subjects and items, as preregistered. We found a significant interaction, meaning that as the proficiency increased, the difference in judgment scores between conventional and unconventional sentences also increased ($\beta = 4.59$,

¹ Due to experimenter error, the judgment task at the final assessment included a random subset of 40 unconventional and conventional sentences instead of 48.

$t = 6.69, p < 0.0001$). Figure 1 displays each class from lower to higher proficiency. As is visible, the effect was driven by the unconventional sentences; the same model shows a significant effect of class on the

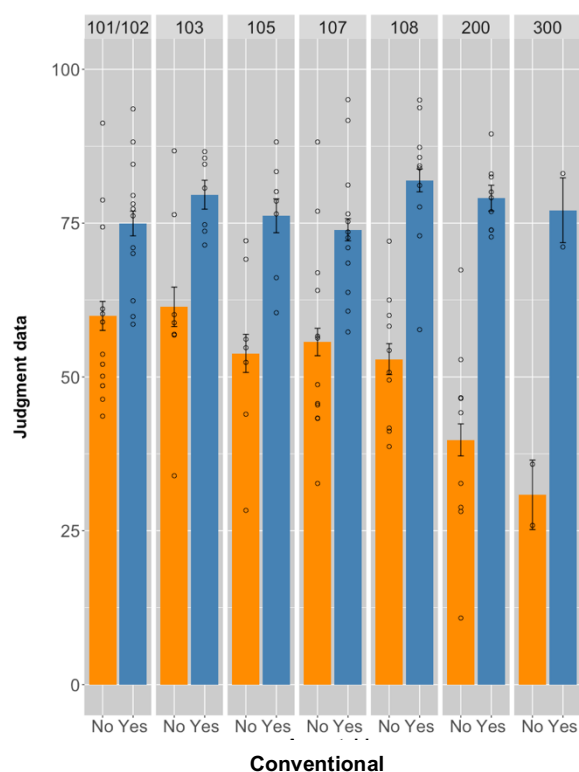


Figure 1: Initial acceptability judgments from students in different class levels (specified at top, x-axis)

unconventional sentences only (unconventional: $\beta = -3.61, t = -3.93, p = 0.0002$; conventional: $\beta = 1.07, t = 1.65, p = 0.11$). In other words, as proficiency increases, judgments for unconventional sentences decreases while judgments on conventional sentences remains largely unchanged.

Our primary hypothesis was that repeated exposure to acceptable sentences would impact judgments on unconventional paraphrases, which would provide evidence of statistical preemption in adult language learners. Results confirm a significant effect of exposure: participants gave lower ratings to those unconventional constructions for which they had read acceptable paraphrase constructions during the three days of exposure (Figure 2).

A linear mixed effects model with judgment score at final assessment as outcome and exposure as fixed effect with random intercepts for subject and item confirmed the effect ($\beta = -3.88, t = -2.79, p = 0.005$). To make sure that the effect was not driven by a single construction, we ran the model with an added random effect of construction type in an exploratory analysis

and again found an effect of exposure ($\beta = -3.90, t = -2.81, p = 0.005$). In other words, we find evidence of statistical preemption in language learners: reading conventional paraphrases led participants to rate unconventional sentences as appropriately less acceptable.

More specifically, we had preregistered an expectation that the effect of statistical preemption would be strongest for learners at the intermediate level. This prediction was borne out as the effect of exposure is significant specifically for intermediate level learners of Spanish ($\beta = -6.25, t = -3.05, p = 0.002$); and not for beginner level ($\beta = 0.68, t = 0.34, p = 0.74$) or advanced level learners ($\beta = -2.11, t = -1.16, p = 0.25$).

The self-paced reading task provided no evidence that learners slowed down when they read unconventional sentences, despite the judgment data showing that they recognized a difference between conventional and unconventional sentences even at the First Assessment. For each of the sentences in the task, a target region was identified prior to data collection. The target word was the first word at which one could detect that the sentence was unconventional. As preregistered, to allow for possible spillover effects, the next two words were included as well. A slower reaction time during the key window would indicate participants were able to detect that the sentence is unconventional (Jegerski, 2014; VanPatten & Jegerski,

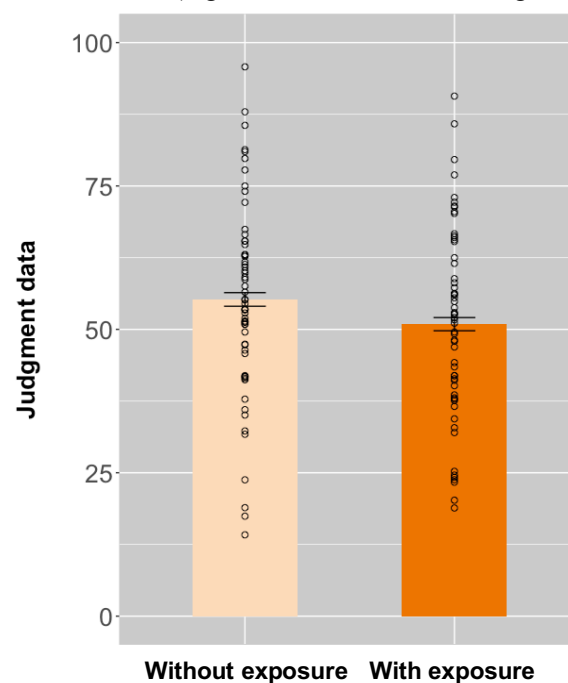


Figure 2: Acceptability judgments on unconventional sentences at final assessment, with and without exposure to conventional paraphrases.

2010). We found no evidence that language learners slowed down during the key window when reading unconventional sentences. Specifically, we ran a mixed effects model with log sum of reading time over target region as our outcome and conventionality as the fixed effect. Random intercepts for subjects and items were included. We found no effect of conventionality ($\beta = 0.019, t = 0.35, p = 0.73$). We also found no effect of exposure on the reading times of unconventional sentences. We ran a mixed effects model with log sum of reading time over target region as our outcome and exposure as the fixed effect on the unconventional sentences from the final assessment. Random intercepts for subjects and items were included. Results show no effect of exposure: participants did not slow down when reading unconventional sentences after being exposed to conventional paraphrases ($\beta = -0.21, t = -1.52, p = 0.13$). The accuracy of comprehension questions on these sentences was high ($M = 94.83\%$), indicating that participants were paying attention to the content of the sentences.

Discussion

Adult language learners tend to treat unconventional formulations more leniently than native speakers do, particularly when a conventional formulation exists to express the identical intended message (Ambridge & Brandt, 2013; Robenalt & Goldberg, 2016; Tachihara & Goldberg, 2020, but cf. Zhang & Mai, 2018). Only at high levels of proficiency do their judgments on unconventional sentences align with native speakers, while their judgments on familiar formulations tend to remain stable across proficiency levels (Robenalt & Goldberg, 2016; Tachihara & Goldberg, 2020). We document this pattern here in Figure 1, which is based on data from the initial judgment task: our classroom learners of Spanish judge conventional sentences as similarly acceptable regardless of their course-level in Spanish; it is their ratings on unconventional sentences that vary across proficiency levels.

Critically, after repeated exposure to conventional paraphrases of a subset of the unacceptable constructions, the judgment data reveals that adult language learners take advantage of the indirect evidence afforded by statistical preemption. That is, repeated exposure to conventional sentences allowed them to better recognize that unconventional sentences were markedly unacceptable. The fact that exposure was provided across multiple days distinguishes the current study from prior work that had failed to find an immediate influence of statistical preemption in language learners after single-shot exposure (Tachihara & Goldberg, 2020).

As predicted, the effect of preemptive exposure was concentrated in the intermediate-level learners. This was expected because statistical preemption relies on competition between sentence formulations, which requires learners to activate competing sentence formulations. Because the strength of activation depends on experience, one needs sufficient experience to activate alternative sentence formulations. Students who are just learning a new language likely have difficulty learning from statistical preemption because they have had too little experience with the language to sufficiently activate competing formulations. Learners who are already highly proficient already behave similarly to native speakers (Robenalt & Goldberg, 2016; Tachihara & Goldberg, 2020). This explains why intermediate speakers, who have enough experience to activate sentence formulations but not enough experience to fully appreciate the unconventional formulations as unacceptable were the students that benefited from exposure.

We designed our current study such that exposure occurred repeatedly over several days, primarily to allow for sleep consolidation to take place. Sleep has been identified as a critical component of lexical competition, which could mean that sleep may also be a critical component of statistical preemption (Gaskell & Dumay, 2003). Since the current design cannot disentangle the effects of repeated exposure and sleep, future work is required to investigate the role of each independently.

While explicit judgment data revealed awareness of a distinction between conventional and unconventional sentences, the online self-based reading time data did not. We had, prior to data collection, expected a slow-down when the learners read the unconventional utterances, as has been found for native Spanish speakers (Tachihara & Goldberg, submitted). The lack of an effect was notable given that the learners had shown some awareness of a distinction between conventional and unconventional sentences even during the initial judgment task, yet no slow-down was evident either initially or after exposure. This implicates a difference between explicit knowledge and online processing of sentence acceptability in classroom language learners (Larsen-Freeman, 2000). Future work is needed to better understand whether to expect a slow-down in reading times for unconventional sentences among language learners of varying proficiency levels, and in various learning contexts.

It is important to note that no explicit feedback was given to participants at any point in the current experiment. While we consider the judgment task at initial and final assessment to be an explicit linguistic task, significant learning occurred from the naturalistic

exposure in which participants read only conventional sentences. Thus, the study demonstrates that, with sufficient exposure, language learners can learn unacceptability without explicit feedback through statistical preemption.

The current results replicate the finding that adult language learners recognize conventional formulations to be fully acceptable long before they fully appreciate the unacceptability of unconventional formulations. This is evident in Figure 2: judgments on unacceptable sentences was lower as proficiency, measured by course level, increased, while judgments on acceptable sentences remained steady across proficiency levels. The current manipulation reveals that adult learners learn to better recognize unconventional sentences as unacceptable, through multiple days of exposure to conventional formulations. These results highlight the fact that learning a new language is, in large part, learning unacceptability, or learning what *not* to say.

References

- Ambridge, B., & Brandt, S. (2013). Lisa filled water into the cup: The roles of entrenchment, preemption and verb semantics in German speakers' L2 acquisition of English locatives. *Zeitschrift für Anglistik und Amerikanistik*, 61(3), 245-263.
- Boyd, J. K., & Goldberg, A. E. (2011). Learning what not to say: The role of statistical preemption and categorization in a-adjective production. *Language*, 55-83.
- Chouinard, M. M., & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of child language*, 30(3), 637-669.
- Dumay, N., & Gaskell, M. G. (2007). Sleep-Associated Changes in the Mental Representation of Spoken Words. *Psychological Science*, 18(1), 35-39.
- Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning & memory*, 13(3), 259-262.
- Gaskell, M.G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, 89, 105-132.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Jegerski, J. (2014). Self-paced reading. In J. Jegerski & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 20-49). New York: Routledge.
- Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 3: 228-238.
- Larsen-Freeman, D. (2000). *Techniques and principles in language teaching*. Oxford University.
- Lindsay, S., & Gaskell, M. G. (2010). A complementary systems account of word learning in L1 and L2. *Language Learning*, 60(SUPPL. 2), 45-63.
- Mattys, S. L., & Clark, J. H. (2002). Lexical activity in speech processing: Evidence from pause detection. *Journal of Memory and Language*, 47(3), 343-359.
- Perek, F., & Goldberg, A. E. (2017). Linguistic generalization on the basis of function and constraints on the basis of statistical preemption. *Cognition*, 168, 276-293.
- 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.
- Tachihara, K., & Goldberg, A. E. (2020). Reduced competition effects and noisier representations in a second language. *Language Learning*, 70(1), 219-265.
- VanPatten, B., & Jegerski, J. (Eds.). (2010). *Research in second language processing and parsing* (Vol. 53). John Benjamins Publishing.
- Zhang, X., & Mai, C. (2018). Effects of entrenchment and preemption in second language learners' acceptance of English denominal verbs. *Applied Psycholinguistics*, 39(2), 413-436.