Learning of gender versus number regularities by native speakers of English

Gabriela Zayas Alom

University of Puerto Rico-Rio Piedras

University of Wisconsin-Madison

**Abstract**

Previous research has shown that people have trouble learning dependencies that their own language does not have. However, this research wasn’t well controlled for properties of the languages and manner of instruction. Using an artificial language paradigm, we investigated whether native English speakers would be better at learning gender or number dependencies. Participants learned an artificial language accompanied by a visual world made up of monsters. We assessed learning of gender and number dependencies by an error monitoring test, which contained gender and number agreement error sentences. Results showed that participants were better at learning gender rather than number dependencies, which was surprising given that English does have number but does not have gender. Further, we explored the role of experience in learning a new language through a language background questionnaire. We found that experience with another language was a significant predictor of how well participants learned these dependencies.

**Introduction**

Learning a second language (L2) can be a difficult, yet not impossible task. People who learn a second language as adults not only have a foreign accent in their new language, they also rarely fully master the grammar of their L2 (Johnson & Newport, 1989). For example, gender (e.g. feminine-masculine, neuter-non-neuter, etc.) and number (e.g. singular-plural, etc.) dependencies are some common grammatical regularities. These grammatical properties are difficult to acquire in an L2 if a speaker’s native language (L1) does not have them.

Several studies have shown that a gender system, such as in Spanish, is difficult to learn for people whose L1 does not have gender. For example, beginning L2 learners of Spanish whose native language did not have gender struggled differentiating a grammatically correct sentence from a grammatically incorrect one (Sagarra & Herschensohn, 2010). On the contrary, L2 intermediate learners of Spanish show sensitivity to grammatically incorrect sentences with regard to gender agreement. This finding suggests that proficiency plays an important role in sensitivity to gender violations in a second language. Thus, while knowledge of grammatical features transfers from L1 to a new L2, there is evidence that adults are capable of acquiring new grammatical features in an L2 that do not exist in their L1 (Sagarra & Herschensohn, 2010).

English second language learners of Spanish also appear to be impaired in their use of helpful gender information. In Spanish the determiner (el/la) is an important cue for native speakers to predict what follows in a sentence. L2 learners are not as effective using these cues (Grüter, Williams & Fernald, 2012). When compared to native speakers, near-native speakers of Spanish do not tend to use these facilitative cues (Grüter et al., 2012). Grüter et al. (2012) argued that native speakers have stronger associations between a noun and its gender than near native speakers do. This explains why advanced L2 learners were sensitive to gender agreement violations and were able to detect mistakes but were slower in retrieving the information during online production (Grüter et al., 2012).

Several researchers developed artificial languages to study how adults learn gender or other complex syntactic relationships. In artificial language studies, participants learn a small, made up language that has some properties that are interesting to investigate. An artificial language paradigm is particularly useful when studying language regularities because it is easier for the experimenter to manipulate and control the dependencies relevant to the research question. Additionally, it provides certainty that participants do not have any prior experience with the target language. Arnon and Ramscar (2011) found in an artificial language learning study that when learning grammatical gender, it is beneficial to first learn multiword units (e.g. determiner-noun, noun-verb-object, etc.) rather than isolated nouns. Moreover, they found that learning multiword units first actually enhanced learning the isolated nouns. In another artificial language study Amato and MacDonald (2010) found that adults were able to implicitly learn complicated contingencies. This suggests that adults are excellent learners of statistical patterns and can learn at least some characteristic dependencies of the language in use.

In the current study, native English speakers will learn an artificial language through a number of different tasks. The novel language that participants learn has gender and number dependencies. The purpose is to determine whether adults are better at distinguishing one of these dependencies than the other. Considering that English is a grammatically ungendered language with number dependencies (singular-plural) we hypothesize that adults should be better at learning number than gender. We will examine this with an error monitoring task. We expect them to be better at detecting number agreement errors than gender agreement errors. We predict this because previous findings (Sagarra & Herschensohn, 2010; Siegelman & Arnon, 2015) support the idea that when learning L2, transfer of properties from L1 occurs, meaning that it should be easier for participants to learn that the artificial language has numerical properties because so does their L1.

**Method**

**Participants**

Thirty native English speakers received payment or course credits for participating in the experiment. Half of the participants were in a production condition and half were in a comprehension condition. The different conditions are not the focus of this study. Exclusion criteria included hearing disabilities and color blindness.

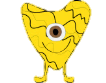
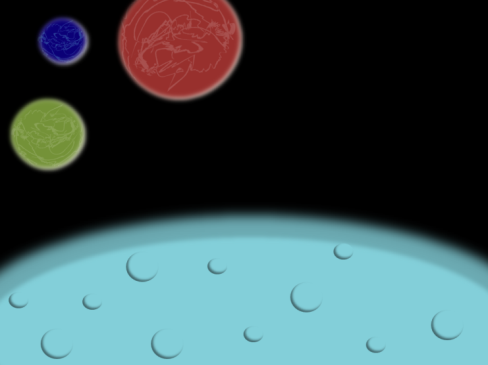
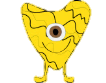
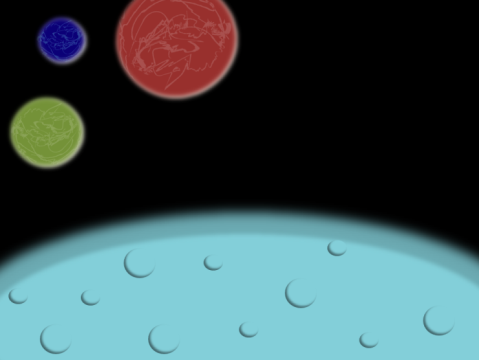
**Materials**

We used an artificial language learning paradigm to assess learning of gender versus number dependencies. The artificial language created consisted of 20 words in total (Table 1; see Appendix A for Tables) and 4 suffixes used in multiple words that expressed both gender and number (see Table 2). Accompanying the language was a visual world that is described by the language (Figures 1 – 5; see Appendix B for Figures and graphs).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Vus** | **fumus** | **teepus** | **traw** | **ot** | **divus** | **kredel** |
| (determiner-singular: one nice monster) | (color adjective-yellow-nice singular) | (monster noun-nice singular) | (pattern-curved lines) | (preposition) | (verb-moving to the right-nice singular) | (location) |

Example of a sentence in the artificial language with the visual world example:

The (singular) yellow monster with curved lines is moving to the right at this location.



There were two types of monsters in the visual world, referred to here as Nice monsters (Figure 1) and Scary monsters (Figure 2). One type of monsters had suffixes with u sounds (-us and –usu), and the other type of monsters had suffixes with o sounds (-ok and –oko). This was randomized for each participant. Some examples of correct and incorrect gender agreement sentences follow:

Correct: Vus fumus teepus traw ot divus kredel.

Incorrect (adjacent): Vus fumus teep**ok** traw ot divus kredel.

Incorrect (non-adjacent): Vus fumus teepus traw ot div**ok** kredel.

Some trials included one monster and others included two. When there were two monsters, they were always the exact same monster twice. Singular monsters had singular suffixes (–us and –ok), and plural monsters had plural suffixes (–usu and –oko). Examples of correct and incorrect number agreement sentences follow:

Correct: Vus fumus teepus traw ot divus kredel.

Incorrect (non-adjacent): Vus fumus teepus traw ot div**usu** kredel.

Incorrect (adjacent): Vus fumus teep**usu** traw ot divus kredel.

Agreement errors could be either adjacent or non-adjacent. Adjacent means that the words that have discord are next to each other (e.g. color –us and monster –ok), non-adjacent means there is one or more words in between the words that do not agree with each other. In the previous examples, there are adjacent and non-adjacent incorrect sentences of both gender and number agreement.

**Procedure**

Experimenters told participants that they were going to learn a new language. They did not provide any other information about the study. The duration of the study was approximately 75 min. The study was divided into passive and active learning trial blocks and three tests at the end of all learning trials to assess participants’ knowledge of the language.

During the passive learning trials participants learned new words. Throughout the experiment, these trials became longer and more complex as new words were added. During the active learning trials participants got to actively practice with new words they learned. Some of the participants had to speak the word out loud and others had to do a picture-sound matching task. Participants in the production condition were asked to describe a picture in the novel language, whereas participants in the comprehension condition had to say whether a picture they saw correctly matched a phrase in the novel language they heard.

After the passive and active learning blocks, participants completed three tests. A vocabulary test and a forced choice test were not the focus of this study and will thus only be briefly described here. The vocabulary test was included to test whether participants knew the meaning of all 18 words of the language. Participants heard a word and saw two pictures of the same type (e.g. two patterns). They had to choose the correct picture. In the forced choice test, participants saw two pictures or videos which differed in only one element of the novel language (different color, pattern, etc.). They heard a phrase or sentence and had to choose the correct picture/video as fast as they could. The forced choice test consisted of 66 trials.

The last test was an error monitoring test. The analyses of this study were focused on these results. In the error monitoring test, participants heard a sentence and had to indicate as fast as possible whether it was correct or incorrect. Participants did not have to wait to hear the last word of the sentence to press the button to indicate if it was correct or not. Once they pressed the button, the sentence stopped playing and the next trial came. There were different types of incorrect sentences within this test however, for the purpose of this study, only the gender and number concord/discord sentences were analyzed. The error monitoring test had 124 trials including other types of errors as well as correct sentences. Specifically, our analysis focused on 48 incorrect sentence trials: 24 gender agreement errors (12 adjacent, 12 non-adjacent) and 24 number agreement errors (12 adjacent, 12 non-adjacent). Dependent variables were reaction time and accuracy.

We predicted that participants would be more accurate on error monitoring trials that contained number disagreement errors than in those that contained gender disagreement errors. Because all of the participants are native speakers of English, and English contains number dependencies, participants should be faster and more accurate identifying number errors in the novel language. Further, we predicted that adults would be faster and more accurate at adjacent than non-adjacent agreement errors because it’s harder to spot the difference on agreement when the error is among words that are closer together than those that are farther apart.

After training sessions and tests, participants completed a language background questionnaire. They were asked if they knew other languages than English and how many years of experience they had with that language (1-3 years, 3-5 years, 5-up years). We predicted that the more years of experience they had with a gendered language, the better they would be at identifying gender agreement errors.

**Results**

The vocabulary test was used as exclusion criterium for the statistical analyses. There were a total of 18 words tested, and based on pilot data, we concluded that participants who scored 15 or higher took the task seriously. Out of 35 participants tested, 5 (14%) did not meet the vocabulary test requirement. The data analysis below is thus based on the data of the remaining 30 participants who scored 15 or higher on the test.

The forced choice test had 66 trials and the vocabulary test had 18 trials. Participants performed well on the forced choice (*M* = 59.93, *SD* = 6.78) and the vocabulary tests (*M* = 16.97, *SD* = 1.09), meaning that by the end of the study they had learned these aspects of the novel language.

Figure 6 shows the mean of correctly identified agreement errors per category in the error monitoring test. In the gender-adjacent category participants correctly identified more agreement errors (*M* = 9.97, *SD* = 2.61) than in the number-adjacent category (*M* = 8.73, *SD* = 3.36). However, participants did not greatly differ in scores when identifying gender-non-adjacent errors (*M* = 7.37, *SD* = 4.11) and number-non-adjacent errors (*M* = 7.57, *SD* = 3.42).

We ran a mixed effects logistic regression analysis on the data to test whether adjacency, gender-number and their interaction significantly predicted correctly identifying agreement errors. Contrary to hypothesis, the factor gender-number was not a significant predictor of whether participants correctly identified the agreement errors (*p* > 0.05). There was no overall difference in how well English native speakers in our experiment were able to detect gender versus number agreement errors. As expected, the factor adjacency was found to be a significant predictor of whether participants correctly identified agreement errors (*p* < 0.001). Participants had more difficulty to identify agreement errors when they were on the verb, non-adjacent to the noun phrase. Additionally, there was a significant interaction between two factors (*p* < 0.01): Participants identified more gender than number adjacent agreement errors, whereas they identified just as few gender as number non-adjacent agreement errors.

Figure 7 shows the mean number of agreement error sentences correct per participant sorted by years of experience with another language based on the language questionnaire. The results show that having more years of experience with another language is a significant predictor (*p* < 0.001) of how well participants do in the gender and number error agreement trials. Thus, the more experience participants had with learning another language, the more agreement errors they were able to identify in our task.

**Discussion**

This study explored whether native English speakers would be better at learning number dependencies than gender dependencies in an artificial language study. We found that there was no significant difference between correctly identified agreement errors in incorrect gender and number sentences. Thus, participants did not identify number dependencies better than gender ones. This unexpected result may be due to the fact that the difference between suffixes in sentences with a gender agreement error (–us vs –ok or –usu vs –oko) were easier to notice than the difference in suffixes in sentences with a number agreement error (–us vs –usu or –ok vs –oko; see also example sentences in the Method section). Previous research has shown that learning a gendered L2 is a very difficult task for native speakers of an ungendered language (Sagarra & Herschensohn, 2010). Having gender suffixes that were easier to distinguish than number suffixes may have given native English speakers an advantage for spotting gender agreement errors in this experiment.

Another possible explanation for these findings may be that experience with a gendered L2 increases sensitivity to agreement violations (Sagarra & Herschensohn, 2010). The questionnaire that participants completed at the end of the experiment indicated that most of our participants had at least some experience learning a gendered L2 in school. Our analyses show that amount of experience with a gendered L2 was a significant predictor of how well people generally did in our task. Thus, more experienced language learners seem to be better at catching agreement errors (Sagarra & Herschensohn, 2010). Our analysis does not show an interaction of language experience with the factor gender/number: thus, participants with more language learning experience are generally better at our task but not specifically better at gender. However, in our sample experience learning gendered languages was completely confounded with general language learning experience, as virtually everyone in our sample who learned a second language learned a gendered language like Spanish or German.

Adjacent agreement errors were significantly easier for participants to detect than non-adjacent agreement errors. This was expected because non-adjacent agreement errors, where the incorrect suffix is not right next to a word with a correct suffix, are harder to spot than adjacent errors where the incorrect suffix is right next to a word with a correct suffix. Additionally, we found that the interaction of adjacency and gender-number was significant, meaning that when the error was gender-adjacent it was easier to catch than number-adjacent. Non-adjacent errors of both types were equally hard to catch.

Future research could further explore the role of experience in identifying gender or number agreement errors. Specifically, it would be interesting to tease apart the helpful role of general language experience from experience with gendered languages. In addition, other variables (e.g. RT) should be taken into consideration for future analyses of this data. We might expect participants with experience with another language to be faster at identifying gender and number agreement errors. Furthermore, it will be interesting to see whether the RTs show the same gender-number with adjacency interaction that the accuracy data analyzed here do.

To conclude, we found that language learning experience is a significant predictor of learning of gender and number dependencies. We also found an expected difference between ease of identifying adjacent and non-adjacent agreement errors. Surprisingly, we did not find a difference between learning of gender and number dependencies, though this factor did interact with adjacency but not in the expected direction. Thus, this study provides preliminary evidence to suggest that gender need not be harder to learn than number in well controlled circumstances like in our artificial language learning experiment. We will continue to explore these questions in future analyses on the bigger dataset that this study is part of.

**References**

Amato, M. S., & MacDonald, M. C. (2010). Sentence processing in an artificial language: Learning and using combinatorial constraints. *Cognition, 116*(1), 143-148. doi:10.1016/j.cognition.2010.04.001

Arnon, I., & Ramscar, M. (2011). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition, 122*, 298-305. doi:10.1016/j.cognition.2011.10.009

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research, 28*(2), 191-215. doi:10.1177/0267658312437990

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Sagarra, N., & Herschensohn, J. (2010). Proficiency and animacy effects on L2 gender agreement processes during comprehension. *Language Learning, 61*(1), 80-116. doi:10.1111/j.1467-9922.2010.00588.x

Siegelman, N., & Arnon, I. (2015). The advantage of starting big: Learning from unsegmented input facilitates mastery of grammatical gender in an artificial language. *Journal of Memory and Language, 85*, 60-75. doi:10.1016/j.jml.2015.07.003

**Appendix A: Tables**

Table 1: Words of the artificial language

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Determiner** | **Color Adjective** | **Monster noun** | **Markings** | **Preposition** | **Verb** | **Location** |
| V- | Fum- | Teep- | Traw | Ot | Div- | Kredel |
|  | Saf- | Zout- | Chag |  | Pov- | Chaftem |
|  |  | Weem- | Plim |  | Zev- | Hullem |
|  |  | Mog- | Stam |  |  |  |
|  |  | Ket- |  |  |  |  |
|  |  | Pex- |  |  |  |  |

Table 1.Words of the artificial language categorized by their grammatical function in the sentence. The order in which the words appear is the correct grammatical order of a sentence in the language. The words that end with a hyphen (-) always carry a suffix (see Table 2).

Table 2: Suffixes of the artificial language

|  |  |  |
| --- | --- | --- |
|  | **Number** | |
| **Gender**  **(Monster type)** | Singular | Plural |
| Nice/Scary | -us | -usu |
| Nice/Scary | -ok | -oko |

Table 2*.* Four suffixes that accompany the determiner, color adjective, monster noun and verb in the artificial language.

**Appendix B: Pictures and graphs**

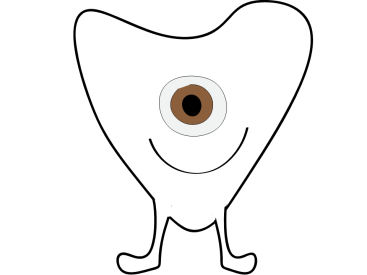
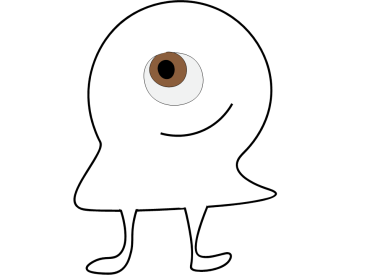
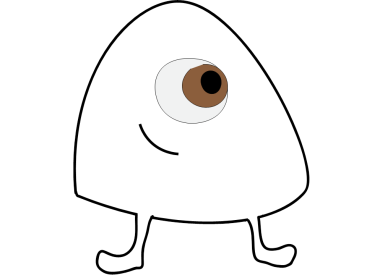


Figure 1*.* Nice monsters of the artificial language. These monsters are single-eyed, smiling, two-legged and have round shapes.

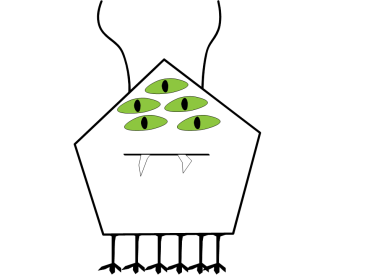
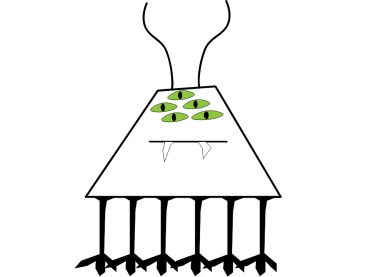
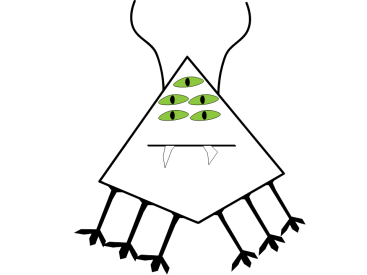


Figure 2*.* Scary monsters of the artificial language. These monsters are five-eyed, have pointy shapes and have six legs and two antennas.



Figure 3.Two possible colors (yellow or purple) that any monster can have.

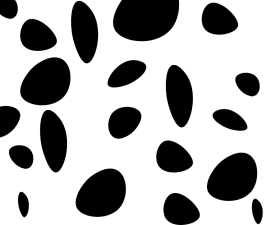
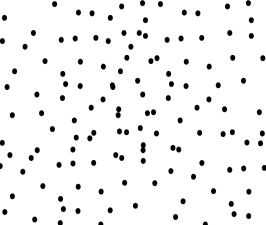
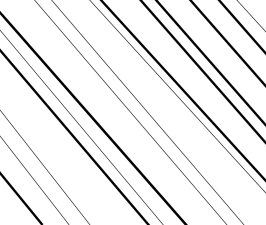
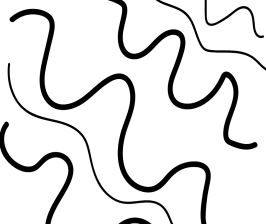


Figure 4*.* Possible patterns that monsters can have on their skin.

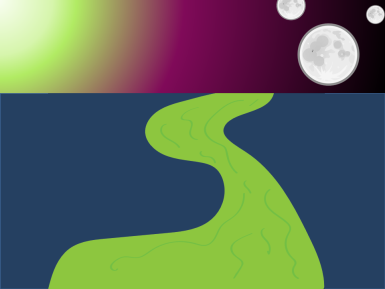
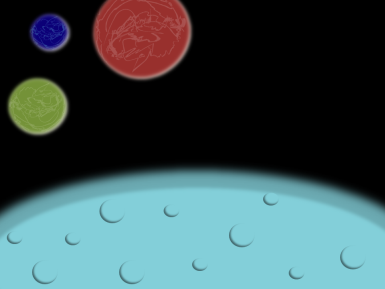


Figure 5*.* Three possible locations where the monsters can be.

Figure 6. Mean of correctly identified agreement errors per type. The error bars represent the standard error of the mean.

Figure 7. Mean of correctly identified agreement errors per years of experience learning a second language. The error bars represent the standard error of the mean.