Syntactic adaptation in adults during online eye-tracking: Replication of Havron et al. (2019)

Background

Havron, de Carvalho, Fiévet, & Christophe (2019) demonstrated that both adults and children showed signs of rapid syntactic adaptation after repeated exposure to a particular sentence structure. Furthermore, participants drew on these expectations to interpret unfamiliar words that were presented in the same syntactic context. We will first briefly review the motivation for the original study before explaining the reasoning behind this replication.

Havron et al.'s work drew on two lines of research in the literature. First, numerous studies have shown evidence for an effect of syntactic priming in both children and adults. Peter et al. (2015) and Rowland et al (2012) found that exposure to either double object datives or prepositional datives primed production of the corresponding syntactic structure for adult and child participants. In addition to priming production of particular structures, evidence from fMRI (Noppeney & Price, 2006) and from ERP (Ledoux, Traxler, & Swaah, 2007) suggests that syntactic adaptation influences comprehension of later sentences. Using a self-paced reading paradigm, Fine et al. (2013) demonstrated that syntactic expectations are updated rapidly during comprehension, and that these expectations can reduce or undo the normal effect of garden-path sentences. Finally, Thothathiri & Snedeker (2008) found effects of syntactic priming on children's comprehension of sentences containing datives, even when the lexical content of the sentences did not match. They used eye-tracking to measure participants' comprehension, which is the same method that was employed in this experiment.

Second, research has established that syntactic bootstrapping allows children to infer the meanings of unfamiliar words based on their syntactic characteristics (Gleitman, 1990). For

example, upon hearing a sentence such as “It’s daxing,” a child can infer that *dax* refers to an action and is therefore a verb. Evidence that children are sensitive to syntactic context has been found with infants as young as 14 to 18 months (He & Lidz, 2017), as well as 24-month-olds (Waxman et al., 2009). Using video stimuli that were similar to those in our study, Bernal et al. (2007) reported that 23-month-old infants used syntactic bootstrapping to infer the meanings of novel words, based on their ability to correctly point to the correct video portraying an action or an object. Although Rabagliati, Gambi, & Pickering (2016) point out that there may not yet be enough evidence to conclude that children use their predictions to learn new words, Huang & Arnold (2016) found that children who were more sensitive to syntactic cues had more accurate interpretations of new words.

Havron et al. (2019) brought together these two lines of research to ask whether priming French children with a particular syntactic structure would influence how they assigned meaning to unfamiliar words in a syntactically ambiguous context. They found that, indeed, children and adults appeared to alter their predictions about which syntactic structure a speaker would use, and draw on these predictions to infer the meaning of a novel word. This replication builds on their work by testing whether these results will replicate in a new context: an eye-tracking study conducted entirely online. Replicating the original findings with adults using online eye-tracking, as opposed to a lab study, is a first step towards verifying that the results will transfer to other contexts; we suggest some possible future directions in the discussion section. As in the original study, we hypothesized that when participants encountered an unfamiliar word, they would look more at the video (action or object) that matched the type of phrases (verb or noun) they had heard during training trials. Our preregistered design and analysis can be found on the Open Science Framework (OSF) at: https://osf.io/wza3k.

Method

Participants

We collected data from 77 participants (31 female; 46 male) using Prolific. All were adults who reported speaking French as their first language. Participants were randomly assigned to a condition (37 in the noun condition; 40 in the verb condition). They were told that the study had originally been designed for children to explain the child-friendly nature of the videos.

Measures

We conducted online eye-tracking using Webgazer, a program that runs a participant's webcam and records the coordinates of their eye movements. Participants needed a laptop or desktop computer with a screen width of at least 1280 pixels to complete the study. We followed the common practice of only analyzing looks that were to either the action video or the object video (57% of the total looks in the dataset). In the analysis, we examine only looks to the action video, since any look not to the action video is to the object video. All stimuli, data, and analyses can be found at: https://github.com/eswanson166/havron-replication.

Procedure

The stimuli used in the study, as well as the structure of the trials, were identical to those used in Havron et al. (2019). Every participant was randomly assigned to either the noun condition or the verb condition. Participants first completed a 9-point calibration, which was adapted from the original study to work with WebGazer. The study consisted of two phases: training trials and test trials. The total experiment included ten trials and lasted about twelve minutes.

During each training trial, all participants saw two videos. One showed a girl performing a familiar action (such as jumping), while the other was of the same girl holding a familiar object (such as a toy car). The structure of each training trial was identical. First, the participant saw a preview of one video only, followed by a preview of the other video. Then, during the contrast phase, the participant saw both videos together. For these parts of the trial, a female narrator told the child to look at the videos in a child-friendly voice, but she did not comment on what the videos depicted. The last part of the trial was the event phase, during which children saw both videos again, but the narrator described what was in just one of the two videos. If participants were in the noun condition, she said a phrase such as *La petite grenouille* ("The little frog"). If participants were in the verb condition, she said a phrase such as *La petite dort* ("The little one (fem.) is sleeping"). Thus, participants in both conditions heard the same syntactic context: *La petite [X],* but it was followed by either a noun (meaning "The little X") or a verb (meaning "The little one is Xing"). Participants were exposed to four training trials. The side the target video appeared on was counterbalanced, and the order of the training trials was randomized.

In between the first two training trials and the last two training trials, participants watched two filler trials. These trials had the same structure as the training trials except that the narrator referred to whichever type of video was *not* referred to in the training trials, using a structure that was unambiguous. Therefore, participants in the noun condition heard a description of the action video in a sentence such as *Elle écrit* ("She writes"), since *Elle…* cannot be followed by a noun. Similarly, participants in the verb condition heard a description of the object video in a sentence such as *C'est une poussette* ("This is a baby-stroller"), because *C'est une…* cannot precede a verb. These filler trials were included so that participants would understand that the narrator could refer to either the action video or the object video. It was simply with the

structure *La petite…* that the narrator was biased toward using either nouns or verbs. Also, we wanted to reduce the possibility that participants would look toward the action or object video on test trials purely because they were used to looking at that type of video.

After the training trials, all participants watched three test trials, which were identical regardless of condition (though the order was again randomized). Test trials had the same structure as training trials, but the two videos depicted a novel object and a novel action. Also, participants heard the narrator's description once before the event phase started so that we could start measuring looks from the beginning of the event phase. The narrator used the same *La petite…* context as before, but it was followed by an unfamiliar word that does not actually exist in French, such as *La petite nuve.* Since *La petite…* can be followed by a noun or a verb, participants could in principle interpret *nuve* as a noun or a verb. However, our hypotheses predicted that participants would adapt to the structure preferred by the speaker during training trials. Therefore, participants in the noun condition were expected to interpret novel words as nouns, and therefore look more at the object video during test trials. Conversely, participants in the verb condition were expected to interpret novel words as verbs, and therefore look more at the action video during test trials. A key linking hypothesis was that more looks to a video indicated that participants interpreted the word as matching what was depicted in the video.

As in Havron et al. (2019), we also included one generalization trial at the end of the experiment. On this trial, participants saw a video with a boy in it and heard the sentence, *Le petit pirdal. Le petit* is the masculine form of *La petite*, so it can similarly be followed by a noun, meaning "The little X," or a verb, meaning "The little one (masc.) is Xing." This was an exploratory trial intended to examine whether participants would generalize the predictions formed during training trials to a slightly different structure with the same underlying syntax.

Results

Test Trials

We calculated each participant's proportion of looks to the action video in each test trial and then averaged these three proportions to obtain each participant's mean proportion of looks to the action video across the three test trials. Figure 1 shows confidence intervals for the overall mean proportion of looks to the action video in each condition, as well as dots representing individual participants' mean proportions of looks. As we hypothesized, participants in the verb condition (M = 0.58, SD = 0.16) appeared to have a higher proportion of looks to the action video than participants in the noun condition (M = 0.41, SD = 0.17).

To confirm this, we conducted two main analyses: a mixed effects linear regression (the same as in the Havron et al. study) and a mixed effects logistic regression analysis. We also preregistered a cluster-based permutation analysis, which is forthcoming. The lme4 package was used to conduct the regression analyses, and the reported p-values come from likelihood ratio tests of the described model compared to the model without the effect of interest.

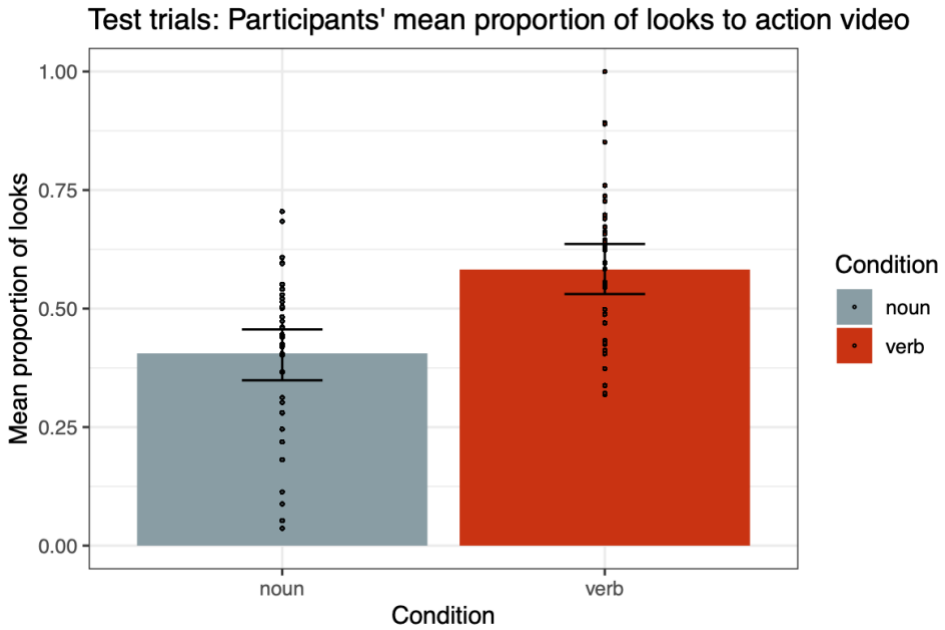

Figure 1. Confidence intervals for the mean overall proportion of looks to the action video in the noun and verb conditions during test trials. Black dots correspond to the mean proportion of looks for individual participants, averaged across the three test trials.

In the mixed effects linear regression, we predicted a participant's arc-sin transformed mean proportion of looks to the action video as a function of condition, with a random intercept for participant. We did not have a random intercept for item because there were only three test items. Also, condition was centered, because it exhibited high collinearity with the intercept. There was a significant main effect of condition in the direction expected: Participants in the verb condition had a significantly higher proportion of looks to the action video than participants in the noun condition (β = 0.20, SE = 0.05, t = 4.39, p < 0.001).

The mixed effects logistic regression predicted the log odds of looking to the action video as a function of condition and previous look (to the action video or not). It included a random intercept for participant and a random slope that accounted for participant differences in the effect of previous look. We centered condition and previous look due to high collinearity. There was a significant main effect of condition (β = 0.68, SE = 0.15, z = 4.52, p < 0.001), such that participants in the verb condition were more likely to look at the action video. There was also a significant main effect of previous look (β = 4.35, SE = 0.20, z = 21.65, p < 0.001) such that if a participant looked at the action video on their previous look, they were more likely to look at the action video on the following look as well.

Training Trials

We were also interested in participants' eye movements on training trials. We wanted to confirm that participants did, in fact, look at the video described during training trials in order to ensure that (a) the eye-tracker was reliably measuring their looks and (b) participants were reacting to the descriptions they heard.

In Figure 2, we present a visualization of participants' mean proportion of looks to the action video during training trials. We see that compared to the test trials, the training trials show an even larger difference between the noun condition (M = 0.31, SD = 0.14) and the verb condition (M = 0.72, SD = 0.14). As expected, participants in the verb condition looked more to the action video than participants in the noun condition. This result is unsurprising given that the training trials used known words that clearly referred to one of the two videos.

It is also important to ensure that the filler trials were effective, so participants would understand that the narrator could use a different linguistic structure to refer to the video that did not match the participant's condition. Figure 3 illustrates participants' mean proportion of looks to the action video during filler trials. It confirms that the pattern of results from test and training trials was switched, as intended: Participants in the noun condition (M = 0.70, SD = 0.19) looked more to the action video than participants in the verb condition (M = 0.34, SD = 0.18).

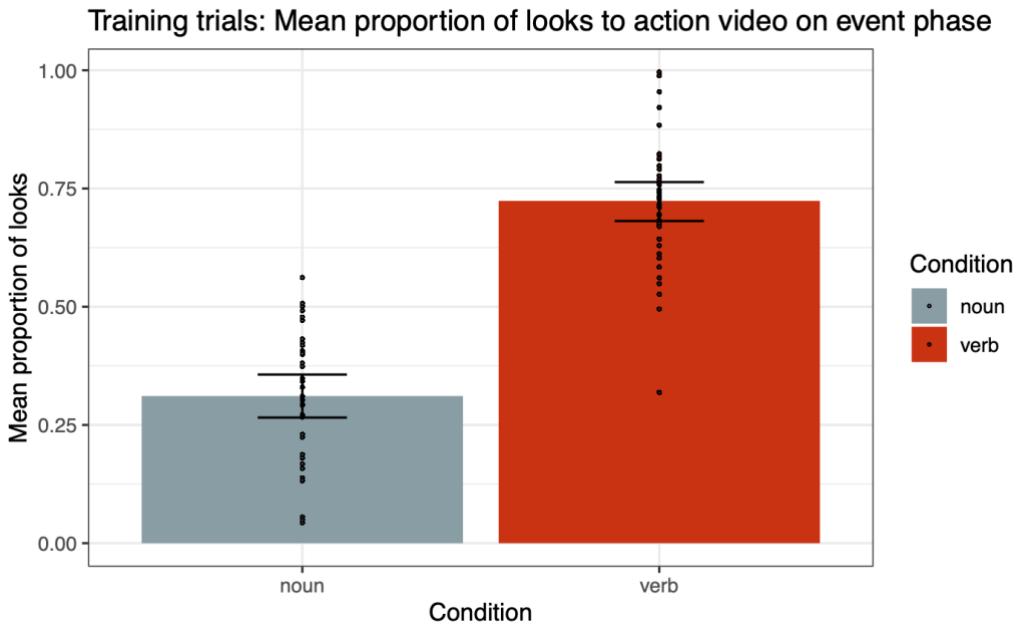

Figure 2. Confidence intervals for the mean overall proportion of looks to the action video in the noun and verb conditions during training trials. Black dots correspond to the mean proportion of looks for individual participants, averaged across the four training trials.

Figure 3. Confidence intervals for the mean overall proportion of looks to the action video in the noun and verb conditions during filler trials. Black dots correspond to the mean proportion of looks for individual participants, averaged across the two filler trials.

Discussion

This study replicated the adult results of Havron et al. (2019), which examined how syntactic priming influenced word learning. The original study found that participants adapted to a syntactic structure after hearing it repeatedly and that they used their adjusted expectations to interpret an unfamiliar word. Our results were similar: A significant effect of condition was found with both the mixed effects linear regression model and the mixed effects logistic regression model. That is, compared to participants who heard *La petite (noun)* on training trials, participants who heard *La petite (verb)* looked significantly more at the action video on test trials. This suggests that participants were interpreting the ambiguous words on test trials to be consistent with the syntactic structure (noun vs. verb) that had previously been used by the narrator.

The difference we found between conditions appears to be smaller than in the original paper. Havron et al. (2019) found a mean proportion of looks of 0.653 in the verb condition (compared to our 0.58) and 0.275 in the noun condition (compared to our 0.41); the size of the standard deviations was similar. However, given that this was a replication and that it used online eye-tracking, which is noisier than an in-lab eye-tracking device, this observation is not surprising. Though the difference was smaller, the effect of condition was still significant.

It is encouraging that the results of the original paper replicated in an online eye-tracking context, where there is less direct control over participants' experience during the study. Furthermore, the fact that participants were quite clearly looking at the expected videos during both training and filler trials, when it was obvious which video was being described, lends support to the idea that online eye-tracking results can be fairly reliable, despite the somewhat high rate of track loss. Although calibration can sometimes be difficult, making it take longer to recruit the desired number of participants than it does with other types of studies, we conclude that it is feasible to conduct studies online using WebGazer.

After replicating the results of Havron et al. (2019), there are several questions about syntactic adaptation and word learning that merit further exploration, either with in-lab studies or with online eye-tracking. First, is this adaptation speaker-specific? Since we used the same narrator for all of the trials, we do not know whether participants' expectations are person-specific or whether they would generalize to a different speaker. In addition, we used the same linguistic context, *La petite…*, for the entirety of the experiment, except for one exploratory trial at the end that used *Le petit....* A visualization of the data for this trial is included as an appendix, but further investigation is needed to see whether the adaptation observed in this study would generalize to different lexical content that had the same underlying syntactic structure as the

training trials. Finally, crosslinguistic experiments, as well as experiments with other types of sentence structures, would allow us to confirm that this phenomenon is not unique to French or to the *La petite…* structure chosen for this work. Studies that investigate all these questions will be essential in order to determine the extent to which syntactic adaption is a mechanism that guides word learning in the real world.

References

Bates, D., Maechler, M., Bolker, B. and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01.

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development, 3*, 325–341. doi: 10.1080/15475440701542609

Fine, A.B., Jaeger, T.F., Farmer, T.A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS One, 8*(10). doi: 10.1371/journal.pone.0077661

Havron, N., de Carvalho, A., Fiévet, A.-C., & Christophe, A. (2019). Three- to Four-Year-Old Children Rapidly Adapt Their Predictions and Use Them to Learn Novel Word Meanings. *Child Development, 90*(1), 82-90. doi: 10.1111/cdev.13113

He, A. X., & Lidz, J. (2017). Verb learning in 14-and 18-month-old English-learning infants. *Language Learning and Development, 13*, 335–356. doi: 10.1080/15475441.2017.1285238

Huang, Y. T., & Arnold, A. R. (2016). Word learning in linguistic context: Processing and memory effects. *Cognition, 156*, 71–87. doi: 10.1016/j.cognition.2016.07.012

Ledoux, K., Traxler, M.J., & Swaab, T.Y. (2007). Syntactic priming in comprehension: Evidence from event-related potentials. *Psychological Science, 18*(2), 135-143. doi: 10.1111/j.1467-9280.2007.01863.x

Noppeney, U. & Price, C.J. (2006). An fMRI Study of Syntactic Adaptation. *Journal of Cognitive Neuroscience, 16*(4), 702-713. doi: 10.1162/089892904323057399

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects

in a structural priming task. *Journal of Memory and Language, 81*, 1–15. doi: 10.1207/s15327817la0101_2

R Core Team. (2017). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. www.R-project.org/

Rabagliati, H., Gambi, C., & Pickering, M. J. (2016). Learning to predict or predicting to learn? *Language, Cognition and Neuroscience, 31*(1), 94–105. doi: 10.1080/23273798.2015.1077979

Rowland, C.F., Chang, F., Ambridge, B., Pine, J.M., & Lieven, E.V.M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition, 125*(1), 49-63. doi: 10.1016/j.cognition.2012.06.008

Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition, 108*(1), 51–68. doi: 10.1016/j.cogni tion.2007.12.012

Appendix

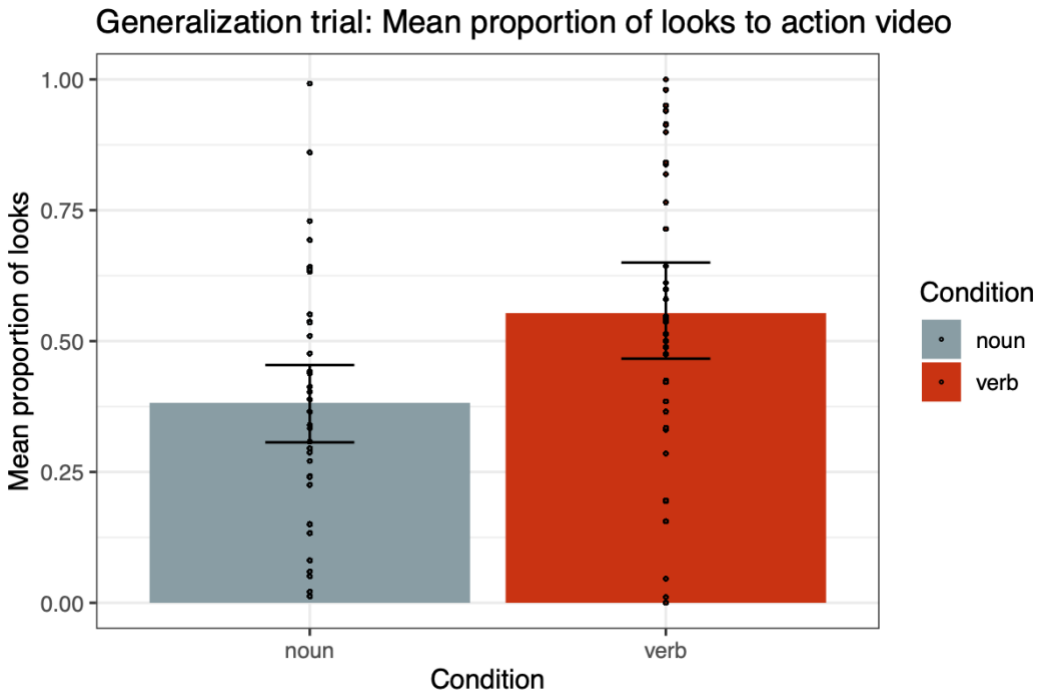

Figure 4. Confidence intervals for the mean overall proportion of looks to the action video in the noun (*M* = 0.38, *SD* = 0.23) and verb (*M* = 0.55, *SD* = 0.29) conditions during the generalization trial. Black dots correspond to the proportion of looks for individual participants. The confidence intervals do not overlap, though they are extremely close (noun: 0.31—0.45; verb: 0.47—0.65). Further investigation is needed to determine whether syntactic adaptation would guide word learning in when the lexical content is different from the original but has the same underlying syntactic structure.