Language priming in the wild: Using a mobile-app for online data-collection

Heyselaar, E.[a], Wheeldon, L.[b], Gilani, S.[c], Segaert, K.[c,d]

a. Communication & Media, Radboud University, Nijmegen
b. Institutt for fremmedspråk og oversetting, Universitetet Agder
c. School of Psychology, University of Birmingham
d. Centre for Human Brain Health, University of Birmingham

Corresponding author:

Evelien Heyselaar

Thomas van Aquinostraat 4,
6525GD Nijmegen
The Netherlands

evelien.heijselaar@ru.nl

Abstract

Syntactic priming paradigms are widely used to investigate language production, yet they are almost exclusively implemented in controlled laboratory settings. This limits scalability, inclusivity, and the study of language behavior in everyday contexts. Here, we present and validate a smartphone-based method for measuring syntactic priming. Our mobile application implements a picture description task commonly used in laboratory studies and incorporates automated speech transcription and syntactic classification. Using our app, we collected data on the production of active and passive sentence structures and evaluated short-term, cumulative, and long-term (across sessions) priming effects. To assess methodological validity, we compared the patterns observed in app-collected data with those obtained previously using an equivalent laboratory paradigm. Results demonstrate that the mobile app reliably elicits short-term syntactic priming, with increased production of passive structures following passive primes, replicating well-established laboratory effects. Within-session analyses further revealed cumulative priming, indicating gradual learning of infrequent syntactic structures. Long-term priming over a one-week interval was not observed in the app data, in contrast to laboratory findings. However, a follow-up experiment with a 24-hour interval between sessions yielded significant long-term priming effects, suggesting that retention of syntactic learning may decay more rapidly in uncontrolled, real-world environments. Together, these findings establish the mobile application as a valid and scalable method for studying syntactic priming outside the laboratory. The approach enables large-scale data collection in naturalistic settings and provides a flexible methodological framework for investigating individual differences in large diverse samples, and potential intervention applications in language research.

Introduction

The field of psycholinguistics has a long tradition of syntactic priming investigations, testing a variety of syntactic structures (Bernolet et al., 2016; Bock, 1986; Bock and Griffin, 2000), in different modalities (Branigan et al., 1999; Holly P Branigan et al., 2000; Hartsuiker et al., 2008), languages (Bock, 1986; Hartsuiker and Kolk, 1998; Sung, 2015), dialogue settings (Branigan et al., 2000), and populations (Branigan and Messenger, 2016a; Hardy et al., 2020). Syntactic or structural priming is the facilitation of syntactic processing that occurs when a syntactic structure is repeated across consecutive sentences. This facilitation can manifest itself in several ways; one of those is the tendency for speakers to reuse the same syntactic structure (e.g., a passive sentence) across sentences. For example, one could hear a conversation partner produce a passive sentence and subsequently re-use this structure in one of their own sentences. Syntactic priming has been modelled as a process that occurs, at least in part, due to an implicit learning effect (Chang et al., 2006; Heyselaar et al., 2020; Jaeger and Snider, 2013), raising the possibility that is could be employed as a tool for language learning or language maintenance. However, unlocking this potential requires broader access to the methodology that goes beyond the time and location constraints of the laboratory. In this paper, we describe and validate a smartphone application to enable syntactic priming to be integrated into daily life.

Before we provide more information on the smartphone application itself, we will elaborate on the link between syntactic priming and implicit learning. Our language system reuses sentence structures (i.e., priming), which aids language production, especially the production of less frequent (usually more complex) sentence structures. For example, the more common active structure ("the girl chases the boy") versus the more complex and less frequent passive structure ("the boy is chased by the girl"). Speakers are unaware of the effects of priming on the sentence structures they produce, i.e., this learning is implicit. Multiple accounts of language production propose that the priming of sentence structures relies on implicit learning. Although they differ in the exact mechanism proposed to explain this learning (Chang et al., 2006; Jaeger and Snider, 2013; Reitter et al., 2011), all involve a mechanism where a speaker's structural preference is updated based on prior experience. These 'updates' are never lost, even with time or other information intervening (Branigan et al., 2000; Kaschak, 2007), allowing a build-up over time with repeated exposures. Structural priming may thus play an important role in language acquisition, language learning, and language change (Chang, 2008). The repetition of recently heard or produced syntactic structures may serve to strengthen representations of a given structure, aiding in the language acquisition process (Tomasello, 2006), as demonstrated in children (Branigan and Messenger, 2016b). Priming may also alter the long-term probability of using a given syntactic structure, thereby contributing to language change across the lifespan (e.g., Chang, 2008; Dell and Chang, 2014), as demonstrated in older adults (Heyselaar and Segaert, 2022).

However, the use of priming for language learning has been established in experimental settings (e.g., Branigan and Messenger, 2016; Heyselaar and Segaert, 2022). Although online methods of collecting data (e.g. Gorilla) have been employed, these methods are geared towards replicating a controlled lab environment in someone's home. To increase the potential for priming to contribute to language learning more generally, there is a need to be able to use priming as part of daily life, in different environments, and in harder to reach populations.

Firstly, priming via a mobile app (rather than in an experimental laboratory setting) may facilitate engaging a larger sample of participants. In several subfields of psychology, there has been a call to

understand individual variability, for example variability in the language profile characteristics that confer advantages of being bilingual (Bialystok, 2016; DeLuca et al., 2020). To answer these types of research questions, large datasets are required. With respect to structural priming, it is established that overall speakers indeed show structural priming effects but that the degree to which speakers are primed is modulated by the wide range of individual differences, e.g. inpersonality characteristics and even social opinions(Schoot et al., 2016; Segaert, 2018). In other words, while the structural priming effect is generally robust, any study exploring modulations of the effect should carefully consider power (Mahowald et al., 2016).

Second, research on priming using a mobile app may enable more diverse participant sampling, benefiting inclusivity. It is notoriously difficult to engage participant samples that are representative of the wider populations in a local area or country, for example with respect to ethnicity and socioeconomic status. On a wider scale, it is known that most research participants in research across the globe come from very narrow subsets of people, i.e. samples drawn from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies (Henrich et al., 2010), which biases findings and limits external validity.

Third, priming via a mobile app is more accessible to a broader population of potential users than laboratory-based research. One example of a more accessible application in the wider population would be older adults. We have previously demonstrated that continued exposure to complex syntactic structures is critical for maintaining syntactic diversity in older age (Heyselaar and Segaert, 2022). However, language spoken to older adults in social interactions is observed to be modified, such that they receive a more simplified linguistic register (Kemper, 1994). More frequent exposure to diverse and complex syntactic structures could have beneficial effects on older adults' syntactic language use. This exposure could be delivered through a mobile app. Commercially available "brain training apps" have become increasingly popular in the wider population, including older adults (Bonnechère et al., 2020; Tetlow and Edwards, 2017). Older adults have higher expectations of brain training exercises than younger adults (Rabipour and Davidson, 2015) and could potentially benefit most from their use, if effective. Therefore, a mobile app that uses priming with the aim to increase the use of infrequent (more complex or harder to use) structures would be an accessible way to remedy an aspect of language decline that older adults are known to struggle with (Kemper et al., 2001b, 2001a). Moreover, whereas some online methods of collecting data (e.g. Gorilla) have been used previously to collect priming data, such methods are geared towards replicating a controlled lab environment in someone's home. Our present approach differs in that the app would be free, accessible, and designed as a fun game integrated into daily life.

Finally, our study directly addresses existing uncertainty about whether syntactic priming even exists outside a controlled laboratory setting. Hence, the aim of the present work is to 1) establish whether structural priming exists when measured using a mobile app in the real world and 2) whether the effects are comparable to those established in laboratory settings. For the laboratory setting, we use structural priming data previously collected and reported in Heyselaar and Segaert (2022), establishing priming effects which were cumulative and long-lasting. We then use a similar structural priming protocol in a mobile application to examine whether we can elicit learning of infrequent structures within a session (short-term priming) but also whether this learning is robust enough to be retained over time (long-term priming). Analyzing the cumulative priming of passive structures allows us to additionally investigate the rate of learning between and within the sessions.

Experiment 1

Methods

Participants

All participants were aged 18 – 30 years, monolingual English speakers with normal or corrected-to-normal vision. Participants were *a priori* excluded from participating if they indicated any neurological deficits, psychiatric disorders, dyslexia or any other language disorders. This study has been approved by the research ethics board of the University of Birmingham.

App. 39 participants (M_{Age}: 25.7 years, SD_{Age}: 3.47; 14 males, 1 non-binary) were recruited via Prolific. In addition to the requirements described above, participants had to be residents of the United Kingdom and have an android phone to participate. Participants were paid GBP7 per hour for their participation.

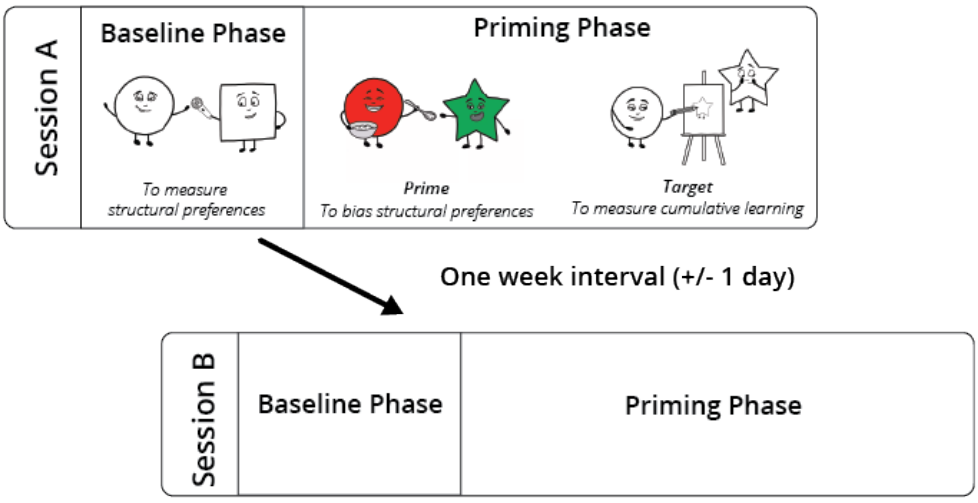
Lab. For the lab data, we used a subset of the data collected, analyzed, and reported for Heyselaar & Segaert (2022). In the present comparison of lab vs. app data, we only use lab data of young participants (more specifically, Session A and Session B of experiment 1). The labdata used were therefore collected from 29 participants (M_{Age}: 19.4 years, SD_{Age}: 0.82; 2 Male) recruited from the Research Participation Scheme of the University of Birmingham. Participants received university credits for their participation.

Study Design and Task

The design is based on that reported in Heyselaar & Segaert (2022). All participants completed two test sessions, separated by a week (+/- 1 day). Each test session consisted of a baseline phase and a priming phase (Figure 1).

Figure 1

Overview of the experiment design



Baseline Phase. In the baseline phase, participants were asked to describe pictures. During each trial two pictures were presented: A picture eliciting an intransitive sentence followed by a picture eliciting a

transitive sentence (actives or passives). These baseline trials allowed us to measure the baseline preference for active and passive structures. There were 20 baseline trials in total. In Session A, the baseline structure preference was measured before any influence from the priming phase. In Session B, the structure preference was measured again to examine whether the priming phase of the previous session (1 week earlier) had biased the structure preference, indicating *long-term priming*.

Priming Phase. In the priming phase, participants completed a priming experiment, similar to Segaert et al. (2011). Prime and target picture were alternated. Prime pictures were always coloured, followed by a target picture that was always grayscale. Participants were instructed to describe pictures with one sentence, naming the green actor before the red actor if the actors were depicted in colour. This allowed us to manipulate, for the primes which were always colour-coded, whether the participant produced an active (e.g. App: *the square is kissing the triangle;* Lab: the man is kissing the woman) or passive (e.g. App: *the triangle is being kissed by the square;* Lab: the woman is being kissed by the man*)* syntactic structure. For the grayscale target pictures, we then measured which structure (active or passive) the participants chose to produce.

For the app experiment, there were 40 passive priming trials (a passive picture followed by a transitive grayscale target), and 40 active priming trials (an active prime followed by a transitive grayscale target), all randomized in one experimental session. We also included 20 filler trials (an intransitive picture followed by an intransitive grayscale target). In total therefore, we had 100 trials consisting of 80 transitive and 20 intransitive pictures. In the priming phase, we measured whether the proportion of passives produced increases from the start to the end of the experimental session, indicating *cumulative priming*.

The trial types are similar to those in the laboratory experiment. Specifically, there were 38 repetitions per condition instead of 40, resulting in 96 trails consisting of 76 transitive and 20 intransitive pictures.

Trial Timing. The baseline phase and priming phase were presented as one continuous task. The change from the baseline phase to the priming phase was not obvious to the participants. The timing of each trial (baseline and priming phase) was as follows: Participants were initially presented with the infinitive form of a verb (to be used in an upcoming utterance, e.g., *to run*) for 500ms. After 500ms of a white screen, a coloured picture would appear. Participants were instructed to describe the picture following the rules described above. The picture was presented until the participant responded (with a time-out after 12 seconds). There was an intertrial interval of 1,500-2,000ms (jittered) before the next verb was presented. Coloured and greyscale pictures were alternated until all pictures were described. No verbs were repeated within a trial, and no picture was repeated within an experimental session. Each experimental session took a total of 45 minutes to complete.

Materials

Stimuli. The pictures used in the lab experiment are described in Heyselaar and Segaert (2022). In short, the picture set depicts transitive events such as to chase, to interview or to serve with an agent and patient (either both adults or children). Each picture has a grayscale version and two colour-coded versions with a green and a red actor. There was always one male and one female actor with each event depicted with each of the two actors serving as the agent, and the agent randomized on the left or right. For the app experiment, we created new stimuli using Procreate. To minimize gender stereotyping, we used anthropomorphized shapes instead of human actors. Specifically, we used a circle, a triangle, a square, and a star (see Figure 1). To ensure equal exposure and avoid priming effects towards specific shape characters, each shape was equally presented as the agent and as the patient, see below.

75 unique verbs were used: 40 depicted transitive verbs such as *to kiss* or *to follow*, and 35 depicted intransitive verbs such as *to sneeze* or *to panic*. The transitive pictures were used to elicit transitive prime and target sentences. Each transitive picture had three versions: one grayscale version and two colour-coded versions with a green and a red actor. Each transitive event was depicted by one of the eight possible combinations of the four characters used. Each shape is always depicted with one of the three other shapes; shapes never appear more than once in a given picture. Each shape-pair appeared twice per verb: once with the shape serving as the agent, once with the shape serving as the patient. To prevent participants forming strategies, the position of the agent (left or right) was randomized. Fillers elicited intransitive sentences, depicting events such as *running* or *singing* with one actor (in grayscale or green)

Lab. Participants were invited to complete the study in the lab. The experiment was completed on a Dell Latitude E5470 Laptop (14" screen) using E-Prime (Scheider et al., 2002).

App. Participants were invited to download the app via the Google Play Store. An interactive schematic of the app can be found here. The app was written using Android Studio Electric Eel (2022.1.1 Patch 1) in Java and accessible for all phones with an API 31 or lower. Instructions and practice trials were identical for the lab and app versions of this experiment.

Data Analysis

Coding. To code which structure the participants used to describe the picture, different strategies were used for the lab and app version. In the lab, the researcher would code the structures used online. If participants made mistakes during the practice sessions or the first 10 trials, the researcher would enter the cubicle and correct the participants. For the app, we used the transcription function (RecognitionListener) in Android Studio to transcribe the descriptions. In this way, we did not need to store any sensitive audio data. This function was manually triggered by the participant pressing and holding a button while they described the picture. This ensured that the participant was in control of when the app would "listen".

Using the package *spacyr* in R (Benoit et al., 2023), we wrote a custom script to identify the syntactic structures used in the transcribed utterance. Specifically: if an utterance included the word 'by', it was marked as a passive. If there was an agent but no patient, then it was marked as an intransitive. Utterances were coded as an active if there was an agent and the patient occurs after the verb. Any utterances that did not fall into these categories were manually checked and assigned a grammatical structure. This was the case for 4.9% of the data.

Pre-processing.

Mixed-Effects Models. To determine whether there is a difference in the proportion of passive responses produced in the baseline sessions, a general linear mixed effects model (GLMER) was run, using the lme3 package. We started with a full model with a maximal random-effect structure as described by Barr et al., (2013): The repeated measures nature of the data was modeled by including a per-participant and per-item random adjustment to the fixed intercept ("random intercept"). We attempted to include as many random adjustments to the fixed effects ("random slopes") as justified by the main effects of the model. We used a backward step reduction procedure to identify the model that best captured the data. Specifically, we used the *anova* function from the stats package to compare models and only continued if there was no significant difference.

To determine whether there is a difference in the production of passive structures over the course of an experimental session, we utilized a GAMM for all our analyses, using the gam function from the itsadug

package (van Rij et al., 2017), as previous experiments have shown that cumulative priming mirrors a growth curve more than a linear correlation over the length of the experimental session (referencing all our papers). Unlike ANOVAs or generalized mixed-effects regression (GLMER), GAMM does not assume linearity (although it can find a linear form if supported by the data). Instead, GAMM strikes a balance between model fit and the smoothness of the curve using either error-based or likelihood-based methods in order to avoid over- or underfitting. Thus, the data guide the functional form (Hastie and Tibshirani, 1990). The p-value provided therefore indicates whether or not the growth of the curve is significantly different from zero (i.e., no growth). Additionally, GAMM also allows the inclusion of random effects to capture the dependencies between repeated measures.

We again, initially included a by-group interaction. The final model included a Session by Group interaction, trial number (modeled with a smoother), and the interaction between trial number and session (modeled with a smoother). The smoothers used were the default underlying base functions, specifically thin plate regressions. We included per-participant and per-item random smooths, with trial number as a random slope for both. We used the function gam.check for model diagnostics.

For both model types, the dependent variable was the syntax used during the target picture descriptions, coded as 0 and 1. We used a sum-to-zero contrast to compare the two sessions.

Results

The data collected in the lab environment ("lab data") are reported elsewhere (Heyselaar & Segaert, 2022) and are used in the current manuscript as a point of comparison for the data collected through the app ("app data"). For each of the following subheadings, we will first provide an overview of our findings in the laboratory setting, to determine whether the data collected in the app are comparable. For the analysis, we are not interested in direct comparisons between the lab and the app collected data, as differences could be due to the methodological differences of non-interest (e.g. during the lab the participant could immediately describe the picture aloud, whereas for the app they had to first press and hold a button). Instead, we are interested in comparing the trends seen in the data between these two collection methods, and whether similar phenomena can be established using the mobile app.

Short-Term Priming Effects

To determine whether the app was able to induce any passive priming, we first explored the data from the first session (Session A). We compared passives produced after a passive prime with passives produced after an active prime in the priming phase. The final model included a main effect of Prime Type (Active vs Passive, sum-contrast coded), as well as a random slope of Prime Type included in the per-participant and per-item random intercept.

Participants showed a significant increase in the number of passives produced after a passive prime compared to after an active prime (B = 11.01, $p < .001$), suggesting that the app was able to induce the standard short-term priming effect as seen in laboratory studies over many years (Mahowald et al. 2016).

Long-Term Priming Effects

Long-term priming effects have been studied less frequently, but in our previous work we established long-term priming in the current task (Heyselaar and Segaert 2022). To investigate long-term priming, we examined the behaviour during the baseline phase and only included trials that had correctly described intransitive primes and transitive targets. Table 1 reports model outputs for the app- as well as the lab-collected data.

Table 1.

Summary of the Binomial Mixed Effects Model to Analyze the Passive Production During the Baseline Phase for the A. Data collected via the App and B. Data collected in the lab.

A. Lab Data (taken from Heyselaar & Segaert, 2022)

	Estimate	*SE*	*z* value	*p* value	
Intercept	-10.42	1.58	-6.61	< .001	***
Session	5.16	1.70	3.04	.002	**

Note: $N = 1655$; log-likelihood = -138; ** $p < .01$; *** $p < .001$

B. App Data

	Estimate	*SE*	*z* value	*p* value	
Intercept	-34.89	11.06	-3.16	.002	**
Session	-8.95	28.53	-0.31	.754	

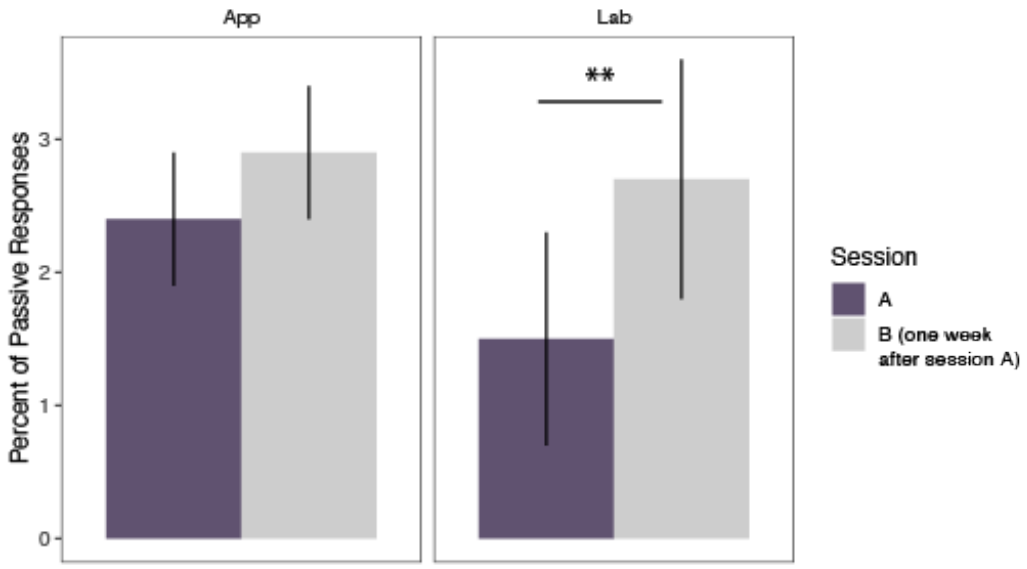
Note: $N = 2139$; log-likelihood = -51.2; ** $p < .01$

Lab Data. The table reports a significant increase in the proportion of passives produced in the baseline phase after the one-week interval (B = 5.16, p = .002).

App Data. Table 1A suggests that there is no significant difference (B = -8.95, p = .754) in the proportion of passives produced in the baseline phase between Session A and Session B (one week interval). This is in contrast to findings from the laboratory. Figure 2 illustrates these differences.

Figure 2.

The percentage of Passive Responses in the Baseline Phase.



Note. Error bars represent standard error. ** p < .01

Cumulative Priming Effects

In the priming phase, we examined the trial-by-trial development of the syntactic preference for active versus passive structures. To accurately capture the probability of producing a passive structure on the current trial, we calculated the *cumulative passive proportion*. This variable was calculated as the proportion of passives produced on the target trials that occurred *before* the current trial, as suggested by Jaeger and Snider (2013). Table 2 shows the model output.

Table 2.

Summary of the Generalized Additive Mixed Effects Model to Analyze the Long-Term Priming Effect for the Priming Phase for A. Data collected via the App and B. Data collected in the lab.

A. Lab Data (taken from Heyselaar & Segaert, 2022)

Parametric coefficients					
	Estimate	*SE*	*t* value	*p* value	
Intercept	0.03	0.01	3.75	< .001	***
Session	0.01	0.00	7.81	< .001	***
Smooth Terms					
	edf	Ref. *df*	*F* value	*p* value	
Trial	1.44	1.76	35.56	< .001	***
Trial by Session	1.00	1.00	0.22	.637	
Random smooth for participants	27.79	28.00	128.70	< .001	***
Random smooth for items	11.27	1118.00	0.01	.168	

Note: $N = 4196$; ** $p < .01$; *** $p < .001$

B. App Data

Parametric coefficients					
	Estimate	*SE*	*t* value	*p* value	
Intercept	0.08	0.09	0.81	.419	
Session	-0.01	0.00	-10.92	< .001	***
Smooth Terms					
	edf	Ref. *df*	*F* value	*p* value	
Trial	2.25	2.63	0.81	.617	
Trial by Session	2.24	2.78	11.42	< .001	***
Random smooth for participants	178.97	388	17.73	< .001	***
Random smooth for items	899.60	3312	3.29	< .001	***

Note: $N = 5817$; ** $p < .01$; *** $p < .001$

Lab Data. The parametric coefficient portion of Table 2 illustrates the difference in intercept points between the sessions. As illustrated in Figure 2, this means that Session B overall had a higher proportion of passive utterances than Session A. This suggests that there is long-term priming from Session A to Session B 1 week later.

The Smooth Terms portion of Table 2B shows the change in bias for active versus passive preference on a trial-by-trial basis - in other words, the slope of the curve. The lack of a significant interaction between trial and the difference between Session A and B means that the rate of change was not significantly

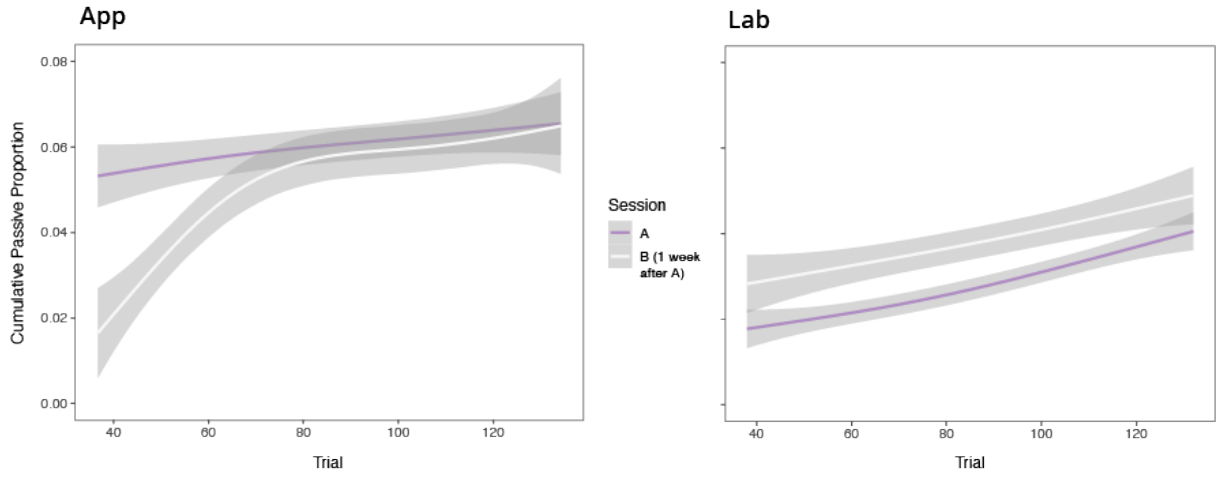
different between these two sessions. In other words, the participants learned at exactly the same rate, with the only difference being that for Session B their baseline preference for passives was significantly higher.

App Data. The parametric coefficient portion also suggests a significant difference in intercept points between sessions. This is illustrated in Figure 2. Indeed, the starting point for Session B is significantly *lower* than the starting point for Session A. This supports the conclusions from Table 1; there is no long-term priming observed for the data collected from the app.

The Smooth Terms portion of Table 2B indicates that the rate of growth is significantly different between Sessions. As can be seen from Figure 3, the slope of Session B is much steeper than that of Session A, and for this reason, the passive preference at the end of each session ends at the same point. This explains the lack of a significant difference reported in Table 1.

Figure 3.

The Cumulative Passive Proportion Per Trial for A. App-collected data and B. Lab-collected data.



Note: Error clouds represent standard error.

Interim Discussion

There are differences in the findings from the data collected in the app environment compared to those collected in the lab environment. Importantly, in the mobile app we do show robust short-term priming (an increase in passives produced after a passive prime), which is consistent with many years of laboratory structural priming research (Mahowald et al., 2016). The key difference seems to be in the retention of this newly learned preference: in the lab, participants still showed a preference for passives a week later, whereas this is not seen in the data collection from the app. A potential explanation could be that the app allows participants to undertake syntactic priming experiments in less controlled environments, such as on public transport or in a waiting room. Moreover, the environment may be variable between and even within sessions. Hence, because the app is mobile (as opposed to a fixed invariable lab-setting) context-based learning contributions would be limited. It could be that, for these reasons, the passive preference learned in Session A decays at a faster rate than the laboratory collected data.

To determine whether this is the case, we ran a new app experiment but now with only 24 hours between the two sessions. If the app is still not able to show long-term priming after such a short interval, then it might not be suited for syntactic priming experiments.

Methods

The exact same set-up was used as reported above. Only novel participants, not involved in the above experiment, were invited to complete the study. Specifically, 38 participants (M_{Age}: 25.29 years, SD_{Age}: 3.71; 17 males, 1 non-disclosed) were recruited via Prolific.

Results

Short-term priming

The new data again showed significant short-term priming effects. There was a significant increase in passives produced in target trials after a passive prime (B = -5.86, $p < .001$).

Long-term priming

Critically, the new data showed a significant increase in the proportion of passives produced in the baseline condition of Session B compared to Session A (24 hours before). This suggests that long-term priming is possible and measurable via the app. The model output is reported in Table 3 and illustrated in Figure 4A.

Table 3.

Summary of the Binomial Mixed Effects Model to Analyze the Passive Production During the Baseline Phase for the Data collected via the App

	Estimate	*SE*	*z* value	*p* value	
Intercept	-34.89	11.06	-3.16	.002	**
Session	-8.95	28.53	-0.31	.754	

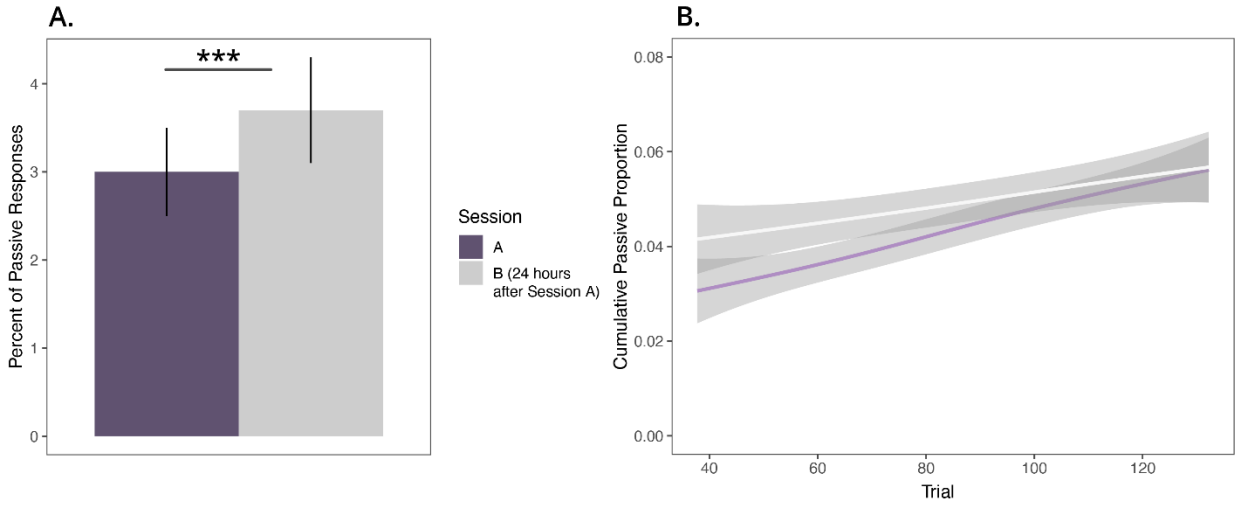
Note: $N = 2139$; log-likelihood = -51.2; ** $p < .01$

Cumulative priming

This suggestion is also supported by the cumulative priming analysis; the intercept for Session B is significantly higher than the intercept for Session A. The slope of the passive preference accumulation is steeper in Session B than Session A suggesting that participants are learning faster during this session. This is similar to the results seen in the one-week interval. Figure 4B illustrates this result and Table 4 reports the model output.

Figure 4

Passive priming production behaviour in the app-collected data with 24 hours between session A and B



Note: A. illustrates the change in passive preference as measured during the baseline phase at the beginning of each session. B. illustrates the cumulative learning of passive preferences over the course of the two sessions. Error clouds/bars represent standard error.

Discussion

The aim of the present work is to establish whether we can measure structural priming using a mobile app employed in the real world. Even though theoretical accounts have emphasized the importance of structural priming for language learning and language acquisition (Chang, 2008; Tomasello, 2006), the majority of priming studies are conducted in a controlled laboratory setting. In addition, many questions remain that can be better answered with a larger and more diverse participant pool, in an everyday environment. In the current study, we conclude that syntactic priming can be observed when participants complete a syntactic priming task on a smartphone app, in an uncontrolled, real-world environment.

To meet ethical and data privacy regulations, our app instantly transcribed the picture descriptions into text, ensuring no sensitive data needed to be stored. During the analysis, we used a custom-made code to automatically mark syntactic structures. These features effectively eliminate the need for any experimenter involvement in transcribing and coding participant responses, making our mobile app very scalable.

In the current study, we tested WEIRD participants aged 18 - 30 in order to compare the syntactic priming effects found in the app with syntactic priming data commonly reported in the literature. The app showed significant short-term priming (increased production of passives over the course of a single session), consistent with many years of syntactic priming research. We did not observe long-term syntactic priming effects lasting over a one-week period (i.e., the passive preference established in the first session was not observed a week later). Such a long-term priming effect has been established in the literature (Kaschak, 2007; Kaschak et al., 2011) as well as in the laboratory data we used as a baseline comparison (Heyselaar & Segaert, 2022). To establish why long-term priming was not established in the app, we conducted a second experiment with only 24 hours between the first and second sessions (as opposed to a week). We hypothesized that the real-world environment was noisier and more prone to distraction (Heyselaar and Segaert, 2019) or decay than the controlled laboratory environment (Chang, 2008; Dell and Chang, 2014). Indeed, when reducing the time between sessions, we do see evidence that suggests that the passive preference established in the first session was still present in the second session (24 hours later). We therefore conclude that the smartphone application is able to measure syntactic priming behaviour. Although the longevity of these effects might be reduced due to the uncontrolled nature of the participant's environment when using the app, this might be a better representation of syntactic priming in the real world.

Above, we highlighted some of the advantages of measuring priming via a mobile app (as opposed to a laboratory setting): one can reach a more inclusive and larger sample. More diverse human research participant sampling increases the external validity of the findings. Enabling larger samples opens possibilities for investigating how characteristics of the individual or context modulate the priming effect. Furthermore, we illustrated above a potential application of an accessible priming app for older adults. Another potential application supported by previous research is for the training of syntactic processes in clinical populations. Van Boxtel et al. (2025) has demonstrated that structural priming can be used as a training mechanism for aphasia patients in a laboratory setting. Our mobile app offers a long-term accessible way of applying such an intervention in practice. This would be in line with broad literature on using smartphone applications for brain-training. Brain training literature suggests that far-transfer effects of training are limited (Sutton et al., 2025), so if the aim of the intervention is to improve language processing, training should focus on the aspect of language processing which is specifically in need of improvement, e.g. syntactic priming for patients with syntactic processing deficits.

In sum, our smartphone application enables large-scale data collection in naturalistic settings and offers a methodology for potentially more inclusive sampling. The app provides a flexible methodology for investigating individual differences, reaching people who may be unable to attend laboratory sessions, and potential intervention applications in language research.

Declarations

Funding

Research was funded by the Leverhulme Trust (RPG-2021-360)

Conflicts of interest/Competing interests

All authors report no competing/conflicting interests

Ethics Approval

This study has been approved by the research ethics board of the University of Birmingham.

Consent to Participate

Informed consent was obtained from all individual participants included in the study.

Consent for Publication

Not applicable

Availability of data and materials

The data and materials for all experiments are available at https://doi.org/10.17605/OSF.IO/3G7UJ

Code availability

The code used for all experiments are available at https://doi.org/10.17605/OSF.IO/3G7UJ

References

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J. Mem. Lang. 68, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Benoit, K., Matsuo, A., Gruber, J., 2023. Wrapper to the "spaCy" "NLP" Library.

Bernolet, S., Collina, S., Hartsuiker, R.J., 2016. The persistence of syntactic priming revisited. J. Mem. Lang. https://doi.org/10.1016/j.jml.2016.01.002

Bialystok, E., 2016. The signal and the noise: Finding the pattern in human behavior. Linguist. Approaches Biling. 6, 517–534. https://doi.org/10.1075/lab.15040.bia

Bock, J.K., 1986. Meaning, sound, and syntax: Lexical priming in sentence production. J. Exp. Psychol. Learn. Mem. Cogn. 12, 575–586.

Bock, J.K., Griffin, Z.M., 2000. The Persistence of Structural Priming: Transient Activation or Implicit Learning? J. Exp. Psychol. Gen. 129, 177–192. https://doi.org/10.10371/D096-3445.129.2.1T/The

Bonnechère, B., Langley, C., Sahakian, B.J., 2020. The use of commercial computerised cognitive games in older adults: a meta-analysis. Sci. Rep. 10, 15276. https://doi.org/10.1038/s41598-020-72281-3

Branigan, H.P., Messenger, K., 2016a. Consistent and cumulative effects of syntactic experience in children's sentence production: Evidence for error-based implicit learning. Cognition 157, 250–256. https://doi.org/10.1016/j.cognition.2016.09.004

Branigan, H.P., Messenger, K., 2016b. Consistent and cumulative effects of syntactic experience in children's sentence production: Evidence for error-based implicit learning. Cognition 157, 250–256. https://doi.org/10.1016/j.cognition.2016.09.004

Branigan, H.P., Pickering, M.J., Cleland, a a, 1999. Syntactic priming in written production: evidence for rapid decay. Psychon. Bull. Rev. 6, 635–640. https://doi.org/10.3758/BF03212972

Branigan, Holly P., Pickering, M.J., Cleland, A.A., 2000. Syntactic co-ordination in dialogue. Cognition 75, B13–B25. https://doi.org/10.1016/S0010-0277(99)00081-5

Branigan, Holly P, Pickering, M.J., Stewart, A., McLean, J.F., 2000. Syntactic priming in spoken production: linguistic and temporal interference. Mem. Cognit. 28, 1297–302.

Chang, F., 2008. Implicit learning as a mechanism of language change. Theor. Linguist. 34, 115–122. https://doi.org/10.1515/THLI.2008.009

Chang, F., Dell, G.S., Bock, J.K., 2006. Becoming syntactic. Psychol. Rev. 113, 234–72. https://doi.org/10.1037/0033-295X.113.2.234

Dell, G.S., Chang, F., 2014. The p-chain: Relating sentence production and its disorders to comprehension and acquisition. Philos. Trans. R. Soc. B Biol. Sci. 369. https://doi.org/10.1098/rstb.2012.0394

DeLuca, V., Segaert, K., Mazaheri, A., Krott, A., 2020. Understanding bilingual brain function and structure changes? U bet! A unified bilingual experience trajectory model. J. Neurolinguistics 56, 100930. https://doi.org/10.1016/j.jneuroling.2020.100930

Hardy, S.M., Wheeldon, L., Segaert, K., 2020. Structural priming is determined by global syntax rather than internal phrasal structure: Evidence from young and older adults. J. Exp. Psychol. Learn. Mem. Cogn. 46, 720–740. https://doi.org/10.1037/xlm0000754

Hartsuiker, R.J., Bernolet, S., Schoonbaert, S., Speybroeck, S., Vanderelst, D., 2008. Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. J. Mem. Lang. 58, 214–238. https://doi.org/10.1016/j.jml.2007.07.003

Hartsuiker, R.J., Kolk, H.H.J., 1998. Syntactic Persistence in Dutch. Lang. Speech 41, 143–184. https://doi.org/10.1177/002383099804100202

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall/CRC Press, London.

Henrich, J., Heine, S.J., Norenzayan, A., 2010. Most people are not WEIRD. Nature 466, 29–29. https://doi.org/10.1038/466029a

Heyselaar, E., Segaert, K., 2022. Structural priming persists for (at least) one month in young adults, but not in healthy older adults. J. Exp. Psychol. Learn. Mem. Cogn. 48, 1219–1234. https://doi.org/10.1037/xlm0001123

Heyselaar, E., Segaert, K., 2019. Memory encoding of syntactic information involves domain-general attentional resources: Evidence from dual-task studies. Q. J. Exp. Psychol. 72, 1285–1296. https://doi.org/10.1177/1747021818801249

Heyselaar, E., Wheeldon, L., Segaert, K., 2020. Structural priming is supported by different components of nondeclarative memory: Evidence from priming across the lifespan. J. Exp. Psychol. Learn. Mem. Cogn. https://doi.org/10.1037/xlm0000955

Jaeger, T.F., Snider, N.E., 2013. Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime' s prediction error given both prior and recent experience. Cognition 127, 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Kaschak, M.P., 2007. Long-term structural priming affects subsequent patterns of language production. Mem. Cognit. 35, 925–937.

Kaschak, M.P., Kutta, T.J., Schatschneider, C., 2011. Long-term cumulative structural priming persists for (at least) one week. Mem. Cognit. 39, 381–388. https://doi.org/10.3758/s13421-010-0042-3

Kemper, S., 1994. Elderspeak: Speech accommodations to older adults. Aging Neuropsychol. Cogn. 1, 17–28. https://doi.org/10.1080/09289919408251447

Kemper, S., Greiner, L.H., Marquis, J.G., Prenovost, K., Mitzner, T.L., 2001a. Language decline across the life span: Findings from the nun study. Psychol. Aging 16, 227–239. https://doi.org/10.1037/0882-7974.16.2.227

Kemper, S., Thompson, M., Marquis, J., 2001b. Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and propositional content. Psychol. Aging 16, 600–614. https://doi.org/10.1037/0882-7974.16.4.600

Mahowald, K., James, A., Futrell, R., Gibson, E., 2016. A meta-analysis of syntactic priming in language production. J. Mem. Lang. 91, 5–27. https://doi.org/10.1016/j.jml.2016.03.009

Rabipour, S., Davidson, P.S.R., 2015. Do you believe in brain training? A questionnaire about expectations of computerised cognitive training. Behav. Brain Res. 295, 64–70. https://doi.org/10.1016/j.bbr.2015.01.002

Reitter, D., Keller, F., Moore, J.D., 2011. A Computational Cognitive Model of Syntactic Priming. Cogn. Sci. 35, 587–637.

Schoot, L., Heyselaar, E., Hagoort, P., Segaert, K., 2016. Does syntactic alignment effectively influence how speakers are perceived by their conversation partner? Plos One 11, e0153521. https://doi.org/10.1371/journal.pone.0153521

Segaert, K., 2018. Priming Effects, in: Zeigler-Hill, V., Shackelford, T.K. (Eds.), Encyclopedia of Personality and Individual Differences. Springer International Publishing, Cham, pp. 1–4. https://doi.org/10.1007/978-3-319-28099-8_479-1

Segaert, K., Menenti, L., Weber, K., Hagoort, P., 2011. A paradox of syntactic priming: why response tendencies show priming for passives, and response latencies show priming for actives. PloS One 6. https://doi.org/10.1371/journal.pone.0024209

Sung, J.E., 2015. Age-related changes in sentence production abilities and their relation to working-memory capacity: evidence from a verb-final language. PloS One 10, e0119424. https://doi.org/10.1371/journal.pone.0119424

Sutton, E., Catling, J., Zanten, J.J.C.S.V.V., Segaert, K., 2025. Practice makes perfect, but to what end? Computerised brain training has limited cognitive benefits in healthy ageing. Psychol. Res. 89, 75. https://doi.org/10.1007/s00426-025-02110-7

Tetlow, A.M., Edwards, J.D., 2017. Systematic Literature Review and Meta-Analysis of Commercially Available Computerized Cognitive Training Among Older Adults. J. Cogn. Enhanc. 1, 559–575. https://doi.org/10.1007/s41465-017-0051-2

Tomasello, M., 2006. Acquiring linguistic constructions, in: Kuhn, D., Siegler, R. (Eds.), Handbook of Child Psychology. Wiley, New York, NY.

Van Boxtel, W.S., Rainey, K.M., Ferreira, V., Martin, N., Bauman, E., Lee, J., 2025. Structural Priming Treatment in Aphasia: The Role of Lexical and Abstract Syntactic Representations. Am. J. Speech Lang. Pathol. 34, 3318–3342. https://doi.org/10.1044/2025_AJSLP-25-00095
van Rij, J., Wieling, M., Baayen, R., van Rijn, H., 2017. itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs.