The development of abstract syntactic representations in beginning L2 learners of Dutch

The development of abstract syntactic representations in beginning L2 learners of Dutch

Rianne van Lieburg[a]*, Edwige Sijyeniyo[a*], Robert J. Hartsuiker[b], & Sarah Bernolet[a]

[a]University of Antwerp, Grote Kauwenberg 18, 2000, Antwerp (Belgium)

[b]Ghent University, Henri Dunantlaan 2, 9000, Ghent (Belgium)

*these authors contributed equally to this work

Corresponding author:

Rianne van Lieburg

Rianne.vanLieburg@UAntwerpen.be

https://orcid.org/0000-0001-5024-8096

Abstract

The developmental account of second language (L2) syntactic acquisition in late learners (Hartsuiker & Bernolet, 2017) predicts that learners start with item-specific syntactic representations, which become abstract over time. We investigated how the transition between item-specific and abstract syntactic representations takes place for transitive structures in a within-Dutch structural priming experiment. In a longitudinal and a cross-sectional design, we tested whether and when late learners show priming for active and passive sentences, and whether the learning of the passive structure can be sped up by means of a lexically-based structural priming intervention. Active priming took place before passive priming, although abstract representations of the passive may be formed quite rapidly after exposure, which seemed to be accelerated by the intervention. Our results suggest that a developmental account of L2 syntactic acquisition should be a hybrid model, incorporating aspects of the residual activation account as well as an implicit learning mechanism.

Keywords

L2 acquisition, late learners, syntax, structural priming, abstract structural representations

Competing interests: none

JEL code: Z13 (Language)

Acknowledgments: This work was funded by the Fonds Wetenschappelijk Onderzoek Vlaanderen, grant K112815N and by the Bijzonder Onderzoeksfonds (BOF), University of Antwerp, grant DOCPRO4.

1. Introduction

Moving or migrating to a new country often entails being confronted with learning the language of one's new home. When a language is learned, syntactic representations are formed that are used for the comprehension and the production of syntactic structures. But how does this late learning of syntactic structures occur and when in the learning trajectory do late learners establish structural representations of syntactic structures ready for production?

In this paper, we ask how late adult learners of Dutch, who are at the very beginning stages of language learning, establish abstract syntactic representations for active (*The girl reads the book*) and passive sentences (*The book is being read by the girl*). Testing the predictions of a developmental account of second language (L2) syntactic acquisition (Hartsuiker & Bernolet, 2017), we hypothesize that late learners have the tendency to produce active sentences spontaneously and not passive sentences, as passives are the less frequent transitive alternant. This would suggest that learners may not have a syntactic representation for passives yet. Here, we investigated *when* in the learning trajectory transitive syntactic representations are formed, and we tested our research question with *structural priming*.

Structural priming is the tendency to reuse previously processed syntactic structures (Bock, 1986), and it allows researchers to investigate how syntactic information is represented and accessed during syntactic processing (Pickering & Ferreira, 2008; Mahowald et al., 2016). For instance, speakers have a stronger tendency to use a passive structure (*The boy is being bitten by the dog*) after hearing a passive sentence (*The cake is being baked by the cook*) than when they have just heard its active counterpart (*The cook bakes the cake*). As such, structural priming is a way to elicit less frequent alternatives that may not be produced

spontaneously, and it is believed that structural priming only occurs if the speaker has a mental representation of a given syntactic structure (Pickering & Ferreira, 2008). According to Pickering and Branigan (1998), the speaker's lexicon contains a distinct combinatorial node of each separate syntactic structure linked to lemma nodes that can be used with that structure. In our case, if beginning learners who do not yet produce passive sentences spontaneously show a tendency to produce a passive sentence instead of an active sentence, after hearing a passive prime, this suggests that they have developed a structural representation for the passive.

Syntactic representations in bilingual speakers

Work on structural priming in bilingual speakers suggests that these speakers share syntactic representations whenever syntactic structures are *similar* enough between the two languages of a bilingual. For instance, in a cross-language structural priming experiment with Spanish-English bilinguals, Hartsuiker et al. (2004) found that participants had a stronger tendency to produce English passive sentences (*The man is bitten by the dog*) after they heard a Spanish passive prime sentence (*La cantante es atendida por el obrero* "The singer is served by the construction worker"), than when they heard a Spanish active prime sentence. Hartsuiker et al. suggested that between-language priming effects are due to a shared syntactic representation of the passive structure between Spanish and English. Following this assumption, they proposed their bilingual lexical-syntactic model, which is rooted in the residual activation theory (Pickering & Branigan, 1998). Pickering and Branigan suggest that priming of the passive occurs due to short-term residual activation of the lexical representation of a transitive verb (e.g., 'to bite'), the syntactic representation for the passive structure in which "to bite" occurs, and the link between the lexical and syntactic representations, which is strengthened upon the activation of both representations. Hartsuiker and colleagues extended the residual activation theory to bilinguals, proposing that activation

of lexical and syntactic representations in one language induces activation of translation equivalents and shared syntactic structures in the other language.

Though the residual activation theory provides an explanation for short-term structural priming effects and for the *lexical boost effect* (there is a larger tendency to repeat a recently processed structure if the same lexical item is used), it does not explain that structural priming effects are long-lived rather than short-lived (Hartsuiker et al., 2008). Structural priming seems to reflect a long-term learning process. Chang et al. (2006) suggested that structural priming effects arise due to error-based learning of syntactic rules, in which learning is induced by a mismatch between the predicted input and the actual input during syntactic processing. This process is influenced by the relative frequency of syntactic alternants, as less frequent structures induce a larger prediction error than more frequent structures. As a result, low frequent structures (e.g., passives), give rise to larger priming effects than high frequent structures (e.g., actives), a phenomenon called the *inverse preference effect* to structural priming (Ferreira & Bock, 2006). Importantly, recently, Khoe et al. (2023) implemented and tested a bilingual version of the implicit learning account, showing that error-based learning plays an important role in L2 learning too.

In bilingual speakers, not only the relative frequency of a structure but also L2 proficiency seems to play an important role in the magnitude of structural priming. Bernolet, Hartsuiker, and Pickering (2013) showed that higher proficiency speakers of English, who had Dutch as their L1, were primed more strongly for genitives (S-genitives: *the girl's shirt* and Of-genitives: *the shirt of the girl*) than lower proficiency speakers. In contrast, within the L2, the less proficient L2 learners showed the strongest priming for items with lexical overlap. These results suggest that L2 learners start with non-shared, item-specific syntactic representations in their L2, which become abstract and shared between the L1 and the L2 as time progresses. As such, the magnitude of abstract priming effects increases together with

L2 proficiency. Less proficient L2 learners rely on item-specific syntactic representations, which explains why lexical overlap between the prime and target sentences within the L2 leads to larger priming effects in less proficient than in more proficient L2 learners.

Developmental account for the acquisition of L2 syntax

In their developmental account for the acquisition of L2 syntax, Hartsuiker and Bernolet (2017) propose a possible account of the process during which similar L1-L2 syntactic structures become shared over time. According to their account, syntactic development takes place in five stages.

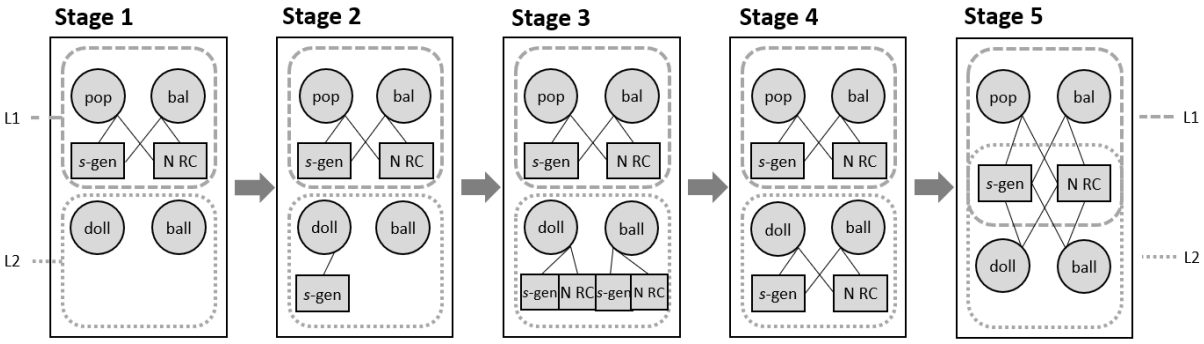


Figure 1. Developmental account for L2 syntax acquisition as proposed by Hartsuiker and Bernolet (2017). The upper part represents the lexical-syntactic network in the L1, and the lower part represents the development of the lexical-syntactic network in the L2.

Stage 1. In the first stage, the learner only has lexical representations without syntactic information connected to them. The learner uses their knowledge of the L1 to formulate sentences in the L2, which may lead to transfer errors. For example, an L2 learner of English with L1 Dutch may produce "the doll from the boy" as a translation of *de pop van de jongen* [correct: "the doll of the boy"], as the preposition *van* is the equivalent of both *of* and *from* in English.

Stage 2. The L2 learner will form item-specific syntactic representations of L2 structures. They may learn the phrase "the doll of the boy", but they might not yet be able to generalize this to other lexical items such as "the ball of the girl".

Stage 3. During this stage, more item-specific syntactic representations are added to the lexicon. This means that the L2 learner can use the lexical item in more than one construction. For example, they may be able to alternate between "the doll of the boy" and "the boy's doll". However, exposure to these structures is still too low to generalize beyond the item-specific syntactic representations. This means that the learner does not know yet whether the construction is a lexical expression or a more general syntactic pattern.

Stage 4. Based on the recurring patterns [object] of [person] and [person]'s [object], the learner will generalize the construction across lexical items, and is able to use the syntactic construction productively.

Stage 5. When this stage has been reached, syntactic structures that are sufficiently similar between languages become shared across languages. A shared syntactic structure means that there is one syntactic representation, for instance, [object] [preposition of/van] [person], which is connected to all Dutch and English nouns stored in a bilingual's lexicon.

The developmental account parallels usage-based models for L1 syntactic development in children (e.g., Tomasello, 2000), according to which children acquire syntactic structures based on item-specific representations called "verbs islands", although adults presumably rely more on explicit memory than children do during the early stages of syntactic learning (cf. Hartsuiker & Bernolet, 2017). In addition, the developmental account for L2 syntax relates to the Relation Morphology framework by Jackendoff and Audring (2020), who claim that there is no sharp distinction between grammatical and lexical representations. Abstract representations differ from lexical representations in terms of their productivity. The developmental account aims to describe how knowledge about the productivity of syntactic structures evolves, and how L2 representations are linked to the already existing representations in the L1.

The developmental account can also be related to previous work by Schwartz and Sprouse (1996). In their Full Transfer Full Access Model/FTFA model, Schwartz and Sprouse assume that second language learners transfer L1 syntactic information for the production of L2 syntactic structures in the initial stages of L2 syntactic acquisition (as is the case in Stages 1 and 2 of the developmental account). However, upon increasing proficiency, late second language learners may depend less on their L1 syntactic knowledge and adapt to a more native-like production pattern (Montero-Melis & Jaeger, 2019), and consequently, to a native-like structural priming pattern (Hartsuiker & Bernolet, 2017). In the developmental account, cross-linguistic influences from, for instance, production preferences are still predicted in higher proficiency learners due to shared syntactic representations.

Moreover, Hartsuiker and Bernolet's account has analogies with the revised hierarchical model (RHM) of the development of the L2 lexicon by Kroll and colleagues (e.g., Kroll & Stewart, 1994; Talamas, Kroll & Dufour, 1999). It is important to note though that RHM is only concerned with lexical and conceptual representations and not with syntax. In RHM, L2 words are initially asymmetrically connected with L1 words (i.e., stronger from L2 to L1 than vice versa) and there are no, or only very weak links, between L2 vocabulary and conceptual representations. Thus, in initial stages processing an L2 word is strongly mediated by L1. However, as the participant becomes more proficient the connectivity pattern changes, with more symmetrical connections between the lexical and the emergence and strengthening of links between L2 words and conceptual representations. Thus, in both Hartsuiker and Bernolet's account and the RHM, an increase in proficiency causes changes in network structure that lead to weaker L1 mediation during L2 processing.

The developmental account predicts different structural priming effects in the different phases of L2 syntactic acquisition. At the first stage, structural priming will only occur immediately after an item if there is lexical overlap between the prime and target sentence. In this case, learners may copy and edit the prime sentence onto their own target response, using the explicit memory of the prime structure as a cue for retrieval from working memory (see Bernolet et al., 2016; Bernolet & Hartsuiker, 2018). At the second and third stage, lexical overlap will still be necessary to induce structural priming since the representations are item-specific. However, the prime and the target structure do not need to follow each other immediately as explicit memory is no longer the only locus of structural priming and residual activation has started to play a role (Bernolet & Hartsuiker, 2018).

During the fourth stage, one may expect to find abstract structural priming within the L2, but no between-language priming yet. Finally, at the fifth stage, abstract structural priming will occur between languages, provided that the syntactic structures are similar

enough to become shared.

Importantly, abstract representations may be formed earlier for more frequent structures (e.g., actives) than for less frequent ones (e.g., passives). Therefore, it might be the case that abstract structural priming effects may be found for frequent structures, whereas lexical overlap is still necessary for the priming of less frequent structures. One could also imagine that verb overlap between several primes and targets may function as a tool to promote the formation of abstract structural priming in primes and targets that do not use verb overlap (for instance, due to implicit learning). As such, the L2 account of syntax suggests that lexically based priming aids the formation of abstract syntactic representations.

The formulation of the developmental account (Hartsuiker & Bernolet, 2017) was based on structural priming studies that recruited university students as late L2 learners, who had usually learned the L2 in a classroom context from an early age and were already quite proficient in the L2. For example, the proficiency effects reported by Bernolet et al. (2013) were found in psychology students with L1 Dutch who learned English during high school. Because these late learners were not at the very beginning stages of L2 acquisition, there is a need to test the validity of the developmental account with late learners who start out with little knowledge of their L2.

One way to investigate the early stages of L2 syntactic learning is by teaching participants a new language from scratch. Muylle et al. (2021a) tested the predictions from the developmental account by teaching participants (with Dutch as their L1) an artificial language. They were subsequently tested in a longitudinal structural priming study with five different sessions on their knowledge of transitive and ditransitive sentences. For the transitive structures, Muylle et al. found significant abstract structural priming effects already during the first session. The magnitude of the structural priming effects did not increase over time. By contrast, in some sessions, the priming effects were weaker than in the first session.

For the ditransitive structure, abstract structural priming within the artificial language and from the artificial language to Dutch was significant during the first session, and abstract structural priming from Dutch to the artificial language was only found after the second session.

As abstract priming effects were found much earlier than expected, the findings of Muylle et al. (2021a) did not provide conclusive evidence for the L2 syntax acquisition account, which predicts that within-language priming occurs before between-language priming, and that lexically based priming is not only found earlier than abstract structural priming, but that it also aids the formation of abstract syntactic representations. Their study shows that abstract structural representations may be developed very rapidly in an artificial language, but this may be different when learning natural languages.

Though using an artificial language has several advantages (e.g., full control of exposure to the language, see Wonnacott et al., 2008), one of its downsides is that the language is only used within one context (i.e., a lab) and this may influence the learning process and the speed of establishing syntactic representations. Therefore, we set out to test the predictions of the developmental account of L2 syntax acquisition in an ecologically valid learning situation, where exposure to the new language also occurs outside of class, with late learners of Dutch.

Current Study

Based on the predictions of the developmental account of L2 syntactic acquisition (Hartsuiker & Bernolet, 2017), we examined how the transition between item-specific (stage 3) and abstract syntactic representations (stage 4) takes place for transitive structures in beginning learners of Dutch. We investigated the following questions in a within-Dutch structural

priming experiment that had a lexically-based intervention block halfway through the experiment:

> **RQ1**: When in the learning trajectory of transitive structures do late learners of Dutch show priming for active and passive sentences?
>
> **RQ2**: To what extent do several instances of verb overlap in passives (the more complex transitive alternant) boost the production of passive sentences in subsequent trials without verb overlap?

In our experiment, we tested abstract structural priming as well as a possible transfer effect of passive items with verb overlap on subsequent prime-target trials without verb overlap. We chose to not directly compare priming effects of items with and without lexical overlap (either in a within-participants design or in a between-participants design), since structural priming effects with verb overlap between prime and target pairs are predicted to occur already from stage 1 of the developmental account. Repetition of verbs between prime and target can serve as a cue to the explicit memory of the prime sentence, and participants may use a copy-edit strategy to describe the target picture (Bernolet et al., 2016). Therefore, one may find priming effects even if there is no abstract representation of the more complex structure yet. Hence, such a design would not be very informative with regard to the transition from specific representations (stage 3) to abstract representations (stage 4). Consequently, our experiment consisted of three blocks: a pre-intervention block (no verb overlap), an intervention block in the middle of the experiment (with verb overlap between prime and target sentences), and a post-intervention block (no verb overlap). In the pre- and post-intervention blocks, we included active and passive prime sentences as well as a neutral baseline condition in order to investigate the presence of abstract structural priming of both the less complex and the more complex transitive alternative (see Bernolet et al., 2009, who

used a baseline too). The intervention block consisted of only passive prime sentences. In this way, we aimed to boost the production of the more complex and less frequent passive structure during and after the intervention block. Based on our research questions and experimental design, we formulate the following hypotheses:

> **H1**: Abstract syntactic representations for transitives may occur earlier for active sentences than for passive sentences because actives are the more frequent transitive alternant. Therefore, we expect to find active priming before passive priming. We expect to find passive priming as the (spontaneous) production of passives increases as a function of proficiency.
>
> **H2**: Learners may benefit from the few instances of lexical overlap in the intervention as it may promote the abstraction of less frequent structures in subsequent trials without verb overlap (due to implicit learning processes). This may result in more passive structures post-intervention, compared to pre-intervention, and thus possibly a stronger passive priming effect.

We tested our hypotheses in two different experimental designs, namely, a longitudinal and cross-sectional design. (1) Similar to Muylle et al. (2021a), our longitudinal design consisted of five sessions, in which we investigated the process of establishing syntactic representations *within* learners. (2) For our cross-sectional design, we used a group of lower proficiency and higher proficiency learners of Dutch to investigate whether the different stages of the developmental account would translate to different abstract structural priming patterns based on different L2 proficiency levels *between* learners.

In addition to our two groups of late Dutch learners, it was necessary to also test a Dutch control group to determine whether a few instances of verb overlap affect abstract structural priming of transitive structures in native language users. This was important to

investigate since we assume that native speakers have already developed and established syntactic representations for active and passive structures. We anticipated that the lexical boost effect induced in the intervention block, which would presumably lead to stronger passive priming, may extend to the post-intervention block.

For the longitudinal study, we expect active priming from the first few sessions and passive priming in the later sessions. Note that priming effects of the active structure can only be measured if participants (attempt to) produce a passive in at least a small part of the trials. It may be the case that we will only be able to detect active priming from Session 2 or 3, even though we assume that participants already have an abstract representation for the active structure as from Session 1 (if they are able to complete the task of describing pictures).

Since the learners in our longitudinal study were explicitly instructed on the passive structure in their language course, shortly before Session 3, we expect passive priming to occur from Session 3 or 4, depending on how fast abstract representations are formed after learning the structure. Similarly, for the cross-sectional study, we hypothesize that the lower proficiency learners may show active priming, and that the higher proficiency learners may show passive priming.

More generally, we expect that active priming will disappear in the later sessions of the longitudinal study, and that we will not find active priming in the higher proficiency learners of the cross-sectional study, due to the inverse preference effect. Therefore, we expect to find a similar priming pattern to native speakers in the higher proficiency learners (see Montero-Melis & Jaeger, 2019) (usually, native speakers do not show active priming but show strong passive priming).

In terms of the effects of the intervention, we predict that participants will use more passives due to a learning effect from the intervention items with verb overlap. First, the

intervention draws attention to the passive structure because of the repeated use of this structure (especially since there were no intervening fillers). Second, when participants produce passives during the intervention using a copy-edit strategy, they processed more passives than they would have done in a priming block without lexical overlap, which may lead to faster learning (see Muylle et al. [2021b], who showed that items with lexical overlap also boost structural priming in subsequent items without lexical overlap).

2. Experiment 1: Control group

2.1 Method

2.1.1 Participants

We recruited a group of native Dutch speakers (N = 19) as a control group. There were 13 females, 5 males and 1 other person with an age range from 18 to 29 years old (*M* = 22.6, *SD* = 4.1). For the sake of group comparison, we equally recruited approximately the same number of Dutch native speakers as L2 learners (see below).

2.1.2 Materials

Our materials, designed for the purpose of this study, comprised color drawings (1,064 pictures in total). All materials are available online in the Supplementary materials [https://osf.io/x8ejm/?view_only=60a652af276843e8835b8ffd797e29c0]**.** To make our materials suitable for the late learners of Dutch in Experiment 2 and 3, we used a limited vocabulary. Based on the learning materials of our late learners, we chose eight professions (animate entities) (e.g., *a baker*, a *singer)* and eight vehicles (inanimate entities) (e.g., *an ambulance*, *a firetruck*). For the pictures that contained animate entities, we chose six different transitive verbs: *to greet*, *to call*, *to help*, *to serve*, *to phone*, and *to carry*. Also, for the pictures with inanimate entities, we chose six different transitive verbs: *to hit*, *to pass*, *to*

block, *to replace*, *to drag*, *to follow*. Next to the transitive verbs for the critical trials, we selected intransitive verbs for the filler trials. We had twelve intransitive verbs for the pictures with animate entities (e.g., *to smile*, *to cry*) and twelve intransitive verbs for the pictures with inanimate entities (e.g., *to smell*, *to stop*). Importantly, our pictures were set up such that, despite the restricted vocabulary, many different noun-verb combinations were possible, while avoiding lexical overlap between items. As the materials were adjusted to the vocabulary knowledge of late learners, the items were easy for the native Dutch speakers.

2.1.3 Design

The experiment included audio prime sentences, a verification task, and target pictures for description. There were 72 prime sentences, 24 were critical prime sentences, divided over three conditions: the active (*De bakker helpt de zangeres* [The baker helps the singer]), the passive (*De zangeres wordt geholpen door de bakker* [The singer is being helped by the baker]) and the baseline condition, which consisted of conjoined noun phrases (*de bakker en de zangeres [the baker and the singer]*). We included a baseline to determine the production preferences of the learners in an unprimed condition (see Bernolet et al., 2009). Apart from the critical prime sentences, we included an intervention block in the middle of the experiment, consisting of four passive prime sentences. In the intervention block, we repeated the verb between the prime and target. The remaining 44 sentences were fillers, which could be described with intransitive verbs. The design of the experiment is illustrated in Figure 2. All prime sentences were recorded into separate audio files.

The verification task, which functioned as a distraction task, was presented simultaneously with the prime trials. Participants listened to the prime sentence while seeing two pictures: the correct picture that matched the prime sentence and a competitor picture. For instance, if participants would hear the prime sentence "*The teacher serves the doctor"*, they would see two pictures. In the correct picture, the teacher is the agent, and the doctor is

the patient. In the competitor picture, the doctor is the agent, and the teacher is the patient. The competitor picture could either be described with an active sentence "*The doctor serves the teacher*" or with a passive sentence "*The teacher is being served by the doctor*". We included the verification picture in the prime trials to test whether learners would interpret the first noun phrase that they heard as an agent (active interpretation) or a patient (passive interpretation). The verification for the filler trials consisted of the same object in the correct and competitor picture, but the verb was different in the two pictures. For instance, for the sentence "The teacher *sings*", the competitor picture would portray a teacher smiling. We manipulated the position of the correct picture and its competitor (i.e., on the left or on the right side of the screen) within each prime item across our experimental lists.

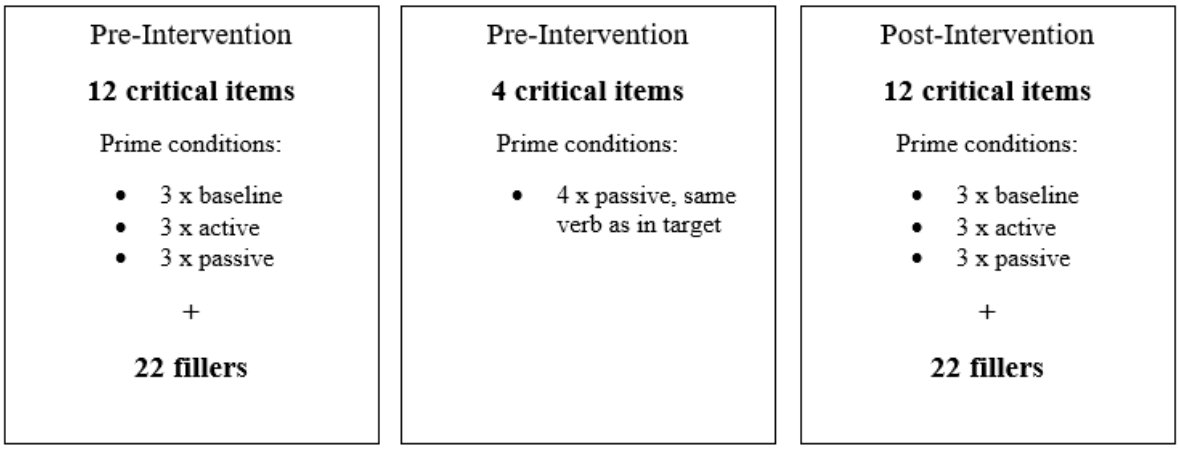
As critical target pictures, we used 24 pictures with transitive verbs that could be used to describe them. For these pictures, we did not repeat the verb of the prime sentence. In this way, we tested abstract structural priming (Pickering & Ferreira, 2008). We counterbalanced the position of the patient within items across our experimental lists, so that passive sentence production could not be related to the position of the patient in the target pictures. We also used 44 filler target items with intransitive verbs. The target verbs were displayed together with the pictures, and participants were told to conjugate the verb to produce a sentence.

Lastly, the prime-target pairs were pseudorandomly mixed with the filler items, with the constraint that the experiment started with three filler items. Each prime-target pair was followed by one, two or three filler items. Based on the experimental conditions, the position of the correct and competitor pictures in the verification task, and the position of the patient in the target items, we created 12 experimental lists to obtain a fully crossed design within items. For the last six lists, we reversed the order of the trials in the experimental lists. That is, all items that preceded the intervention block in the first six lists followed the intervention block in the last six lists. Similarly, all items that followed the intervention block in the first

six lists, preceded the intervention block in the last six list. Importantly, we used the same four trials in the intervention block across all experimental lists.

The experiment was programmed in *PsychoPy* v.3.4 (Peirce et al., 2019) and was run on the online platform Pavlovia. The Pavlovia experiment was embedded in a Qualtrics (Qualtrics, Provo, UT) survey, which contained the instructions, the request to provide informed consent and some demographic questions (age, gender, language background).

Figure 2. Design of the experiment.



2.1.4 Procedure

The native speakers received a link to the Qualtrics survey and were asked to provide the researchers with a recording of their spoken utterances after their participation. After the participants had read the instructions and had given their consent, they were instructed to turn on an audio recording and to start the experiment in Pavlovia. At each trial, participants saw a white screen with an audio button. After pressing the audio button, the prime sentence would start to play, and participants saw the correct picture and its competitor picture. Participants had to click on the picture that, according to them, displayed what they had just heard. Once participants had provided an answer, they saw a white screen with a "speak" button. After pressing the speak button, participants saw a picture for the target item accompanied with the

main verb at the bottom of the target picture. They had to conjugate the main verb to formulate a sentence which described the target picture. The native speakers took about 20 to 25 minutes to complete the experiment.

2.1.5 Coding

The responses were manually coded as active, passive or "Other" responses. A response was coded as active if the agent of the transitive event was mentioned first, followed by a conjugated verb and the patient. Passive sentences were coded when the patient was mentioned first, followed by an auxiliary verb (*to be*), a form of the past participle and if the sentence ended with a prepositional by-phrase using the preposition *door* (*by*). For instance, a correct passive response for Figure 3 would be: *De ambulance wordt gevolgd door de tram* - [The ambulance is being followed by the tram]. All other responses, including 'short passives', in which the agent was not overtly realized (e.g., *de bakker wordt gegroet* [the baker is being greeted]), were coded as "Other" responses.

Figure 3. An example of a target picture.

2.1.6 Analysis

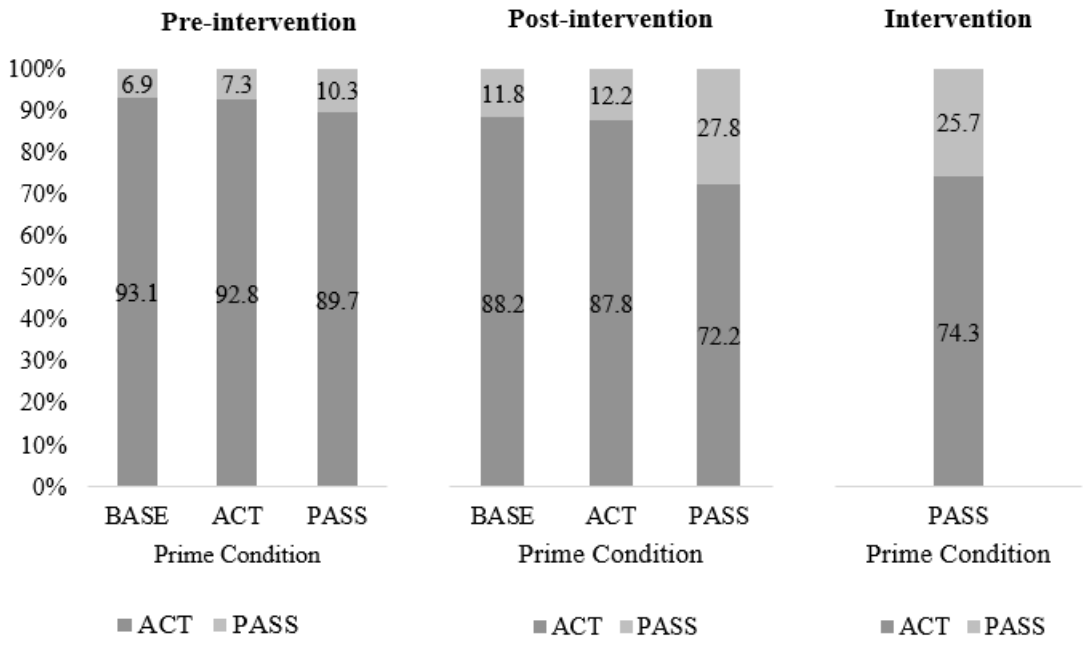
We measured the priming effects by comparing the proportions of active and passive responses after an active or passive prime sentence to the proportion of active and passive

responses after a baseline prime. The target responses were fit to a generalized linear mixed model (R-package lme4, Bates et al., 2015) with a Bobyqa optimizer to increase convergence (Powell, 2009) in *R* (version 4.1.3). We ran a model with *Prime Condition* (baseline/active/passive) and *Intervention* (pre/post) and its interactions as fixed factors. The baseline *Prime Condition* and *Pre-intervention* served as the reference level. Following the maximal random effects structure as proposed by Barr et al. (2013), we added random slopes and random intercepts for *Participants* and *Items*. The maximal model was simplified in a stepwise way due to convergence and singularity issues. We first simplified the random effects structure by testing if the random slope terms could be omitted without decreasing the fit of the model. We removed random slopes for *Item* before removing any random slopes for *Participant* because the variance in items is usually smaller than the variance in participants (Segaert et al. 2016). For the final model, we calculated the conditional and marginal R^2 values, which are measures of the effect size, with the *rsquared* function from the piecewiseSEM package (version 2.1.0., Lefcheck, 2016).

2.2 Results

The native speakers produced 505 responses, of which 395 were active responses (74.2%), 110 passive responses (20.7%) and 27 "Other" responses (5.1%). Figure 4 shows the proportion of active and passive responses of the native speakers per *Prime Condition* before and after the intervention block, and the responses during the intervention block.

Figure 4. Proportion of active and passive responses before, after and in the intervention block for the control group (experiment 1).



The final model included random intercepts for *Participant* and *Item* and no random slopes. In the control group, we found significant passive priming ($p < .05$) and significantly more passives after the intervention block ($p < .01$) (see Table 1) than before the intervention block. There was no significant interaction between *Prime Condition* and *Intervention*[1]. The fixed effects of the final model explained 3.94% of the variance (marginal R^2; Nakagawa & Schielzeth, 2013) and conditional on the random effects, they explained 13.40% of the variance.

1 Descriptively, there seems to be an interaction between *Prime Condition* and *Intervention*. However, this was not confirmed in our model, probably due to low statistical power.

Table 1. Summary of fixed effects of generalized linear mixed model (N = 431, log-likelihood = -149.5, experiment 1).

	Coefficient	*SE*	Wald's *Z*	*p*-value
(Intercept)	-3.19	0.46	-6.95	<.001***
Condition(ACT)	0.05	0.41	0.11	0.91
Condition(PASS)	0.93	0.37	2.50	<.05
Intervention(Post)	0.95	0.32	2.94	<.01

2.3 Discussion

As expected, we found only passive priming for the control group. This result is in line with the inverse-preference effect (Chang et al., 2006), where the less frequent syntactic alternant induces stronger priming than the most frequent syntactic alternant due to surprisal. Interestingly, the native speakers were strongly influenced by the intervention block: they produced significantly more passives after the intervention compared to before the intervention block. Moreover, whereas they produced 25.7% passives *during* the intervention block, this increased to 27.8% after the intervention. To our knowledge, this effect of our methodological manipulation has not been found for native speakers yet. Even though native speakers have already established a firm syntactic representation for passives, a few instances of verb overlap in the same syntactic structure boost passive sentence production such that more passive responses (in all conditions) are observed in subsequent trials without verb overlap. This shows the importance of our priming manipulation and, presumably, this

manipulation will have a stronger effect in late learners as they might not have yet established an equally strong syntactic representation for the passive structure.

3. Experiment 2: Longitudinal Study

3.1 Method

3.1.1 Participants

For the longitudinal study, we recruited participants who were enrolled in a one-year Dutch language course at a language institute in Antwerp, Belgium. Seventeen participants (10 female, 7 male, between 18-45 years [M = 24.1, SD = 6.3])) volunteered to participate five times throughout the academic year. The participants had varying first languages (3 Arabic, 3 Russian, 2 Persian, 2 Spanish, 2 Turkish, 1 French, 1 Tajiks, 1 Thai, 1 Afrikaans/English and 1 Ukrainian/Russian). All participants spoke English as their L2 (sometimes in addition to other languages) and learned Dutch as their L3, L4 or L5. At the end of the fifth session, participants received monetary compensation for their participation.

3.1.2 Materials

The materials of Experiment 2 were identical to those of the control group experiment. We used the same materials in each session.

3.1.3 Design

The design of Experiment 2 was similar to that of the control group experiment (Experiment 1) for the Sessions 2, 3, and 4. However, at Session 1 and Session 5, we added comprehension trials to the experiment to collect more data on whether participants

interpreted the prime sentences correctly[2]. We created another 72 items, of which 24 were critical items. Both the comprehension prime trials and the comprehension target trials had the same design as the production prime trials. That is, participants would hear the prime sentence while performing a verification task. Half of the comprehension target trials were actives, and the other half were passives. Each of the active and passive target trials were preceded by a baseline, an active or a passive prime sentence. We divided both the production and the comprehension trials into blocks of 12 items each, leading to six production blocks and six comprehension blocks, which alternated each other. There was still an intervention block of four items (two were comprehension trials and two were production trials) halfway through the experiment. By reversing the order of the trials in the experimental list, we did not only establish that all items preceding the intervention in the first six lists followed the intervention in the last six lists, it also ensured that half of the lists started with a production block and half of the lists started with a comprehension block.

Even though all participants participated in the same language course, individual proficiency still varied, for instance, due to differences in exposure to Dutch outside the language course. Therefore, at Session 4, we had participants take the LexTALE language test (Lemhöfer & Broersma, 2012), to objectively measure participants' general Dutch proficiency. Importantly, the LexTALE test has been validated and tested to be a reliable predictor of L2 proficiency. Lemhöfer and Broersma showed that the LexTALE scores strongly correlate with self-rating scores on writing, reading, listening and speaking proficiency. Moreover, we used the LexTALE for practical reasons too: it only takes approximately 5 minutes to complete, and since we were testing beginning learners, we did

[2] Originally, we had planned to add the comprehension trials in all sessions. However, during Session 1 it turned out that the experiment was too long and intensive for the participants. Therefore, we decided to leave out the comprehension trials in Session 2, 3 and 4. To still be able to measure the increase of correct interpretations of the transitive prime sentences, we decided to reinsert the comprehension trials at Session 5. By then, participants had become more proficient in Dutch and the average completion time was much lower than in Session 1.

not want to subject them to a long language experience questionnaire (e.g., LEAPQ test - Marian et al, 2007). We assumed that an estimation of their Dutch vocabulary size would be informative enough regarding their familiarity with Dutch words.

3.1.4 Procedure

Participants were tested five times over the course of eight months. There were about six weeks between each session. Participants completed the same experiment during each session, but they always received a different list per session.

The procedure was similar to that of the control group, except that we provided them with assistance. Due to the COVID-19 pandemic, participants were assisted remotely through the phone. We employed research assistants who recorded and noted down all target sentences produced by participants. Prior to Session 1, we sent participants a booklet with illustrations and translations of all the vocabulary used in the experiment to allow them to familiarize themselves with the vocabulary. We also provided participants with a short demonstration video that demonstrated how the experiment would look. For each session, the research assistant called the participant, and first verified whether they understood the setup of the experiment and the assistant repeated the instructions if necessary. Crucially, our research assistants had to adapt to each participant differently depending on how well each participant comprehended Dutch. Naturally, during the earlier sessions, more help was needed than during the later sessions.

The experiment lasted approximately 50 minutes in Session 1. The duration decreased over the course of time. At Session 5, the average completion time was 20 minutes. At Session 1, participants were asked to complete a short questionnaire with questions on their demography and language background. At Session 4, the experiment was followed by the LexTALE language test (Lemhöfer & Broersma, 2012), which took approximately 5 minutes.

3.1.5 Coding

The coding of responses was identical to that of the control group, except for the treatment of grammatically incorrect responses. Because conjugating the past participle in Dutch is regarded to be difficult for L2 learners due to irregular verbs, we allowed all attempts of the past participle (e.g., **De auto wordt achter**ge**volgd door de motor* [correct: De auto wordt achtervolgd door de motor – The car is being chased by the motorcycle]). If participants used the infinitive form of the verb rather than an (attempted) conjugated form or produced ungrammatical sentences, we coded these responses as "Other". If participants did not produce a target sentence at all, we coded these responses as "null" responses. "Other" and "null" responses were disregarded in our analyses.

3.1.6 Analysis

Priming effects were measured in a similar way as for the control group (i.e., comparing proportions of active and passive responses after a prime condition vs. the baseline [unprimed] condition) in a GLMER model. Our model consisted of the factors *Prime Condition* (baseline/active/passive), *Session* (1 to 5) and *Intervention* (pre/post) and its interactions as fixed factors. The baseline *Prime Condition* and *Pre-intervention* served as the reference level. *Session* was treated as an ordinal variable. We started with a maximal random effects structure and simplified it until convergence.

3.2. Results

We collected 2232 responses, of which 1105 were actives (49.5%), 666 passives (29.9%) and 461 "Other" responses (20.7%). Table 2 shows the responses per session.

Table 2. Active, Passive and Other responses per session (experiment 2).

Session	Active	Passive	Other	Total
1	192 (46.2%)	31 (7.5%)	193 (46.4%)	416 (*n = 16)*
2	272 (57.1%)	78 (16.4%)	126 (26.5%)	476 (*n* = 17)
3	195 (43.5%)	174 (38.8%)	79 (17.6%)	448 (*n = 16)*
4	232 (48.7%)	195 (41.0%)	49 (10.3%)	476 (*n* = 17)
5	214 (51.4%)	188 (45.2%)	14 (3.4%)	416 (*n = 16)*

The total number of participants differs per session due to technical issues or illness of participants.

We excluded the "Other" responses from further analyses[3]. Note that the number of "Other" responses strongly decreases, while the proportion of passive responses increases over time. At the same time, the proportion of active responses remains somewhat around the same percentage throughout the five different sessions. This suggests that the increase in passives mainly comes from the decrease in "Other" responses. However, we excluded the "Other" responses in our statistical analysis (see below). This means that we have more observations in the later sessions than in the earlier sessions, mainly due to the increase in passive responses. In our analysis, we only focus on the *proportion* of active and passive responses rather than absolute numbers. Thus, an increase in passive responses naturally goes at the expense of the proportion of active responses.

Figure 5 shows the proportion of passive responses per prime condition, per session, per

[3] The proportion of "Other" responses strongly decreased over time. In the earlier sessions, participants often skipped targets, saying that they did not know how to produce a sentence. In addition, they sometimes produced alternative, non-transitive structures to describe the sentence, such as prepositional phrases. In the later sessions, participants were familiar with the experiment and the sentence structures used in it (participants may have interpreted the prime sentences as examples of desired target responses) and produced fewer alternative structures.

prime condition before and after the intervention. Figure 6 shows the proportion of passive responses per session during the intervention.

Figure 5. The proportion of passive responses per session, per prime condition before and after the intervention (experiment 2).

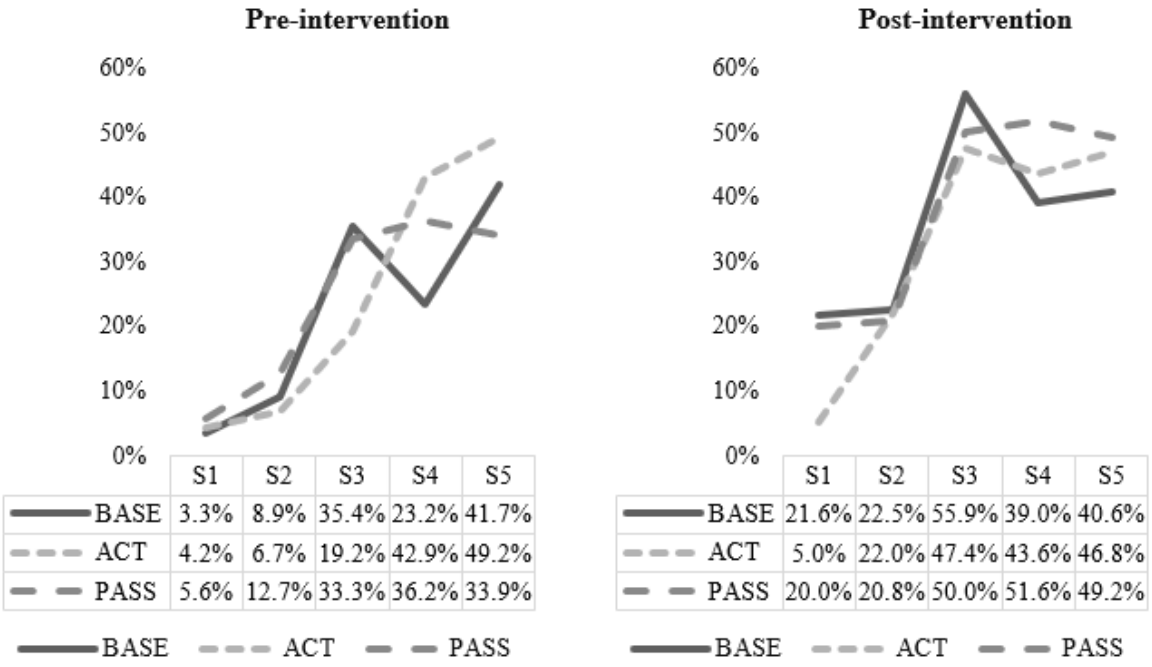
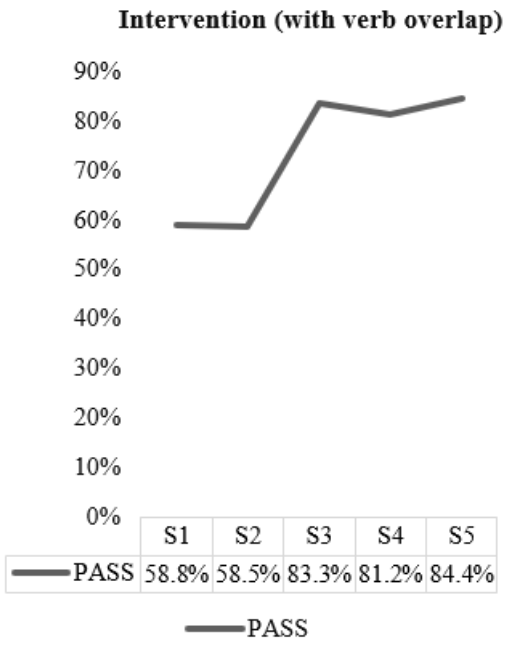


Figure 6. The proportion of passive responses per session in the intervention.



We fit the target responses to a generalized linear mixed model. The final model included random intercepts for *Participant* and *Item* but no random slopes. The fixed effects of the final model explained 18.68% of the variance (marginal R^2; Nakagawa & Schielzeth, 2013) and conditional on the random effects, they explained 46.65% of the variance. In Table 3, we report the model output.

Table 3. Summary of fixed effects of generalized linear mixed model (N =1551, log-likelihood = -710.4, experiment 2).

	Coefficient	*SE*	*Wald's Z*	*p*-value
(Intercept)	-2.04	0.41	-5.01	<.001***
Condition(ACT)	-0.26	0.21	-1.22	0.22
Condition(PASS)	0.11	0.18	0.60	0.55

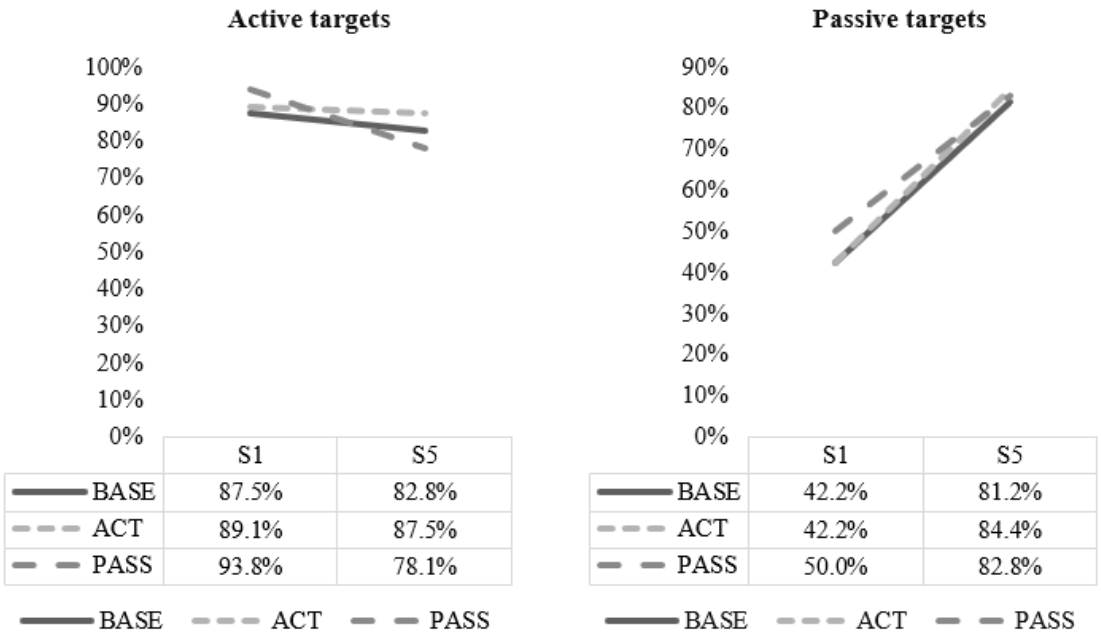
Sessions(L)	2.52	0.47	5.35	<.001***
Sessions(Q)	-0.60	0.42	-1.42	0.15
Sessions(C)	0.17	0.38	0.44	0.66
Sessions(4)	0.85	0.33	2.61	<.01**
Intervention(Post)	1.10	0.18	6.13	<.001***
Condition(ACT)*__Sessions__(L)	1.55	0.56	2.78	0.01
Condition(PASS)*__Sessions__(L)	0.27	0.45	0.59	0.55
Condition(ACT)*__Sessions__(Q)	-0.50	0.50	-0.99	0.32
Condition(PASS)*__Sessions__(Q)	-0.01	0.42	-0.02	0.98
Condition(ACT)*__Sessions__(C)	0.08	0.45	0.17	0.86
Condition(PASS)*__Sessions__(C)	-0.52	0.41	-1.27	0.20
Condition(ACT)*__Sessions__(4)	-1.01	0.38	-2.65	<.01**
Condition(PASS)*__Sessions__(4)	-0.65	0.37	-1.76	<.1.
Sessions(L)*__Intervention__(Post)	-1.12	0.46	-2.42	<.05*
Sessions(Q)*__Intervention__(Post)	-0.21	0.42	-0.50	0.62
Sessions(C)*__Intervention__(Post)	-0.03	0.37	-0.10	0.92
Sessions(4)*__Intervention__(Post)	0.15	0.32	0.48	0.63

Note: the variable Sessions is an ordinal variable. The model uses polynomial contrasts. L refers to a linear predictor, C to a cubic, Q to a quadratic and 4 to the fourth derivative.

The results indicate that the proportion of passives increases linearly over time ($p < .001$), then stabilizes ($p < .01$). Participants also produce more passives after the intervention than before the intervention ($p < .001$), but this effect decreases linearly over time ($p < .05$). There is a significant interaction between the active prime condition and the fourth derivative of *Session* ($p < .01$), indicating that participants produce more active sentences after an active prime sentence (i.e., active priming) in the earliest session. The active priming effect diminishes in the next two sessions. In Session 4, participants produce more passive sentences after an active prime than after a baseline prime. This effect disappears again in Session 5.

At Session 1 and Session 5, we also included blocks with comprehension target items. At Session 1, participants indicated the correct picture corresponding to the target item in 67.4% of the cases. Accuracy was much higher for active (90.1%) than for passive sentences (44.8%). In the intervention, they were correct in 37.5% of the sentences (note that there were only passive sentences in the intervention, which were more difficult to understand for our participants). At Session 5, the accuracy increased to 82.5%, and the accuracy was identical for active and passive sentences. In the intervention, participants responded correctly in 87.5% of the sentences. Figure 7 shows the proportion of correct responses per Prime condition for both sessions.

Figure 7. Correct responses in the comprehension targets, per prime condition per Session (experiment 2).



	S1	S5
BASE	87.5%	82.8%
ACT	89.1%	87.5%
PASS	93.8%	78.1%

	S1	S5
BASE	42.2%	81.2%
ACT	42.2%	84.4%
PASS	50.0%	82.8%

At Session 4, participants performed the LexTALE test to measure their proficiency. Scores varied between 48.75% and 78.75%. The mean score was 58.0% (SD: 7.0%). Exploratory analyses showed that the LexTALE scores did not affect the outcome variable in the priming experiment.

3.3 Discussion

The results of the longitudinal study show an increase of the use of the passive structure over sessions. The growth decreases after Session 4. We find structural priming effects of the active structure at Session 1 after the intervention and significantly more passives after an active prime than after a baseline prime at Session 4 before the intervention. Especially for the earlier sessions, participants produce more passives after the intervention than before the intervention. This pattern is in accordance with our expectations, as the intervention block seems to promote the learning of passives, which is especially relevant for learners at earlier stages of language development who still rely on item-specific representations rather than abstract structural representations.

Nevertheless, in contrast to our expectations, we did not find priming of the passive structure in the later sessions of the longitudinal study (although the increase in passives over sessions may be partially the result of long-term priming). During the later sessions, the learners were certainly proficient enough to produce the passive structure, given the overall high proportion of passive responses (especially after session 3 and onwards, after the learners have received explicit instruction on the use of the passive) and the high accuracy rates in comprehension at Session 5. Because the passive structure is more complex than the active sentences and the intransitive filler sentences, participants may have become aware of the fact that we were interested in the passive structure. Interestingly, in the debriefing survey after the final session, participants reported to have noticed our interest in the passive structure during the experiment. This may have led to a relatively high overall proportion of passive sentences in the later sessions. Importantly, the surprisal effect that plays a role in inducing priming effects may thus have been weaker due to the repeated sessions, although it is also possible that other mechanisms may have been at play. Therefore, it is important to cross-validate our findings in a cross-sectional study.

The awareness of the passive structure may also explain the significantly higher proportion of passive sentences after an active prime than after a baseline prime at Session 4 before the intervention. A transitive prime sentence, be it an active or a passive prime sentence, may have made participants more aware of the choice between the active and the passive sentence while producing the target sentence. The motivation of the participants to demonstrate their abilities and to learn from the experiment themselves may have triggered them to try to produce the more complex passive structure. As such, the longitudinal design turned out to be useful to test the very early stages of language learning, that is, to prime a structural alternation before one of the alternatives was learned. However, the repeated measures design did not allow us to follow the developmental path of language learning.

4. Experiment 3: Cross-Sectional Study

4.1 Method

4.1.1 Participants

We recruited participants who were taking Dutch classes at a center for adult education based in Antwerp (Belgium). For the lower proficiency group, we recruited students who were learning Dutch in level 2 (A1/A2 level). For the higher proficiency group, we recruited students from level 3 and 4 (B1/B2 level).

Eighteen participants from the low proficient group and 20 participants from the high proficient group participated in our study. There were 23 females and 15 males, who were between 19 and 53 years (M=33.7, SD= 7.6). The participants had varying first languages (Tigrinya, Turkish, Twi, Urdu…), but Arabic occurred the most (there were 7 native speakers of Arabic). Most participants indicated English as their L2 (amongst other languages), which means Dutch was either their L3, L4, L5 or L6. All participants gave their consent before participating in our experiment and received a monetary reward for their participation.

4.1.2 Materials

The materials of Experiment 3 were identical to those of Experiment 2.

4.1.3 Design

The design was similar to the design of Experiment 2. We did not include the comprehension blocks from the longitudinal study. Different from Experiment 1 and 2, however, is that this experiment was run in Qualtrics (Qualtrics, Provo, UT). We employed Qualtrics for this experiment because a pilot test (N = 4) showed that using *PsychoPy* was too effortful for the participants. As a result, we were not able to collect a single complete datafile during the pilot

study. In addition, Qualtrics is not only computer friendly but also mobile phone friendly, as some participants in this study only had a mobile phone at their disposal.

4.1.4 Procedure

The procedure of Experiment 3 was largely similar to that of Experiment 2. Participants were also tested remotely through the phone with the help of research assistants. For Experiment 3, we filmed two short demonstration videos: one for personal computer users and one for mobile phone users.

Participants had to manually press the play button to listen to the prime sentence in Qualtrics, and thus, the research assistants emphasized that they could only play the audio prime sentences once. At the end of the experiment, participants were asked to fill out the same background questionnaire as in Experiment 2. The experiment lasted approximately 50 minutes for the learners in level 2 and 25 minutes for the learners in level 3 and 4.

4.1.5 Coding

We coded the responses of Experiment 3 according to the same coding scheme as that of Experiment 2.

4.1.6 Analysis

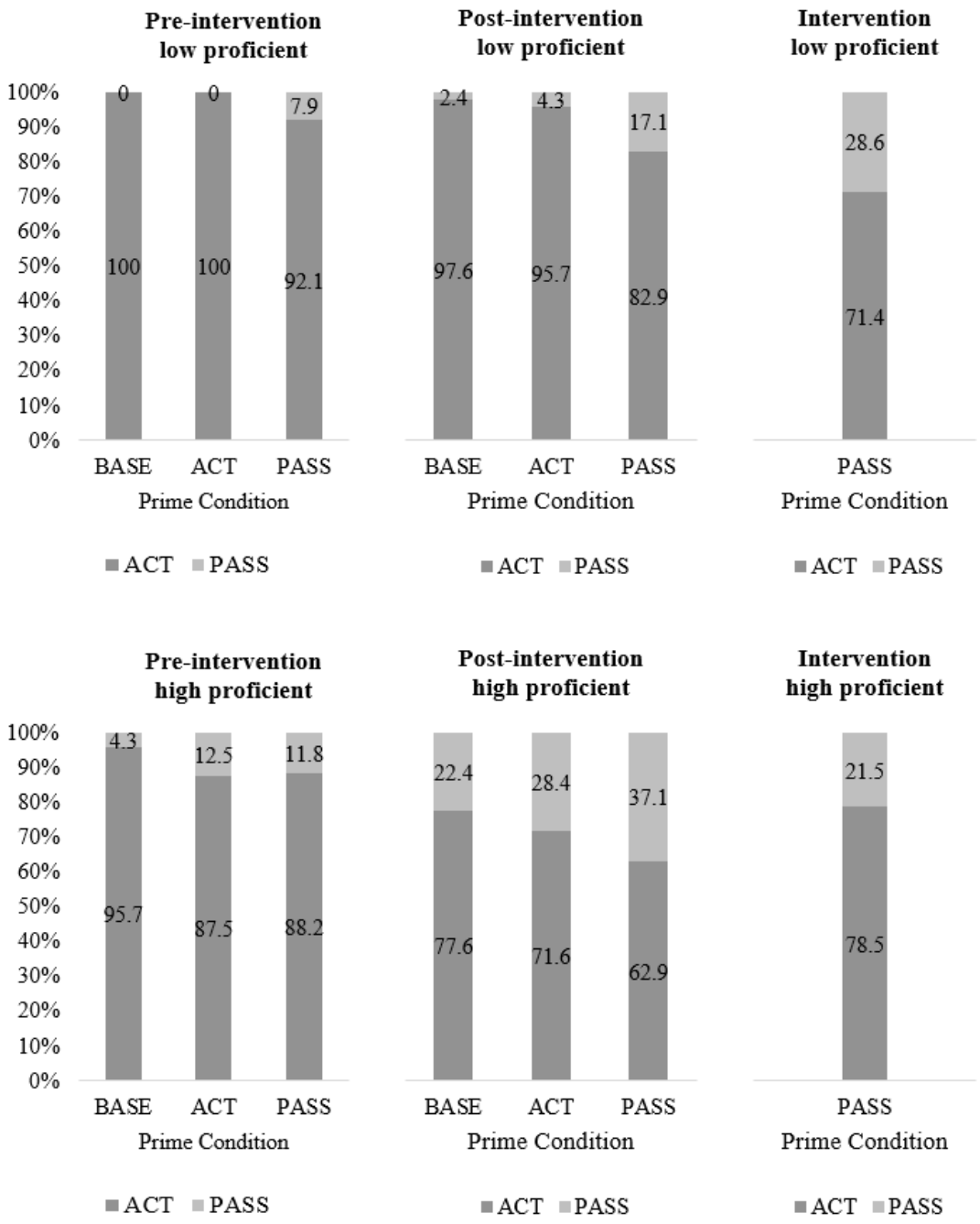
Our full model consisted of a three-way interaction between *Condition* (baseline, active, passive) * *Proficiency* (lower proficiency learners vs. higher proficiency learners) * *Intervention* (pre vs. post). The baseline condition, the lower proficiency speakers and the pre-intervention were the reference levels. The dependent variable *Target response* (active vs. passive, with active as reference level) was binary. Similar to the analysis of Experiment 1 and 2, we used a Bobyqa optimizer to increase convergence. The maximal random effects

structure consisted of *Condition* and *Intervention* as random slopes within *Participant*, and *Condition, Proficiency,* and *Intervention* as random slopes within *Item*. Moreover, we simplified our maximal model until it converged, and no singularity issues were detected (similar to Experiment 1 and 2).

4.3 Results

The 38 participants produced a total of 1,064 responses, of which 616 (57.9%) were active sentences, 169 (15.9%) passive sentences and 279 (26.2%) 'Other' responses. There was one participant in the low proficient group who only produced 'Other' responses, and thus, we excluded this participant from our analyses. Figure 8 shows the proportion of active and passive responses of the lower and higher proficiency speakers before, after and during the intervention. Before the intervention, the lower proficiency speakers only produced passives when they were primed with a passive structure, whereas the higher proficiency speakers produced passives in all conditions. Interestingly, before the intervention, the higher proficiency speakers produced slightly more passives in the active prime condition than in the passive prime condition. After the intervention, both groups of speakers produced passives in all conditions, with the highest proportion of passives in the passive prime condition.

Figure 8. Proportion of active and passive responses before and after the intervention block for the lower and higher proficiency speakers (experiment 3).



Our final model consisted of significant main effects for *Condition* ($p < .001$), *Proficiency* ($p < .001$) and *Intervention* ($p < .001$). The fixed effects of the final model explained 14.84% of the variance (marginal R^2; Nakagawa & Schielzeth, 2013) and conditional on the random effects, they explained 28.20% of the variance.

We found significant passive priming for the lower and higher proficiency learners (*p*

< .001). We also observed a main effect of *Proficiency*: higher proficiency speakers produced significantly more passives in the baseline condition compared to lower proficiency speakers (p < .001). Lastly, the number of passives increased significantly in the baseline condition in the post-intervention block compared to the baseline condition in the pre-intervention block (p < .001) for both proficiency levels. Interestingly, we observed a marginal negative structural priming effect for actives (see third row, Table 4). This marginal effect suggests that the learners produced more *passives after an active prime condition than after the baseline condition.* This tendency can also be seen in Figure 8 where, for instance, descriptively, the higher proficiency speakers produced more passives when they were primed with an active prime than when they encountered a baseline prime. We did not find interactions between our predictors, probably because of the small number of observations in the data.

Table 4. Summary of fixed effects of generalized linear mixed model (N = 664, log-likelihood = -218.9, experiment 3).

Fixed effects	Coefficient	SE	Wald's Z	*p*-value
Intercept	-5.04	0.65	-8.30	<.001***
Condition(ACT)	0.63	0.34	1.83	<.1.
Condition(PASS)	1.21	0.33	3.65	<.001***
Proficiency(HighProficient)	2.04	0.59	3.43	<.001***
Intervention(Post)	1.50	0.28	5.50	<.001***

4.3 Discussion

The results for the cross-sectional study show that proficiency, the intervention with lexical overlap and abstract priming of the passive all contribute to the production of passives.

Participants who are more proficient, produced more passives than participants who are less proficient. The intervention positively affected the production of passives in both groups, which is noticeable in the increase of passives in the post-intervention block. This highlights the importance of our intervention manipulation. Though we expected to find different priming patterns for the lower and higher proficiency learners, the results show no difference between the two groups, as both types of learners showed significant passive priming. Importantly, the lower proficiency speakers were probably not at the very beginning stages of learning Dutch (like the learners in the first session of the longitudinal study), since they might have already been exposed to the passive sentence structure and may have formed a syntactic representation for this structure. Moreover, both the lower and higher proficiency learners produced more passives after the intervention block than before the intervention, which goes against our expectations. This is because we predicted that the lower proficiency learners would show a larger increase in the number of passives post intervention than pre-intervention compared to the higher proficiency speakers. In fact, note that numerically, the higher proficiency learners display a larger increase in passive responses than the lower proficiency learners *within* the intervention (see General Discussion).

5. General Discussion

Revisiting the research questions

RQ1*: When in the learning trajectory of transitive structures do late learners of Dutch show priming for active and passive sentences?*

We found active priming only in the very beginning stage of learning (i.e., only at Session 1 of the longitudinal study), implying that active priming becomes weaker with increasing proficiency. Note that it is not surprising that active priming disappears over time, since L2 production patterns might become more native-like with increasing proficiency (Hartsuiker & Bernolet, 2017; Montero-Melis & Jaeger, 2019). In native speakers, due to the inverse

preference effect, active priming is usually not found, while strong passive priming is observed. Nevertheless, in Experiment 2, the disappearance of active priming over time is probably not the result of the inverse preference effect since active priming disappeared but, unexpectedly, we did not observe passive priming (note, however, that the proportion of passives increased in later sessions compared to earlier ones, see *Discussion* of Experiment 2). Active priming may therefore have decreased for other reasons.

More specifically, in Session 4 of the longitudinal study and in the cross-sectional study, we found a negative structural priming effect in the active prime condition. That is, participants produced more passive sentences after an active prime than after a baseline prime. A transitive prime sentence, be it an active or a passive prime sentence, may have made the high proficient learners more aware of the choice between both transitive structures while producing a target sentence. They might also have practiced the alternation between active and passive sentences explicitly during their language course. Consequently, an active prime may have reminded them of the passive structure. As most participants indicated that they believed that we were testing their knowledge of the passive, they might have chosen to produce the more complex alternative even after an active prime. As a result, the structural priming effect of the active structure might have become weaker over subsequent sessions, and even negative in Session 4.

The number of passives produced across prime conditions is larger for higher proficiency learners than for lower proficiency learners, both in the longitudinal study and the cross-sectional study, suggesting that the abstract structural representation for passives becomes stronger upon increasing proficiency. This is also reflected in the passive priming effects. In the longitudinal study, we do not find passive priming, although some passives are produced across prime conditions already in Session 1. We assume that abstract syntactic

representations for the passive structure are not present in the very early stage of language learning, and the production of passives may be the result from L1 transfer or a copy-edit strategy (Bernolet et al., 2016; Bernolet & Hartsuiker, 2018). In the cross-sectional study, we find significant passive priming in both the lower proficiency and the higher proficiency speakers. The lower proficiency speakers in the cross-sectional study, who were probably not at the very beginning stages of language learning, may have already been exposed to the passive structure (e.g., during reading), and therefore, might have formed a syntactic representation for this structure, at least in comprehension.

RQ2*: To what extent do several instances of verb overlap in passives (the more complex transitive alternant) boost the production of passive sentences in subsequent trials without verb overlap?*

Participants produce more passives after the intervention than before the intervention, which is the case in our L1 control group as well as in the L2 learners in the longitudinal and the cross-sectional study. This suggests that the intervention boosts the production of passive sentences in subsequent trials without verb overlap, at least partly due to increased attention towards the passive structure. In addition, participants in the longitudinal study showed a stronger increase of passives after the intervention in the earlier sessions than in the later sessions, which implies that very low proficient learners benefit to a greater degree from the intervention in subsequent trials without verb repetition than more proficient learners. This proficiency effect suggests that the intervention may accelerate the development of abstract structural representations in participants who have not yet developed an abstract structural representation of the passive.

We also looked at the effects of the intervention with regard to what happens *within* the intervention block. In the intervention items itself, descriptively, more passives were

produced by the higher proficiency learners in the cross-sectional study (experiment 3) than by the lower proficiency learners. This was unexpected since the developmental account of Hartsuiker and Bernolet (2017) predicts that lower proficiency learners rely more on verb overlap between primes and targets than higher proficiency speakers. Similarly, participants produced more passives in the intervention in the later sessions of the longitudinal study than in the earlier sessions. Thus, although we observed a strong lexical boost effect in the lower proficiency speakers, it seems that the passive was boosted stronger in the higher proficiency learners than in the lower proficiency speakers during the intervention block. There are possibly two main components leading to the lexical boost effect, namely explicit memory, and residual activation of the structural representation of the passive. Whereas explicit memory plays a role in both the lower and higher proficiency learners (explaining the increase in the use of passives during the intervention in both groups), residual activation may only be present in learners with a stored structural representation of the passive. Lower proficiency learners who only display structural priming due to a copy-edit strategy (i.e., explicit memory) may lack an actual mental representation of the passive which is activated during processing, and this may be why higher proficiency learners display a larger lexical boost effect than lower proficiency learners. Indeed, we also find a large lexical boost effect in the intervention in our control group of native speakers. Moreover, the higher proficiency speakers probably used more passives than the lower proficiency speakers because the passive is a complex structure which required conjugating the past participle. It could have simply been the case that the lower proficiency speakers did not produce as many passives as the higher proficiency speakers during the intervention due to the complexity of the passive structure.

6. Limitations and Future Directions

A limitation of the current study is that our sample size did not meet the recommendations of

Mahowald et al. (2016) to reach sufficient statistical power. We were not able to find enough participants, since we tested a very specific group of learners; and we could not increase the number of test items as the task was cognitively demanding for the participants. For this reason, our results should be interpreted with some caution. Nevertheless, we still have an estimated power to detect abstract structural priming of more than 60% for the longitudinal group (16 critical items (excluding baseline items * 5 sessions * 17 participants), and about 50% in the cross-sectional experiment (16 items * 38 participants), not taking into account the long-lasting lexical boost effect of the intervention. Importantly, we believe that our findings are a first step in answering how L2 syntactic representations are acquired in a natural language learning setting, but future studies should include more participants to better understand the process of late L2 syntactic learning.

We are aware that our testing method happened in an unconventional manner (i.e., participants were tested at a distance, while being assisted through the phone) due to the COVID-19 pandemic. Although most participants used their personal computer to participate, a few only had a mobile phone at their disposal. This could have caused more noise in the data compared to conventional lab-testing. However, despite the circumstances, our method highlights the robust strength of the structural priming paradigm: the tendency to repeat syntactic structures does not only occur in a lab setting, where participants may be aware of experimental manipulations, but it also occurs in people's homes, where experimental manipulations may be less apparent. Also, our method may open new avenues for conducting online research (see for example, Slim & Hartsuiker, 2022).

Our data pinpoints a possible shortcoming of the developmental account (Hartsuiker & Bernolet, 2017), namely that it is based on the residual activation model of Pickering and Branigan (1998). Not only the low proficient speakers, but also the high proficient speakers as well as the native speakers produced more passives after the intervention than before the

intervention. So, a few instances of lexical overlap boost the passive structure in subsequent items without lexical overlap between prime and target. The residual activation model does not predict long-lasting lexical boost effects. At the same time, the implicit learning model does not have a straightforward explanation for the lexical boost effect itself (though Chang et al. [(2006]) argue that lexical enhancement effects are due to explicit memory traces of the prime structure). Still, the implicit learning model may predict these effects indirectly, assuming that the items with verb overlap induce stronger explicit memory traces that enhance implicit learning (Chang et al., 2006). Because of the lexical boost effect, participants produced more passives during the intervention items than they would have done in critical items without verb overlap. As a result, the number of passives heard and produced by participants is higher, which may have led to stronger implicit learning. Because of this enhanced implicit learning, they produced more passives after the intervention than they did before the intervention. A bilingual model of L2 development should therefore probably be a hybrid model (cf. Reitter et al., 2011; Momma, 2022; for monolingual hybrid models of structural priming), integrating both implicit learning mechanisms and residual activation. This suggestion has also been made by Hartsuiker and Bernolet (2017), and the results of this study strengthen the need for such a hybrid model.

7. Conclusion

Altogether, our results suggest that priming of the active structure takes place before priming of the passive structure, in accordance with our hypotheses. Nevertheless, abstract representations of the passive structure seem to be formed quite rapidly after exposure to the structure. The very early learners, who did not show passive priming, displayed a larger increase in their production of the passive in the intervention block than the learners who had acquired the passive, suggesting that the intervention items with lexical overlap speeded up

the formation of abstract representations due to implicit learning. Our results indicate that, ideally, a developmental model of L2 syntax should be a hybrid model, incorporating aspects of the residual activation theory as well as an implicit learning mechanism.

Materials, data, and analyses are available online: https://osf.io/x8ejm

Ethics approval: Approval was obtained from the ethics committee of Social Sciences and Humanities at the University of Antwerp (SHW_18_77).

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2009). Persistence of emphasis in language production: A cross-linguistic approach. *Cognition*, *112*(2), 300-317. https://doi.org/10.1016/j.cognition.2009.05.013

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition*, *127*(3), 287-306. https://doi.org/10.1016/j.cognition.2013.02.005

Bernolet, S., Collina, S., & Hartsuiker, R. J. (2016). The persistence of syntactic priming revisited. *Journal of Memory and Language*, *91*, 99-116. https://doi.org/10.1016/j.jml.2016.01.002

Bernolet, S. & Hartsuiker, R. J. (2018). Syntactic representations in late learners of a second language: a learning trajectory. In *Bilingual cognition and language: the state of the science across its subfields* (Vol. 54, pp. 205-224). John Benjamins.

Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 575-586. https://doi.org/10.1016/0010-0285(86)90004-6

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234. https://doi.org/10.1037/0033-295X.113.2.234

Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes, 21*(7-8), 1011–1029. https://doi.org/10.1080/016909600824609

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. *Journal of Memory and Language*, *58*(2), 214-238. https://doi.org/10.1016/j.jml.2007.07.003

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, *15*(6), 409-414. https://doi.org/10.1111/j.0956-7976.2004.00693.x

Hartsuiker, R.J., & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition, 20*, 219-234. https://doi.org/10.1017/S1366728915000164

Khoe, Y. H., Tsoukala, C., Kootstra, G. J., & Frank, S. L. (2023). Is structural priming between different languages a learning effect? Modelling priming as error-driven implicit learning. *Language, Cognition and Neuroscience*, *38*(4), 537-557. https://doi.org/10.1080/23273798.2021.1998563

Lefcheck, J.S. (2016). piecewiseSEM: Piecewise structural equation modeling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573-579. https://doi.org/10.1111/2041-210X.12512

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325-343. https://doi.org/10.3758/s13428-011-0146-0

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5-27. https://doi.org/10.1016/j.jml.2016.03.009

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940-967. https://doi.org/10.1044/1092-4388(2007/067)

Momma, S. (2022). Producing filler-gap dependencies: Structural priming evidence for two distinct combinatorial processes in production. *Journal of Memory and Language*, *126*, 104349. https://doi.org/10.1016/j.jml.2022.104349

Montero-Melis, G., & Jaeger, T. F. (2020). Changing expectations mediate adaptation in L2 production. *Bilingualism: Language and Cognition*, *23*(3), 602-617. https://doi.org/10.1017/S1366728919000506

Muylle, M., Bernolet, S., & Hartsuiker, R. J. (2021a). The development of shared syntactic representations in late L2-learners: Evidence from structural priming in an artificial language. *Journal of Memory and Language*, *119*, 104233. https://doi.org/10.1016/j.jml.2021.104233

Muylle, M., Bernolet, S., & Hartsuiker, R. J. (2021b). On the limits of shared syntactic representations: When word order variation blocks priming between an artificial language and Dutch. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000997

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133-142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195-203. https://doi.org/10.3758/s13428-018-01193-y

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*(4), 633-651. https://doi.org/10.1006/jmla.1998.2592

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*(3), 427. https://doi.org/10.1037/0033-2909.134.3.427

Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, 26-46.

Qualtrics (2021, first release 2005). *Qualtrics*. Provo: Utah. Available at https://www.qualtrics.com.

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, *35*(4), 587-637. https://doi.org/10.1111/j.1551-6709.2010.01165.x

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research*, *12*(1), 40-72. https://doi.org/10.1177/026765839601200103

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*(10), e24209. https://doi.org/10.1371/journal.pone.0024209

Segaert, K., Wheeldon, L., & Hagoort, P. (2016). Unifying structural priming effects on syntactic choices and timing of sentence generation. *Journal of Memory and Language*, *91*, 59–80. https://doi.org/10.1016/j.jml.2016.03.011

Slim, M. S., & Hartsuiker, R. J. (2022). Moving visual world experiments online? A web based replication of Dijkgraaf, Hartsuiker, and Duyck (2017) using PCIbex and WebGazer. js. *Behavior Research Methods*, 1-19. https://doi.org/10.3758/s1342802201989-z

Wonnacott, E., Newport, E. L., & Tanenhaus, M. K. (2008). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*, *56*(3), 165-209. https://doi.org/10.1016/j.cogpsych.2007.04.002