A Double Mind

Syntactic representations in bilingual speakers

Rianne van Lieburg

Supervisors **dr. Sarah Bernolet | prof. dr. Robert J. Hartsuiker**

Thesis submitted for the degree of Doctor of Linguistics
Faculty of Arts | Department of Linguistics | Antwerp, 2023

University of Antwerp

Faculty of Arts
Department of Linguistics

A Double Mind

Syntactic representations in bilingual speakers

Thesis submitted for the degree of
Doctor of Linguistics
at the University of Antwerp
to be defended by

Rianne van Lieburg

Supervisors
dr. Sarah Bernolet
prof. dr. Robert J. Hartsuiker

Antwerp, 2023

Jury

Chairman
prof. dr. Frank Brisard, University of Antwerp, Belgium

Supervisors
dr. Sarah Bernolet, University of Antwerp, Belgium
prof. dr. Robert J. Hartsuiker, Ghent University, Belgium

Members
dr. Gerrit Jan Kootstra, Radboud University Nijmegen, the Netherlands
prof. dr. Dominiek Sandra, University of Antwerp, Belgium
dr. Katrien Segaert, University of Birmingham, United Kingdom
prof. dr. Reinhild Vandekerckhove, University of Antwerp, Belgium

Contact
Rianne van Lieburg

University of Antwerp
Faculty of Arts
Department of Linguistics
Prinsstraat 13, 2000, Antwerp, Belgium
Rianne.vanLieburg@UAntwerpen.be

Universiteit Antwerpen

Faculteit Letteren en Wijsbegeerte
Departement Taalkunde

Twee gedachten

Syntactische representaties in tweetalige sprekers

Proefschrift voorgelegd tot het behalen van de graad van
doctor in de taalkunde
aan de Universiteit Antwerpen
te verdedigen door

Rianne van Lieburg

Promotoren
dr. Sarah Bernolet
prof. dr. Robert J. Hartsuiker

Antwerpen, 2023

Preface

"Mark my words, those four years will fly by", my supervisor said to me on my first day. And she was right, although surely not in the way she meant, due to all the unlikely and unfortunate events that happened in the past years. During that first summer, everyone was obsessed with the "major relocation", as we switched offices from the first floor to the second floor. This event caused quite a stir, which is a big contrast with last summer, when we suddenly had to leave the second floor and even the entire building after the fire on the Stadscampus. Given the pandemic in between those two moves, one could imagine that my permanent office in D.221 was actually not that permanent.

Nevertheless, after four years there somehow is a thesis, and I have done my research with joy and enthusiasm. This is not least due to the presence and help of many. First, I would not have been able to perform my experiments without my participants, of whom 2,330 are reported in this thesis, and neither have I forgotten the several hundreds of people who participated in one of my pilot experiments. My participants remain anonymous, but I would like to mention the names of some people who contributed in many ways to the creation of my thesis.

I would like to especially express my gratitude to my supervisors Sarah Bernolet and Rob Hartsuiker. This thesis would not have looked the same without the freedom they gave me to shape the project and experiments, and their valuable feedback on my texts. I also really appreciated the pleasant yearly sessions with the members of the doctoral committee, Dominiek Sandra and Gerrit Jan Kootstra.

Many thanks go to my colleague Edwige Sijyeniyo. It was really great to work together on the same project. We did not only have fruitful discussions on the interpretation of our usually confusing results, but we also encouraged and inspired each other, or helped each other to relax or relativize, and the latter has perhaps been most crucial in order to deliver this thesis. I also would like to thank Hanne Surkyn, who never ran out of stories, and therefore was responsible for some welcome entertainment in the office. I enjoyed the lunch breaks with my colleagues Edwige, Hanne, Lisa and Lisan so much that on days working from home I would still come only for lunch.

I also want to say thanks to the colleagues from Ghent University. I have not been able to see them often in real life, but they were always willing to offer their help. The practical aid from Chi Zhang, Merel Muylle, Mieke Slim and Xuemei Chen for using the labs in Ghent or finding participants in China was indispensable. Their work on structural priming has continued to inspire me and I am grateful that I was able to do my PhD in the prime of the research group.

Finally, I wish to express my gratitude to my family and friends. Although they would have been thanked in the preface anyway, they have in fact helped me out often by, to

mention just a few things, debugging the code of my experiments or taking me for a walk to maintain a good work-life balance. Their unconditional support means a lot to me.

You all, and anybody else who has contributed to my thesis in some way, thank you for being part of my process, which has resulted in this thesis.

Rianne van Lieburg

Antwerp, 2023

Voorwoord

"Je zult zien, deze vier jaar vliegen voorbij", zei mijn promotor op mijn eerste werkdag. En zo is het ook gegaan, hoewel ze zeker niet doelde op alle gebeurtenissen die de tijd vooruit zouden doen snellen. In die eerste zomer was de grootste consternatie rondom de "verhuis" – meteen een van de eerste Vlaamse woorden die ik zou leren –, waarbij de kantoren op de eerste verdieping verwisseld werden met die op de tweede verdieping. Die maandenlange opschudding vormt een groot contrast met de afgelopen zomer, waarin we van de een op andere dag ook de tweede verdieping hebben moeten verlaten na de brand op de stadscampus. Door de pandemie die we in de tussentijd doorstaan hebben, is mijn vaste werkplek in D.221 bijna meer fictie dan werkelijkheid geweest.

Toch is er in de afgelopen vier jaar een proefschrift ontstaan, en heb ik mijn onderzoek met veel plezier en enthousiasme kunnen doen. Dat is niet in de laatste plaats dankzij de aanwezigheid en hulp van velen. Allereerst had ik mijn experimenten niet kunnen uitvoeren zonder mijn participanten, van wie 2.330 het tot dit proefschrift geschopt hebben en nog eens enkele honderden belangeloos hebben bijgedragen door hun deelname aan diverse pilotstudies. Mijn participanten zijn anoniem, maar ik noem graag de namen van een aantal mensen die op allerlei manieren veel hebben bijgedragen aan de totstandkoming van mijn proefschrift.

In het bijzonder wil ik graag mijn promotoren Sarah Bernolet en Rob Hartsuiker bedanken. De vrijheid die zij mij gaven om invulling te geven aan de verschillende studies samen met alle uitgebreide feedback op mijn teksten, hebben geleid tot het eindproduct dat er nu ligt. Ook de aangename jaarlijkse sessies met de leden van de begeleidingscommissie, Dominiek Sandra en Gerrit Jan Kootstra, heb ik erg gewaardeerd.

Veel dank gaat uit naar mijn collega Edwige Sijyeniyo. Het was ontzettend fijn om samen op hetzelfde project te werken. We konden niet alleen inhoudelijk overleggen en onze soms verwarrende resultaten bespreken, maar ook elkaar aanmoedigen en inspireren, of juist laten ontspannen en relativeren, en dat laatste was misschien wel het belangrijkste om dit eindproduct te kunnen afleveren. Ook bedank ik graag Hanne Surkyn, die altijd de nodige verhalen meebracht naar kantoor en zo voor soms welkome afleiding zorgde. De lunchpauzes met mijn collega's Edwige, Hanne, Lisa en Lisan waren zo gezellig dat ik daar op mijn thuiswerkdagen nog wel eens voor terugkwam.

Ik wil ook de collega's van de Universiteit Gent bedanken, die ik niet vaak live zag, maar die op afstand altijd voor me klaarstonden. De praktische hulp die Chi Zhang, Merel Muylle, Mieke Slim en Xuemei Chen mij geboden hebben bij onder meer het vinden van participanten of het gebruiken van de labs in Gent, was onmisbaar. Hun werk op het gebied van priming heeft mij steeds geïnspireerd en ik ben dankbaar dat ik in de bloeitijd van de priminggroep mijn onderzoek kon doen.

Tot slot bedank ik graag mijn familie en vrienden. Hoewel zij er niets voor hadden hoeven doen om in dit voorwoord bedankt te worden, hebben ze mij toch steeds met raad en daad bijgestaan in de afgelopen jaren, variërend van het oplossen van programmeerproblemen in mijn experimenten tot de vele wandelingen als tegenhanger voor al het achter de computer zitten. Hun onvoorwaardelijke steun betekent veel voor mij.

Jullie allemaal, en iedereen die ik niet bij naam genoemd heb, maar die wel op de een of andere manier een bijdrage geleverd heeft, bedankt dat jullie onderdeel wilden zijn van mijn traject, dat geleid heeft tot het proefschrift dat nu voor jullie ligt.

Rianne van Lieburg

Antwerpen, 2023

Contents

General introduction

Bilingualism is a specialist field in linguistics, but in reality, bilingualism is everywhere. This is best illustrated by means of the context in which this thesis originated, namely in Antwerp, Belgium's largest city. Although Antwerp is a Dutch-speaking city, many of its 530,000 inhabitants are able to communicate in the three official languages of the country, namely Dutch, French and German. Being a globalized city in the European Union, most people are also very familiar with the English language. Moreover, over half of the Antwerp citizens have an origin outside of Belgium, and many brought a foreign language along with them. For instance, 40 to 50 percent of the children in school report that they speak a different language than Dutch at home.[1] The city of Antwerp is harboring many different kinds of bilinguals and multilinguals, and is illustrative for the ubiquity of bilingualism in general. Indeed, the majority of people in the world are bilingual or even multilingual, being able to speak two or more languages to a certain extent.

As such, studying how multiple languages are represented in the mind is highly topical, and a better understanding of the multilingual mind comes down to a better understanding of humanity. This thesis focuses on one aspect of the multilingual mind, namely on how grammatical structures are processed during production and comprehension at different stages of learning. In five studies, we investigated how representations of grammatical structures are formed during the process of learning a second language, and how the syntactic representations of one language affect the syntactic representations of another language. These studies contribute to the evaluation and further refinement of a developmental account of syntactic representations in second language (L2) learning (Hartsuiker & Bernolet, 2017).

The structural priming paradigm

A way to investigate mental representations of syntactic structures is by means of the structural priming paradigm. Structural priming is "the phenomenon by which processing one utterance facilitates processing of another utterance on the basis of a repeated syntactic structure" (Branigan, 2007, p.1). One of the first observations of this phenomenon was by Levelt and Kelter (1982), who called Dutch shopkeepers by phone, inquiring about the closing time of the store. They found that the way they formulated their sentence affected the way shopkeepers replied to their question (see example 1).

(1) Om hoe laat gaat uw winkel dicht? -Om vijf uur.
"At what time does your shop close? -At five o'clock."

[1]Statistics are from *Stad in Cijfers* (2022), `https://stadincijfers.antwerpen.be/Databank`.

Hoe laat gaat uw winkel dicht? Vijf uur.
"What time does your shop close? -Five o'clock."

Although these findings indicate that shopkeepers echoed the formulation used in the question asked, they leave it unclear whether shopkeepers repeated the structure of the question or merely the preposition *om* "at". Bock (1986) developed an experimental paradigm that arguably does provide evidence for repetition of structure, showing that the phenomenon of structural priming can be exploited in order to study structural representations in production. Structural priming experiments require syntactic alternants, that is, at least two syntactic alternatives that can be used to describe a particular situation. For instance, the picture in Figure GI.1 can be described by either an active sentence *The mouse is eating the cheese* or a passive sentence, *The cheese is being eaten by the mouse*.

Figure GI.1: Example of a stimulus item in a structural priming experiment.

During a typical experiment, participants are exposed to active and passive prime sentences, always followed by a target picture displaying an action and a transitive verb. Participants are asked to describe the picture in one sentence, using the verb provided. This procedure is illustrated in Figure GI.2. Indeed, Bock (1986) found that participants were more likely to produce a passive sentence to describe the target picture after a passive prime than after an active prime. Note that participants also needed to perform a running memory task during the priming experiment in order to disguise the goal of the experiment. Before the experiments, they had been given a study list of pictures and sentences. During the experiment, participants were asked to evaluate whether the prime sentence or the target picture had occurred in the study list.

Structural priming effects have been demonstrated with several syntactic alternants, in multiple languages and with different demographic groups, including children, people with aphasia, and L2 speakers (see Pickering & Ferreira, 2008 for an overview and Mahowald et al., 2016 for a meta-analysis). Crucially, structural priming occurs between lexically unrelated prime and target sentences, implying that the phenomenon reflects the cognitive processing of abstract syntactic representations. For example, the prime sentence *The elephant is treated by the veterinarian* primes the target sentence *The cheese is being eaten by the mouse*, meaning that a passive prime sentence leads to an increase in the proportion of passive rather than active target sentences, even if there is no lexical overlap between the prime and target sentence. The presence of structural priming in the

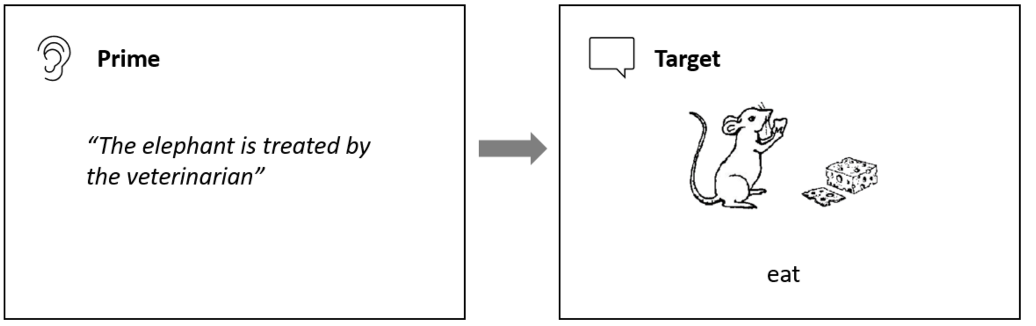


Figure GI.2: The structural priming paradigm. As the prime sentence contains the passive structure, the participant is more likely to produce a passive sentence when describing the target picture than if the prime sentence had been an active sentence (*The veterinarian treats the elephant*).

absence of lexical overlap between prime and target is called abstract structural priming. Nevertheless, structural priming is stronger when the verb is repeated between the prime and target sentence, the so-called lexical boost effect (Pickering & Branigan, 1998). To illustrate, stronger priming will occur for the target sentence *The cheese is being eaten by the mouse* if the prime sentence is *The elephant is eaten by the veterinarian* instead of *The elephant is treated by the veterinarian*. This suggests that there is a link between lexical items such as verbs and abstract syntactic structures, and that this link may be primed in addition to the abstract priming of the syntactic structure.

The structural priming effect is often measured in terms of the difference in the relative proportions of the syntactic alternants between the different prime conditions. For instance, Bock (1986) measured the proportion of active/passive responses after an active prime compared to the proportion of active/passive responses after a passive prime. However, this approach does not allow us to separate between priming of the active structure and priming of the passive structure. It may theoretically be possible that the passive prime has a larger modulating effect on the proportion of active/passive responses in the target sentences than the active prime. In the footsteps of Bernolet et al. (2009), we therefore measured the structural priming effect in our experiments in comparison to a baseline prime. Baseline prime sentences are unrelated to the syntactic alternants under investigation and may, for instance, be intransitive sentences (e.g., *The elephant is walking*). For example, priming of the passive structure is measured by subtracting the proportion of passive responses after a baseline prime from the proportion of passive responses after a passive prime. Importantly, including a baseline allows us to assess the magnitude of priming of the syntactic alternants separately, and to observe the polarity of the priming effects. Previous experiments including a baseline have for example shown that structural priming is modulated by the relative frequency of the syntactic alternants. The less frequent alternative is often primed stronger than the more frequent alternative, a phenomenon referred to as the inverse preference effect (V. S. Ferreira & Bock, 2006).

Structural priming in comprehension

The structural priming paradigm has originally been developed for production experiments, and the majority of the work on structural priming as a way to investigate mental representations of syntactic structures is on sentence production. However, structural priming can also be found in comprehension. A range of methods have been used to test structural priming in comprehension, including self-paced reading (e.g., Fine & Jaeger, 2016; Traxler & Tooley, 2008), fMRI (e.g., Segaert et al., 2012; Weber & Indefrey, 2009), EEG (e.g., Ledoux et al., 2007; Tooley et al., 2009) and eye-tracking (e.g., Arai et al., 2007; Thothathiri & Snedeker, 2008a).

For example, Arai et al. (2007) measured eye movements in a comprehension experiment using a visual world paradigm. First, participants were asked to read a prime sentence, which was a direct object (DO) sentence (e.g., *The assassin will send the dictator the parcel*) or a prepositional object (PO) sentence (e.g., *The assassin will send the parcel to the dictator*). Then they listened to either a DO (*The pirate will send the princess the necklace*) or a PO (*The pirate will send the necklace to the princess*) target sentence. Meanwhile, they looked at a scene which displayed the three entities mentioned in the sentence (i.e., a pirate, a princess and a necklace). Participants looked longer and earlier to the recipient (i.e., the princess) after a DO prime sentence than after a PO prime sentence. This suggests that they anticipated that the target sentence would again be a DO sentence after a DO prime sentence, while they expected another PO target sentence after a PO prime.

The procedure for testing structural priming in comprehension is essentially similar to the way structural priming is tested in production experiments: participants are exposed to prime-target pairs of two different syntactic structures. Like in production experiments, these syntactic structures may be syntactic alternants, meaning that the two structures express the same meaning (such as the active and the passive structure). Alternatively, temporarily ambiguous structures may be used. In this case, the structures have the same beginning before deviating, and they have different meanings. For instance, Traxler and Tooley (2008) found that it was easier for participants to process a reduced relative clause (e.g., *The defendant examined by the lawyer was unreliable*) after another reduced relative clause (e.g., *The engineer examined by the doctor had a large mole*) than after a main clause (e.g., *The engineer examined the license in the doctor's office*).

Given the different methodologies used in comprehension, the way structural priming effects are measured is diverse. Self-paced reading experiments measure the time a participant needs to read each word in the sentence. Here, priming occurs if participants are faster to read a primed sentence than an unprimed sentence. With fMRI and EEG, activation of certain brain regions is measured over the course of exposure to a trial. Priming in fMRI experiments is detected by means of repetition suppression effects, meaning that there is less neural activation for the repeated structure than for the unprimed structure. This shows that priming leads to facilitated processing of mental representations of structures. In EEG experiments, priming can be detected as a reduced P600 effect. The P600 effect is a positive peak in the brain activity around about 600 milliseconds after the start of the trial, and its magnitude is correlated with syntactic complexity: sentences that are more difficult to process, elicit a larger P600 effect than sentences that are easier to process. The P600 effect for a particular sentence is smaller in primed conditions than in unprimed conditions, which implies that structural priming facilitates processing. Finally, eye-tracking studies using a visual world paradigm may detect when participants start looking at an object and how long this gaze persists. Participants may look earlier and longer to a particular object in the primed condition than in the unprimed condition. Alternatively, eye movements may be tracked in a reading experiment, and fixation times on words are measured. Such experiments often use a disambiguating area, and participants may look shorter to disambiguating words in primed conditions than in unprimed conditions, reflecting faster processing of primed sentences.

Due to the way structural priming in comprehension is measured, a baseline prime condition is not required in order to assess priming of the syntactic alternants separately. For instance, in Traxler et al. (2014), participants showed priming for reduced relative clauses (e.g., *The defendant examined by the lawyer was unreliable*), but not for main clauses (*The defendant examined the glove but was unreliable*). Participants looked shorter at reduced relative clause targets after a relative clause prime (*The engineer examined by the doctor had a large mole*) than in an unprimed condition. The looking times for target sentences containing main clauses, on the other hand, was not significantly different after a main clause prime (*The engineer examined the license in the doctor's office*) than after a relative clause prime. However, in an experiment which primes both syntactic alternants, it may still be useful to insert a baseline prime condition, as long-lasting priming may affect the processing of the structures under investigation in the prime sentences as well. In other words, the instances of the syntactic alternants preceding the prime sentence may also have an effect on the processing of the target sentence, and without a baseline it may be difficult to know to what extent a measured priming effect should really be attributed to the prime sentence. A baseline prime condition therefore can serve to measure the magnitude of priming for the structures more accurately.

As in production, the lexical boost effect has been observed in comprehension as well (e.g., Traxler et al., 2014). There is also some evidence for the inverse preference effect (cf. X. Chen et al., 2022). It must be noted, however, that the results in comprehension are much less consistent than in production. This may be due to methodological differences between modalities or to the way sentences are processed in comprehension. Still, because of similarities in structural priming effects between production and comprehension, it has been argued that structural priming relies on the same cognitive mechanisms in the different modalities.

Mechanisms behind structural priming

Structural priming effects in production and in comprehension seem to be the product of cognitive processes during speech processing. Two different mechanisms have been proposed to account for structural priming effects, namely residual activation and implicit learning. Hybrid models of structural priming integrate both mechanisms. We will first discuss the residual activation account and the implicit learning account separately. Then we will illustrate the two-stage competition model, which is an example of a hybrid model.

Residual activation account

The residual activation account (Pickering & Branigan, 1998) is a lexicalist account, which fits in Levelt's model of speech production (Roelofs, 1992, 1997; Levelt et al., 1999). According to this model, speech processing involves cognitive processing at three different stages, namely conceptual preparation, lexical selection and form encoding (cf. Indefrey & Levelt, 2004). During these stages, stored information needs to be accessed in the neural network. Roelofs (1992, 1997) assumes that the relevant information is stored in three strata, each one corresponding to one of the stages of cognitive processing: a conceptual stratum, a lemma stratum (also called lexical stratum) and a form stratum. The strata consist of nodes, which are connected to each other. Each node represents a piece of linguistic information, such as a lexical concept, a lemma, or a word form. The nodes also have a level of activation, which depends on the input from other nodes.

During speech production, activation spreads through this network of nodes. At the stage of conceptual preparation, lexical concepts are activated, which in turn leads to the activation of their corresponding lemma nodes. The lemma reaching the highest level of activation is selected during the lexical selection stage. The selected lemma is then passed on to the form encoding stage, where the morphophonological form of the word is determined.

The residual activation model (Pickering & Branigan, 1998) elaborates upon processing at the lemma stratum, the stage during which structural priming effects presumably arise. Lexical items, such as verbs, are stored in lemma nodes. These lemma nodes are linked to combinatorial nodes, which contain syntactic information. During the processing of a sentence, activation spreads from the lemma nodes to the combinatorial node of the syntactic structure being processed. For instance, the verb *give* is connected to a double object node and a prepositional object node. Processing the phrase *give the dog a bone* activates the double object node, whereas the phrase *give a bone to the dog* activates the prepositional object node.

Importantly, the residual activation account assumes that the activation of the nodes and the link between nodes only decays gradually after the processing of a sentence: there remains some residual activation for some time, and this residual activation facilitates re-activation of the nodes. When a speaker produces the double object phrase *give the dog a bone*, the double object node and the link between the lemma *give* and the combinatorial node of the double object node will be activated. Because of residual activation in the combinatorial node, the speaker is more likely to produce a double object phrase again, which causes structural priming. This is also the case when the speaker uses a different ditransitive verb, such as *show*. Since *show* is connected to the same double object node as the verb *give*, residual activation of the double object node due to producing *give the dog a bone* also leads to an increased use of the double object node with the verb *show*. The lexical boost effect is the consequence of the residual activation in the lemma node and the link between the lemma node and the combinatorial node. If the speaker re-uses the verb *give*, the use of the double object construction is not only facilitated by residual activation of the double object node, but also by residual activation of the link between *give* and the double object node. As a result, structural priming is stronger when the verb is repeated between prime and target than when a different verb is used between prime and target. Figure GI.3 illustrates the process of residual activation.

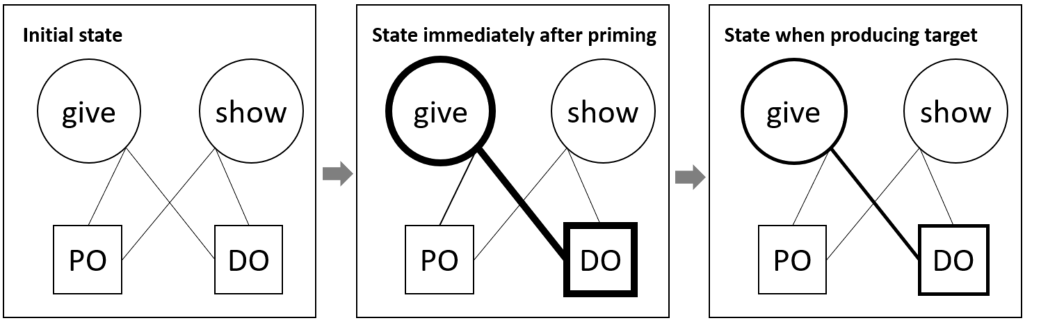


Figure GI.3: The residual activation account (Pickering & Branigan, 1998). The first panel shows the initial state of the network. The lemma nodes of the verbs *give* and *show* are connected to the combinatorial nodes of the prepositional object dative and the double object dative. The second panel shows the state directly after processing the prime sentence *The man gives the dog a bone*. There is activation of the lemma node of *give*, the combinatorial node of the double object dative and the link between these nodes. The third panel shows the state of the network during production of the target, when there is still residual activation available.

Implicit learning account

The implicit learning account (Chang et al., 2006) is an error-based learning account, which is based on a connectionist model of speech production. It assumes that, during language processing, listeners predict what they are going to hear. When the actual incoming speech does not match the predicted input, listeners adapt their system, and learning occurs.

These assumptions are implemented computationally in a dual pathway neural network, consisting of a meaning system and a sequencing system. This is illustrated in Figure GI.4. The meaning system contains the message and maps concepts to words. The sequencing system is a recurrent neural network, which consists of hidden layers with units of different relative weights. The system applies algorithms that generate an output from an incoming input. The sequencing system maps a word to the next word in the sequence, thus sequencing the words in a syntactically appropriate way. During incremental sentence processing, the sequencing system predicts word by word the next word in the sentence. The predicted word is compared to the actual word, and when there is a mismatch, the sequencing system is adapted by means of backpropagation. Backpropagation means that the relative weights of units in the hidden layers are permanently adapted on the basis of prediction error, which affects the predicted output in the future. In this way, prediction error leads to long-term learning.

Chang et al. (2000) argue that structural priming reflects the learning effect of the prediction error that may be caused by processing the prime sentence. During production, the sequence of predicted words becomes the produced sentence to formulate the intended message (Chang et al., 2006). When through exposure to a prime sentence, a double object phrase such as *give the dog a bone* is processed, the relative weight of the link between the message and the double object construction is strengthened. It therefore becomes more likely to predict the double object dative, and thus, to produce the double object dative in

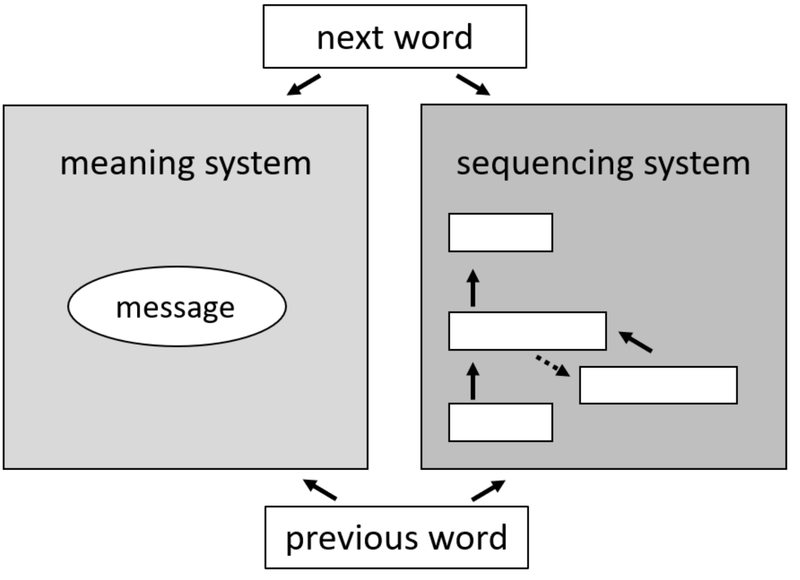


Figure GI.4: The implicit learning account (Chang et al., 2006).

the subsequent target item.

The implicit learning account explains why structural priming effects are long-lasting and can be found even after 10 intervening trials (Bock & Griffin, 2000). In addition, implicit learning accounts for the inverse preference effect, as there is larger prediction error for less frequent structures than for more frequent structures. As a result, processing a less frequent structure leads to a larger adaptation in the relative weights of the sequencing system, which implies a stronger learning effect than for more frequent structures.

A hybrid model for production: Two-stage competition model

Long-lasting priming and the inverse preference effect cannot be explained from the residual activation model (Pickering & Branigan, 1998). On the other hand, the lexical boost effect does not follow immediately from the implicit learning model, as the model does not assume a direct link between lexical items and syntactic structures. Hartsuiker and Pickering (2008) show that while abstract structural priming is long-lasting, the lexical boost effect decays shortly after exposure. Several authors therefore attribute the lexical boost effect to residual activation and/or explicit memory (e.g., Bernolet et al., 2016). In order to integrate a residual activation and an implicit learning mechanism, hybrid models have been proposed (e.g., Reitter et al., 2011; Segaert et al., 2011).

The two-stage competition model (Segaert et al., 2011, 2014, 2016) is such a hybrid model. It assumes that the combinatorial nodes of syntactic structures have a certain resting level of activation, which represents the baseline frequency of structures. For instance, the active structure has a higher base-level frequency than passive structures, as active sentences are much more frequent than passive sentences. The height of the base-level frequency is established through implicit learning: processing a structure leads to the increase of the base-level frequency.

Sentence production proceeds in two stages: the selection stage and the planning stage. When describing a transitive event, both the active and the passive structure are activated. During the selection stage, there is competition between the combinatorial nodes of the activated structures. The structure of which the combinatorial node reaches the selection threshold first (that is, the activation level which leads to firing of the node), is the structure that gets selected and prepared for production during the planning stage.

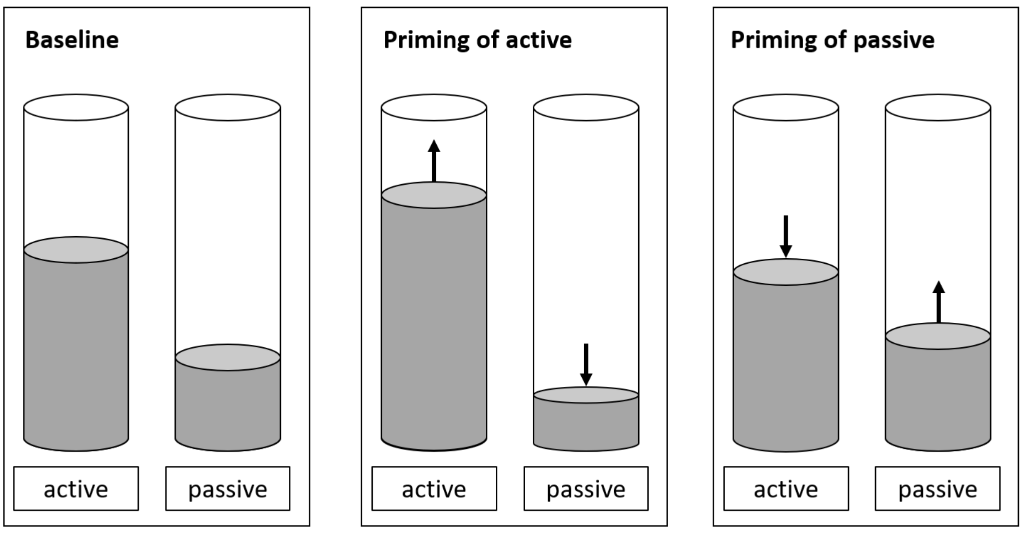


Figure GI.5: The two-stage competition model (Segaert et al., 2011).

The two-stage competition model cannot only account for the lexical boost effect (through residual activation) and the inverse frequency effect (through implicit learning), but also explains why participants are faster to start producing their target sentence when primed (Segaert et al., 2011). Segaert et al. found that this effect is stronger for the more frequent active structure than for the less frequent passive structure. During priming, the activation level of the primed structure increases temporarily. When the active structure is primed, which already has a higher base-level activation than the passive structure, the gap between the activation levels of the active and the passive structure increases, which decreases the competition and consequently, the selection time for the active structure. Instead, when the passive structure is primed, the activation level of the passive structure increases, decreasing the gap between the activation levels of the active and the passive structure. Consequently, there is more competition, increasing the time needed for sentence selection. The planning stage proceeds faster after priming for both the active and the passive structure, accumulating the facilitation of the active structure and compensating for the inhibition of the passive structure during the selection stage. As a result, Segaert et al. were only able to observe structural priming of the onset latencies of active sentences, but did not find significant priming for the passive structure. The structural priming effect for the active structures is an additive effect of a faster selection time and faster planning for active sentences after an active prime than after a baseline prime. The null effect for the passive structures is due to a slower selection time and faster planning for passive sentences after a passive prime than in an unprimed condition, and these effects level each other out.

A hybrid model for comprehension

Although it is believed that structural priming in production and comprehension rely on the same mechanisms of residual activation and implicit learning, and more generally, that production and comprehension processing recruit overlapping neural circuits (cf. Pickering & Garrod, 2013), there are modality-specific components affecting structural priming as well. For instance, at least the planning stage of the two-stage competition model may be specific to production. Presumably, there are also aspects specific to priming in comprehension.

First, structural priming in comprehension seems to be modulated by "message predictability" (Ziegler & Snedeker, 2019). In comprehension experiments, target structures which are more predictable may be primed less than structures that are less predictable in the provided context. For instance, in Arai et al. (2007), the visual scene only included the three entities (a princess, a pirate and a necklace) that were mentioned in the target sentence *The pirate will send the necklace to the princess*. Therefore participants could accurately predict the continuation of the sentence soon after the onset of the word *necklace*. By contrast, the visual display used in the otherwise very similar experiment by Thothathiri and Snedeker (2008a) also included entities that were not mentioned in the target sentence. In addition, one of the extra entities had a similar phonological onset as the critical word in the target sentence. To illustrate, the target sentence *Now, he's gonna feed the baby the apple* was accompanied by a visual display showing a baby, an apple, a bagel and a girl. The critical word of the target sentence is *baby*, as the direct object dative *feed the baby the apple* and the prepositional object dative *feed the apple to the baby* start to deviate at this point. As *baby* and *bagel* have the same first syllable, participants are only certain of *baby* as the correct picture and thus of the direct object dative structure after the entire word is pronounced. This set-up makes the target structure much less predictable than the visual scene used in Arai et al. Indeed, Arai et al. only found structural priming of ditransitive structures when there was verb overlap between prime and target, while Thothathiri and Snedeker were able to observe abstract structural priming as well. F. Ferreira et al. (2002) proposed a "good enough" approach to sentence comprehension, meaning that a listener may not need a full syntactic analysis in order to process a sentence. Therefore the residual activation of primed structures may be weaker in comprehension than in production.

Second, rather than choosing and producing one of the syntactic alternants, the listener is confronted with a particular structure, which may or may not be in line with their expectations. Listeners need to decide which structural analysis is needed to comprehend the processed structure. This is called syntactic ambiguity resolution. Note that the implicit learning account is in fact based on this aspect of comprehension. During comprehension, listeners predict the upcoming structures. When the prediction does not match the processed structure, the relative weights of syntactic structures are updated. Production only reflects the relative weights that have been adjusted during comprehension.

There is no model of structural priming that is specifically developed to account for structural priming effects in comprehension. However, there are two competing accounts on how listeners deal with the process of syntactic ambiguity resolution. Garden-path models for sentence comprehension (e.g., Frazier, 1979; Pickering et al., 2000) assume that during processing, the listener adopts a particular syntactic analysis of the sentence. When the continuing incoming input cannot be fit in this syntactic analysis, reanalysis needs to take place. To illustrate, the verb *realize* can be followed by an object (*The athlete realized her potential*) or by a sentential complement *The athlete realized her potential might make her a world-class sprinter*). Listeners may initially analyze *her potential* as an object, as this structure may be more plausible to follow after *realize*. When the verb *might* follows in the incoming input, listeners need to adapt their initial syntactic analysis in order to parse the phrase as a sentential complement to *realize* (Pickering et al.). Constraint-based models (e.g., MacDonald et al., 1994; Spivey & Tanenhaus, 1998), on the other hand, assume that there is competition between different structures. So, listeners consider multiple syntactic analyses in parallel, and the probability of these analyses is weighted. For example, when listeners hear the phrase *The athlete realized her potential*, they may initially consider both the object analysis and the sentential complement analysis. When the sentence continues with the verb *might*, the sentential complement analysis becomes more probable and ultimately becomes the selected syntactic interpretation.

The garden-path and constraint-based accounts may both fit into a model of structural priming in comprehension. When primed, residual activation of structures may affect which structure is the structure that is initially assumed during syntactic analysis (according to garden path models), or may have an influence on the competition between syntactic analyses of incoming structures (according to constrained-based models). Within the garden-path model, the prediction error that the implicit learning account assumes may be responsible for inducing syntactic reanalysis. In constraint-based models, the relative weights of different syntactic analyses may be determined and updated by the implicit learning mechanism.

If the aim is to have a model of structural priming that applies to both production and comprehension, especially the selection stage of the two-stage competition model by Segaert et al. (2011) may be compatible with constraint-based models of comprehension. The two-stage competition model also assumes that multiple structures compete with each other before one structure is selected for production. It is more difficult to integrate the two-stage competition model with a garden-path model, as the latter does not assume competition between structures. However, there may still be aspects of structural priming that are specific to one of the modalities. While the planning stage of the two-stage competition model may be specific to production, "good enough" processing and syntactic ambiguity resolution may only apply to comprehension. A joint hybrid model of structural priming should therefore not only integrate the mechanisms that are shared between modalities, namely residual activation and implicit learning, but also account for the modality-specific aspects of both production and comprehension. Importantly, due to those modality-specific aspects such as the planning stage in production, and "good enough" processing and syntactic ambiguity resolution in comprehension, under certain circumstances, differential priming effects can be expected between the modalities. Research on such differential effects is still in its infancy, even for structural priming in the L1.

Structural priming in bilingual speakers

As residual activation concerns residual activation of stored lemmas and combinatorial nodes, and implicit learning exploits the prediction of familiar words and structures, the structural priming paradigm can be used to investigate the knowledge of syntactic structures in bilingual speakers. Structural priming effects may thus inform us about the nature of mental representations of syntactic structures in L2 speakers.

First, an important finding is that structural priming effects do not only occur within languages, but also between languages. In a production experiment with Spanish-English bilinguals, Hartsuiker et al. (2004) used Spanish active and passive prime sentences, and asked participants to describe the target pictures in English. Participants produced more passive structures in English after a passive prime in Spanish than after an active prime. Between-language priming has also been found in comprehension. For instance, Kidd et al. (2015) tested the interpretation of German relative clauses, such as *die Frau, die das Mädchen küsst* [lit. the woman that the girl kisses]). In German, relative clauses are ambiguous between subject relative clauses (*the woman that kisses the girl*) and object relative clauses (*the woman that the girl kisses*). Bilingual speakers of English and German were more likely to interpret the German relative clause as an object relative clause after an English object relative clause prime than after an English subject relative clause prime.

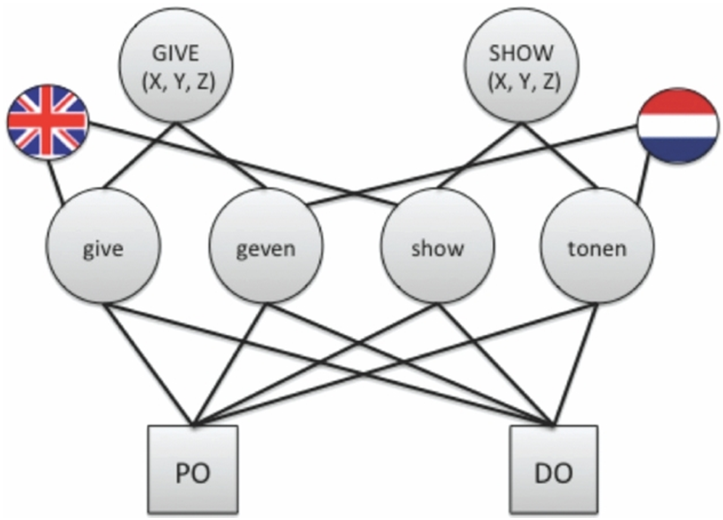


Figure GI.6: The bilingual lexical network (from Hartsuiker & Bernolet, 2017). The lemma nodes of the English ditransitive verbs *give* and *show* and Dutch *geven* and *tonen* are all connected to a single prepositional object node and direct object node.

Between-language priming effects may be explained by the shared syntax account (Hartsuiker et al., 2004), which is rooted within the framework of the residual activation account. The lemma nodes of the lexical items of multiple languages are connected to the same combinatorial nodes representing overlapping grammatical structures. This is illustrated in Figure GI.6. A prime sentence in one language activates the shared combinatorial node of its structure. Due to residual activation, participants are more likely to reproduce that structure when describing a target picture in the other language.

An alternative to the shared syntax account is the connected syntax account (van Gompel & Arai, 2018), which assumes that similar grammatical structures of different languages may at least sometimes have separate combinatorial nodes that are connected to each other. This may particularly be the case for similar structures that differ between languages in terms of the surface word order, such as the by-phrase final passive in English *The cheese is being eaten by the mouse* and the by-phrase medial passive in German *Der Käse wird von der Maus gefressen*. The shared syntax account can explain findings showing comparable between-language priming and within-language priming effects (Kantola & van Gompel, 2011), whereas the connected syntax account fits better with studies finding stronger within-language priming than between-language priming (Cai et al., 2011). Recently, Khoe et al. (2021) performed several between-language structural priming experiments using computational models with simulated bilingual participants, showing that between-language priming can also be explained by an implicit learning mechanism rather than residual activation. Their models showed stronger within-language priming than between-language priming.

L2 priming seems to be modulated by proficiency, at least in production. Bernolet et al. (2013) found that L2 speakers of English with Dutch as their L1 who were relatively proficient showed stronger abstract priming for *s*-genitives (*the girl's shirt*) and *of*-genitives (*the shirt of the girl*) than less proficient speakers. This was the case both in a between-language experiment and in an experiment within the L2. By contrast, the less proficient participants were primed stronger within the L2 when the head noun (*shirt*) was repeated between prime and target. The results suggest that L2 learners might start with item-specific and language-specific syntactic representations in the L2. These representations become abstract and shared between the L1 and the L2 over time.

So far, similar proficiency effects have not been found in structural priming in comprehension. Wei et al. (2018) investigated priming of reduced relative clauses in L2 speakers of English with Chinese as their L1, but did not find evidence for a correlation between the L2 proficiency and the magnitude of the priming effects. The reason may be that the participant group was too homogeneous with regard to the L2 speakers' proficiency in English.

Hwang et al. (2018) suggested that the shared representations between the L1 and the L2 become stronger as proficiency increases on the basis of more interference of the L1 upon increasing proficiency in a non-priming study. Hwang et al. showed pictures of a causative event, together with either an active (e.g., *Jen fixed a computer*) or causative (e.g., *Jen had her computer fixed*) sentence in English to participants with Korean as their L1. In Korean, in contrast to English, active structures are used for both causative and transitive events. Participants decided whether the sentence described the event accurately. Paradoxically, participants who were more proficient in English were more likely to (incorrectly) accept an English active sentence for the causative event than the less proficient participants. This suggests that cross-linguistic influence from Korean to English increases upon increasing proficiency, implying that syntactic structures become shared between the L1 and the L2 over time in comprehension as well.

Developmental account of L2 syntax

The notion that L2 learners might start with item-specific and language-specific syntactic representations has been captured in the developmental account of L2 syntax (Bernolet & Hartsuiker, 2018; Hartsuiker & Bernolet, 2017), which aims to explain how syntactic representations develop during L2 acquisition in five stages (see Figure GI.7).

Stage 1. The L2 learner has lexical representations of words, without syntactic information attached to it. Sentences are formulated by means of the L1 grammar, which may lead to transfer errors. To illustrate, an L2 speaker of English with Dutch as their L1 may translate *de pop van de jongen* "the doll of the boy" into *the doll from the boy* rather than *the doll of the boy*, as the preposition *van* is the equivalent of both *of* and *from* in English.

Stage 2. The learner forms item-specific and language-specific representations, without generalizing the structure to other lexical items. For instance, they will learn that *the doll of the boy* is the correct equivalent of *de pop van de jongen*, but will not generalize this structure in order to produce *the ball of the girl*.

Stage 3. The L2 speaker adds more item-specific, language-specific representations, meaning that they will be able to use a particular lexical item in different constructions. For example, they can now alternate between *the doll of the boy* and *the boy's doll.*

Stage 4. Syntactic constructions are generalized across lexical items. From now on, the L2 learner is able to use the syntactic structures productively. So, they will be able to produce *the doll of the boy* and *the ball of the girl*.

Stage 5. Structures become shared across languages, if they are similar enough. The combinatorial node of syntactic structures is connected to the lemma nodes of both the L1 and the L2 items in the lexicon. For instance, the combinatorial nodes of the *of*-genitive and the *s*-genitive are connected to the L1 Dutch noun *jongen* "boy" and the L2 English noun *boy*.

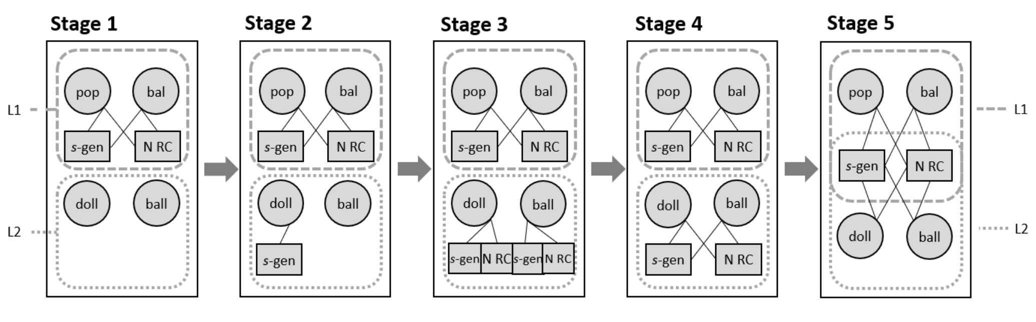


Figure GI.7: The five stages of the developmental account (Hartsuiker & Bernolet, 2017).

The developmental account predicts different structural priming effects at the different stages. At stage 1, some structural priming effects may already occur, but only in the presence of lexical overlap, and if the speaker can rely on explicit memory (e.g., there are no intervening items between prime and target). The speaker may then apply a copy-edit strategy to the prime sentence to produce the target. During stage 2 and 3, structural priming effects still only occur if there is lexical overlap between prime and target, but priming may now be induced by residual activation of the item-specific representations instead of on the basis of explicit memory only. Due to the generalization across lexical items, abstract structural priming effects within the L2 may start to occur at stage 4. Between-language priming of similar structures may only be found at stage 5, when structures have become shared.

Research questions

The aim of this thesis is to evaluate and elaborate upon the developmental account of L2 syntax (Hartsuiker & Bernolet, 2017) in two different ways. First, the developmental model was largely based on late L2 learners at an intermediate to advanced level. That means that the L2 syntax of the more extreme sides of the continuum of the language learners, namely beginning learners and native-like L2 speakers are not well investigated. Second, the developmental model is based on structural priming of L2 structures in production, and it is not clear to what extent the learning trajectory also applies to comprehension.

Muylle (2020; also see Muylle et al., 2021a, 2021b) aimed to test the predictions of the developmental account of L2 syntax in beginning L2 learners. However, as they used an artificial language paradigm, learners were very fast to reach high proficiency in the miniature language. An additional issue they encountered is that the developmental account of L2 syntax is an extension to the shared-syntax model of Hartsuiker et al. (2004), which in turn is embedded in the framework of the residual activation account. Although Hartsuiker and Bernolet (2017) do suggest that there is an implicit learning mechanism which leads to the formation of syntactic representations, the model does not specify how error-based implicit learning may interact with structural priming effects in particular ways at different stages. For instance, prediction error may be larger in less proficient speakers than in more proficient speakers, suggesting that structural priming effects may decrease rather than increase over time. Indeed, some of the findings of Muylle were better explained by means of an implicit learning mechanism rather than by residual activation. For example, Muylle (2020) found inverse preference effects for the priming of ditransitive structures from the artificial language to L1 Dutch, as there was stronger priming for the direct object dative (which is less frequent in Dutch) than for the prepositional object dative (the more frequent structure in Dutch). As such, there is a need to better understand structural priming effects at different stages of L2 learning, in order to revise the developmental account of L2 syntax and turn it into a hybrid model.

In addition, the developmental account of L2 syntax was only based on previous results from production experiments. As mental representations of syntactic structures may be activated during both production and comprehension, it is also important to investigate L2 structural priming in comprehension, and also more specifically, in relation to structural priming effects in production, in order to fully understand the development of these mental representations. Importantly, the developmental account makes predictions with regard to abstract structural priming versus lexically-modulated priming. However, it is not well understood how abstract structural priming works in comprehension in the L1, and whether the magnitude of abstract structural priming in comprehension is similar to the magnitude of abstract structural priming in production. In addition, some of the observed proficiency effects may be specific to production, as less experienced learners might avoid more complex syntactic structures despite having a mental representation for those structures. Nevertheless, the magnitude of structural priming effects in production, as being modulated by proficiency, has played an important role in the architecture of the developmental account. Therefore structural priming effects in comprehension in comparison to structural priming in production should be investigated in the L1 as well as the L2 in order to extend the developmental account of L2 syntax to an integrated account for production and comprehension.

Overview of the dissertation

This thesis consists of two different parts, corresponding to the two broader research questions. The first part is on structural priming in production, in which we investigate L2 structural representations in groups representing the two extremes of the continuum of language learners. The second part is on structural priming in comprehension as compared to structural priming in production. In this part, we explore the relationship between priming in comprehension and production in both L1 and L2 speakers.

The Production part consists of two chapters. In Chapter 1, we tested beginning L2 learners of Dutch in a longitudinal and a cross-sectional study. We studied the transition from item-specific, lexically-dependent representations to abstract syntactic representations for Dutch active and passive sentences. The results show that abstract priming of the more frequent and easier active structure takes place before priming of the more complex and less frequent passive structure, although abstract representations of the passive are formed quite rapidly after exposure. In addition, the results suggest that the mechanisms of residual activation and implicit learning are both present already in beginning L2 learners, and should therefore both be part of a developmental model of L2 syntax.

In Chapter 2, we investigated the syntactic representations for Dutch passive sentences in highly proficient speakers of Dutch, who are proficient heritage speakers of another language. Our aim was to examine the influence of the production preferences of the heritage language on the production preferences in Dutch. We found that in comparison to a Dutch control group, Arabic/Berber-Dutch speakers were more likely to produce agentless passives in Dutch, which is the common passive structure in Arabic and Berber. By contrast, Turkish-Dutch speakers produced fewer by-phrase-medial passives in Dutch than the control group did, although the by-phrase-medial passive is the most frequent passive structure in Turkish. We interpret this inhibition effect as an indication that highly proficient bilinguals may have connected syntactic representations rather than shared representations.

The Comprehension part contains three chapters. In Chapter 3, we studied how priming effects in production relate to priming effects in comprehension, testing ditransitive sentences in L1 speakers of English. More specifically, we investigated to what extent modality-specific aspects affect the magnitude of structural priming by comparing priming from comprehension to production to priming from comprehension to comprehension. We detected priming effects in production, but not in comprehension. Although the absence of priming in comprehension may partly be attributed to the experimental design, priming in production seems to be enhanced by self-priming, suggesting that modality-specific aspects may lead to larger priming in production than in comprehension, despite a common mechanism causing priming.

In Chapter 4, we compared structural priming of ditransitives between production and comprehension in L1 speakers of English. By simultaneously collecting choice data and reaction time data in both modalities, we investigated to what extent any differences in priming effects should be attributed to modality or measurement type. Using the maze task paradigm, we were able to detect abstract structural priming in comprehension both in choice data and in reaction time data. The results show that structural priming does not only involve a facilitatory effect, but may also has an inhibitory effect on the processing of alternative structures. In particular, participants were slower to process a direct-object dative target sentence after a prepositional-object dative prime sentence than after a baseline prime. This suggests that there is competition between structural alternatives, and the relative frequency of structures rather than modality determines the polarity and the magnitude of priming.

Finally, in Chapter 5, we examined structural priming of ditransitives in comprehension in two different populations of intermediate to advanced L2 speakers of English, namely Chinese-English and Dutch-English bilinguals. We asked whether any interaction of L2 proficiency with structural priming effects in comprehension and in the onset latencies of production parallels the known interaction between proficiency and structural priming in production choice data, namely decreasing lexically-dependent priming and increasing abstract structural priming upon increasing proficiency. Indeed, proficiency modulated priming in the onset latencies in production and in comprehension as well. Nevertheless, the observed polarity of the correlation between priming and proficiency was not consistent. The proficiency effects did not always interact with lexical overlap, and in some comparisons abstract priming was stronger for less proficient participants than for more proficient participants. The complex data pattern suggests that there may be an interplay between explicit memory, residual activation, implicit learning and the relative frequency of structures, which determines the magnitude of priming at a particular stage of L2 development.

We will end the thesis with a general discussion in which we present a synthesis of our findings and consider their implications for the developmental account of L2 syntax. We will discuss how a hybrid model of L2 syntax should include the mechanisms that are also present in L1 priming, such as residual activation, implicit learning, competition between structures and modality-specific aspects like self-priming and syntactic ambiguity resolution. On the basis of our conclusions, we will provide some suggestions for future research towards a better understanding of bilingual processing and the development of structural representations in bilingual speakers.

Part I

Production

Chapter 1

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

The developmental account of 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: late L2 syntax acquisition, structural priming, abstract structural representations

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

This chapter is co-authored by Edwige Sijyeniyo.

1.1 Introduction

When one moves or migrates to a new country, this often entails that one is 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 is reading 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 less frequent. 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 (Mahowald et al., 2016; Pickering & Ferreira, 2008). 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 is baking 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 the syntactic structure (Pickering & Ferreira). In our case, if beginning learners 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 bilinguals 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, and the link between these two representations. Hartsuiker et al. 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 & Pickering, 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 depends on the difference between predictions and expectations of syntactic rules 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 (V. S. Ferreira & Bock, 2006). Importantly, recently, Khoe et al. (2021) have implemented and tested a bilingual version of the implicit learning account, showing that error-based learning indeed plays an important role in L2 learning.

In bilingual speakers, not only the relative frequency of a structure but also proficiency seems to play an important role in the magnitude of structural priming. Bernolet et al. (2013) showed that more proficient L2 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 less proficient speakers. In contrast, 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 over time. 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 within the L2 lexical overlap between prime and target 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 syntactic structures become shared over time. According to their account, syntactic development takes place in five stages (see Figure 1.1).

Stage 1. 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* "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. 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: 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. 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.

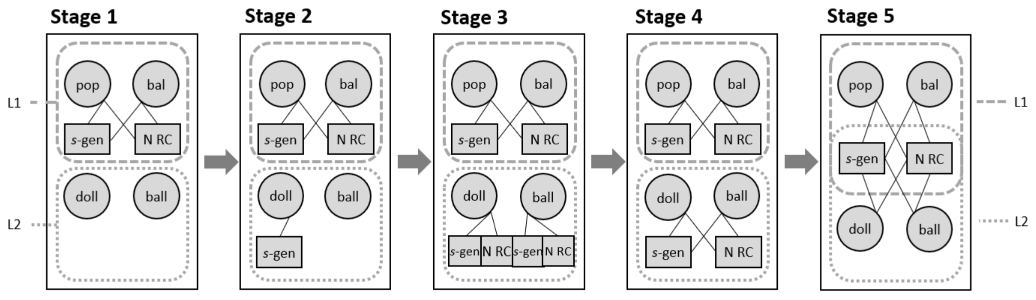


Figure 1.1: Developmental account for L2 syntax acquisition as proposed by Hartsuiker and Bernolet (2017). The upper part represents the lexical-syntactic network in the L1, and the lower part represents the development of the lexical-syntactic network in the L2.

The developmental account of Hartsuiker and Bernolet (2017; see also Bernolet & Hartsuiker, 2018) predicts different structural priming effects at the different stages. 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). 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 implicit learning has started to play a role (Bernolet & Hartsuiker). 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 implies 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 from natural L2 learning.

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:

Research Question 1. When in the learning trajectory of transitive structures do late learners of Dutch show priming for active and passive sentences?

Research Question 2. 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 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 item-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). 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:

Hypothesis 1. Abstract syntactic representations for transitives may occur earlier for active sentences than for passive sentences because actives are more frequent. 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.

Hypothesis 2. Learners may benefit from the few instances of lexical overlap in the intervention in the sense that 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 affects abstract structural priming of transitive structures in native language users, since we assume that they have already developed and established syntactic representations for active and passive structures. The lexical boost effect induced in the intervention block, which will 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 (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 following from the intervention items with verb overlap. First, the intervention draws attention to the passive structure due to the repeated use of this structure. 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).

1.2 Experiment 1: Control group

Method

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 (mean = 22.6, *SD* 4.1). For the sake of group comparison, we recruited approximately the same number of Dutch native speakers as L2 learners (see below).

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.

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., *bakker* "baker", *zangeres* "singer") and eight vehicles (inanimate entities, e.g., *ambulance* "ambulance", *brandweer* "firetruck"). For the pictures that contained animate entities, we chose six different transitive verbs: *groeten* "to greet", *roepen* "to call", *helpen* "to help", *bedienen* "to serve", *bellen* "to phone", and *dragen* "to carry". Also, for the pictures with inanimate entities, we chose six different transitive verbs: *raken* "to hit", *inhalen* "to pass", *blokkeren* "to block", *vervangen* "to replace", *slepen* "to drag", *volgen* "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., *lachen* "to smile", *huilen* "to cry") and twelve intransitive verbs for the pictures with inanimate entities (e.g., *stinken* "to smell", *stoppen* "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 speakers.

Design

The experiment included audio prime sentences, a verification task, and target description pictures. There were 72 prime sentences, of which 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. 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 *De leraar bedient de dokter* “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 *De dokter bedient de leraar* “The doctor serves the teacher” or with a passive sentence *De leraar wordt bediend door de dokter* “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 *De leraar zingt* “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 pseudo-randomly 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 survey (Qualtrics, Provo, UT), which contained the instructions, the request to provide informed consent and some demographic questions (age, gender, language background).

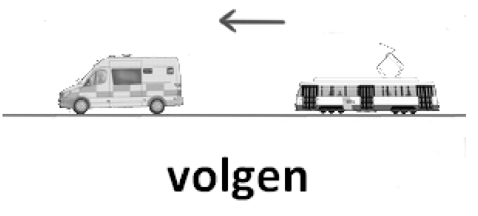


Figure 1.2: An example of a target picture.

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 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, and they had to formulate a sentence. The native speakers took about 20 to 25 minutes to complete the experiment.

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 (*worden* "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 1.2 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.

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 fitted to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer to increase convergeability (Powell, 2009) in *R* (version 4.1.3). We ran a model with Prime Condition (baseline/active/passive) and Intervention (pre/post) and their interactions as fixed factors. The baseline Prime Condition and Pre-intervention served as the reference level. Conform 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).

Results

The native speakers produced 505 responses, of which 395 active responses (74.2%), 110 passive responses (20.7%) and 27 "Other" responses (5.1%). Figure 1.3 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.

The final model included random intercepts for Participant and Item and no random slopes. The model output is reported in the appendix in Table A.1.1. In the control group, we found significant passive priming (β = 0.93, *SE* = 0.37, $p < .05$) and significantly more passives after the intervention block (β = 0.95, SE = 0.32, p < .01) 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.

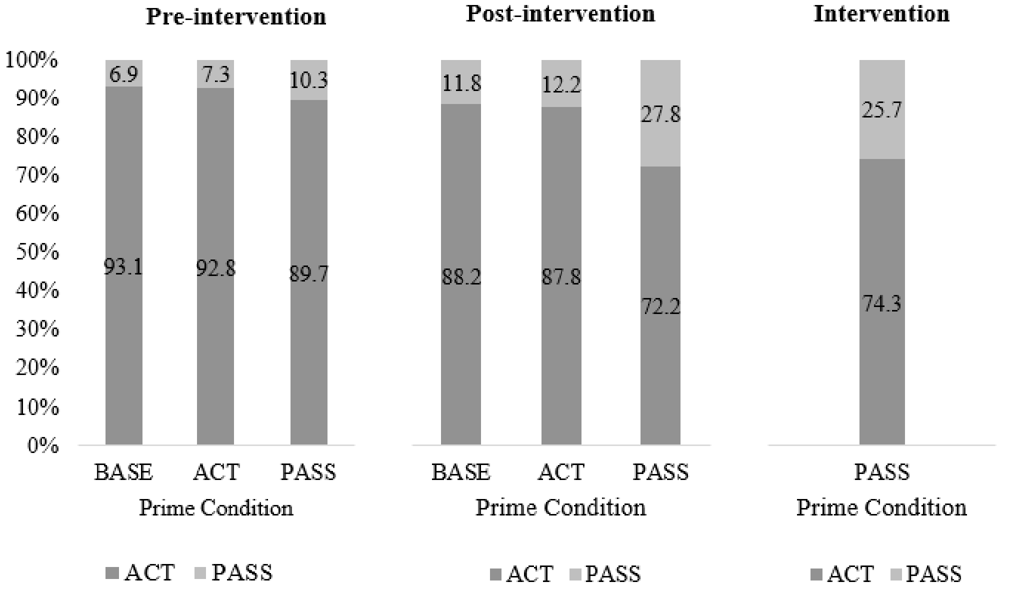


Figure 1.3: Proportion of active and passive responses before, after and in the intervention block for the control group.

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 shows 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% in the passive prime condition 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

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

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

1.3 Experiment 2: Longitudinal study

Method

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 [mean = 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.

Materials

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

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

[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.

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

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 always received a different list. 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. 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.

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.

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 generalized linear mixed model. Our model consisted of the factors Prime Condition (baseline/active/passive), Session (1 to 5) and Intervention (pre/post) and their 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.

Results

We collected 2232 responses, of which 1105 were actives (49.5%), 666 passives (29.9%) and 461 “Other” responses (20.7%). Table 1.1 shows the responses per session. We excluded the “Other” responses from further analyses.[3] Figure 1.4 shows the proportion of passive responses per prime condition, per session, per prime condition before and after the intervention. Figure 1.5 shows the proportion of passive responses per session during the intervention.

Table 1.1: Active, Passive and Other responses per session.

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 differ per session due to technical issues or illness of participants.

[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.

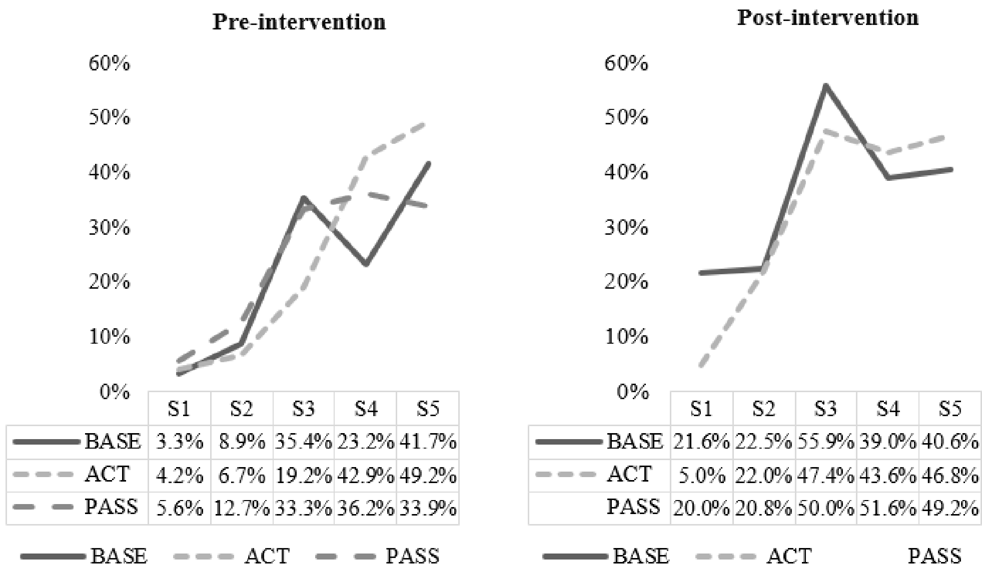


Figure 1.4: The proportion of passive responses per session, per prime condition before and after the intervention block.

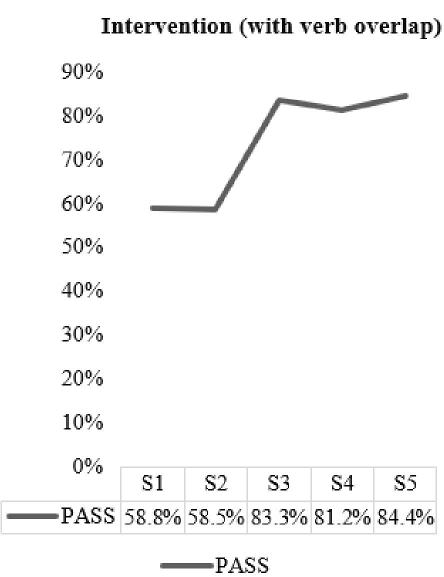


Figure 1.5: The proportion of passive responses per session in the intervention block.

We fitted the target responses to a generalized linear mixed model. The final model included random intercepts for Participant and Item, but no random slopes. In Table A.1.2 of the appendix, we report the model output. 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.

The results indicate that the proportion of passives increased linearly over time ($\beta = 2.52$, $SE = 0.47$, $p < .001$), then stabilized ($\beta = 0.85$, $SE = 0.33$, $p < .01$). Participants also produced more passives after the intervention than before the intervention ($\beta = 1.10$, $SE = 0.18$, $p < .001$), but this effect decreased linearly over time ($\beta = -1.12$, $SE = 0.46$, $p < .05$). There was a significant interaction between the active prime condition and the fourth derivative of Session ($\beta = -1.01$, $SE = 0.38$, $p < .01$), indicating that participants produced more active sentences after an active prime sentence (i.e., active priming) in the earliest session. The active priming effect diminished in the next two sessions. In Session 4, participants produced more passive sentences after an active prime than after a baseline prime. This effect disappeared 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 1.6 shows the proportion of correct responses per Prime condition for both sessions.

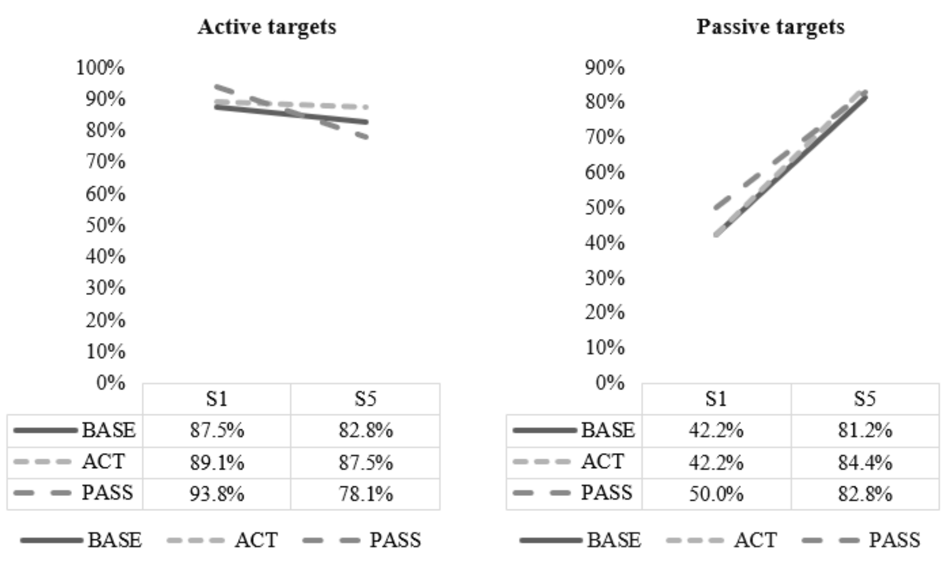


Figure 1.6: Correct responses in the comprehension targets, per prime condition per Session.

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.

Discussion

The results of the longitudinal study showed an increase of the use of the passive structure over sessions. The growth decreased after Session 4. We found 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 produced 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 at this stage, the learners are certainly proficient enough to produce the passive structure, given the overall high proportion of passive responses and the high accuracy rates in comprehension at Session 5. This may be due to long-term learning effects of the earlier sessions. Because the passive structure is more conspicuous than the active sentences and the intransitive filler sentences, due to its length and complexity, participants may have become aware of the fact that we were interested in the passive structure, as participants reported in the debriefing survey after the final session. 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.

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, but the repeated measures design did not allow us to follow the developmental path of language learning.

1.4 Experiment 3: Cross-sectional study

Method

Participants

We recruited participants who were taking Dutch classes at a center for adult education based in Antwerp. 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 lower proficiency group and 20 participants from the higher proficiency group participated in our study. There were 23 females and 15 males, who were between 19 and 53 years (mean = 33.7, *SD* 7.6). The participants had varying first languages (e.g., 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.

Materials

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

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, and 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, and some participants in this study only had a mobile phone at their disposal.

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.

Coding

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

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 (Powell, 2009) to increase convergeability. 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).

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 1.7 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.

Our final model consisted of significant main effects for Condition, Proficiency and Intervention. The model output is reported in the appendix in Table A.1.3). 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 ($\beta = 1.21$, $SE = 0.33$, $p < .001$). We also observed a main effect of Proficiency: higher proficiency speakers significantly produced more passives in the baseline condition compared to lower proficiency speakers ($\beta = 2.04$, $SE = 0.59$, $p < .001$). 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 ($\beta = 1.50$, $SE = 0.59$, $p < .001$) for both proficiency levels. Interestingly, we observed a marginal negative structural priming effect for actives ($\beta = 0.63$, $SE = 0.34$, $p < .1$). 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 1.7 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 condition. We did not find interactions between our predictors, probably because of the small number of observations in the data.

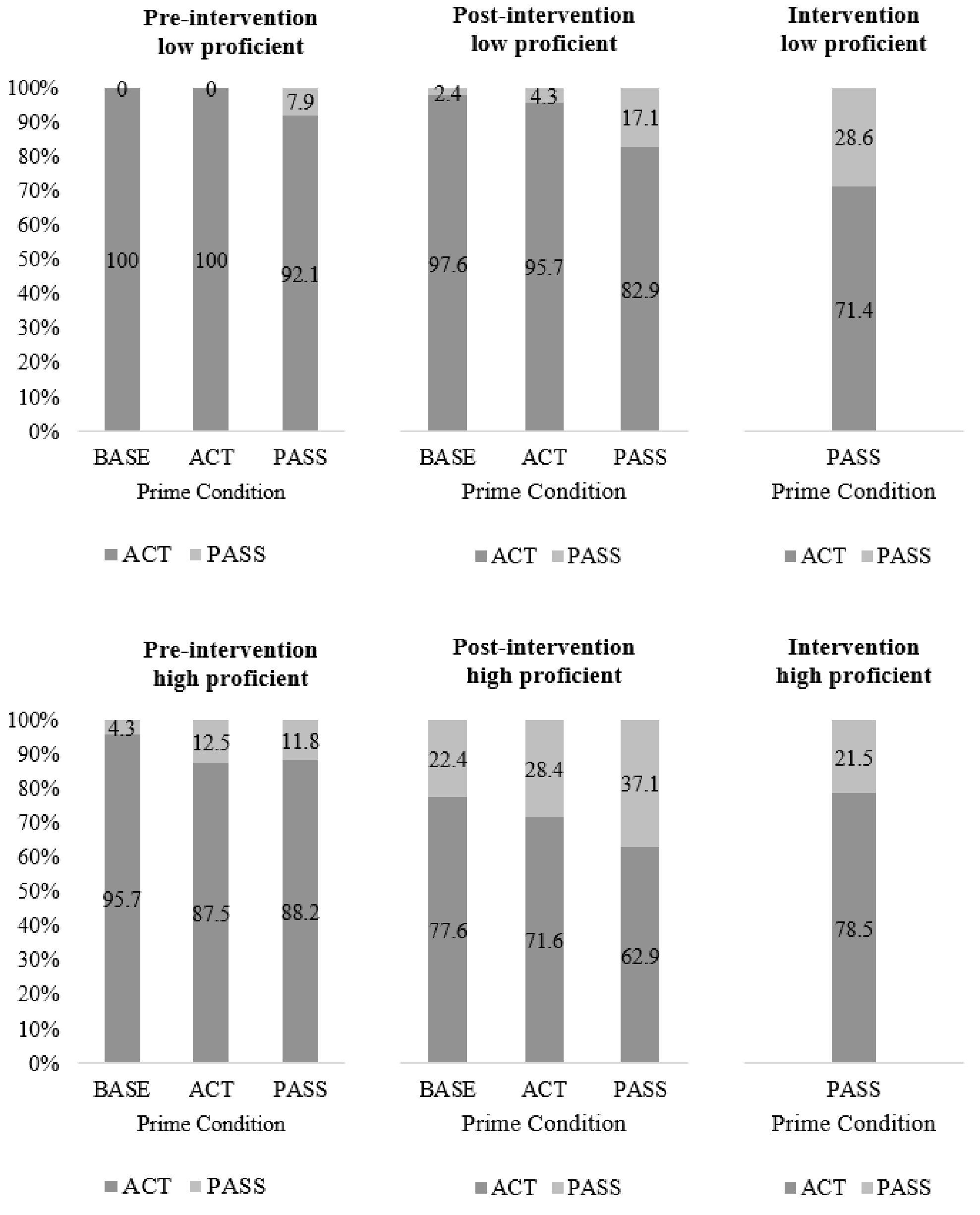


Figure 1.7: Proportion of active and passive responses before and after the intervention block for the lower and higher proficiency speakers.

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 (different from 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 (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).

1.5 General discussion

Revisiting the research questions

Research Question 1: 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 in the absence of significant passive priming. 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 was 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 did not find passive priming, although some passives were 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. In the cross-sectional study, we found 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.

Research Question 2: 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 more 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 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 stronger structural representations of the passive, 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.

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 reach conclusive insights into the process of late L2 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.

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.). 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. Momma, 2022; Reitter et al., 2011; Segaert et al., 2011 for monolingual hybrid models of structural priming), integrating both implicit learning mechanisms and the residual activation model.

1.6 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 sped 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.

Chapter 2

Shared and connected syntactic representations. The production preferences and priming effects of Dutch passives in Arabic/Berber-Dutch and Turkish-Dutch heritage speakers

Cross-linguistic structural priming effects suggest that bilinguals have shared or connected memory representations for similar syntactic structures. This predicts an influence of the production preferences of one language in the other language (Bernolet & Hartsuiker, 2018). We hypothesized that shared structures will lead to a facilitatory effect on production frequencies, whereas connected structures may sometimes lead to an inhibitory effect due to competition between structures. We compared the production preferences and priming effects in Dutch for the frequent by-phrase-final and the uncommon by-phrase-medial passive between Arabic/Berber-Dutch and Turkish-Dutch heritage speakers and native speakers of Dutch. Arabic/Berber-Dutch speakers produced more agentless passives, that is, the alternative shared between their two languages. In contrast, Turkish-Dutch speakers produced less by-phrase-medial passives, although these are less uncommon in Turkish. This inhibition effect suggests that syntactic structures may sometimes be connected rather than shared, although the exact mechanisms behind the inhibitory effects require further research.

Keywords: structural priming, shared syntax, passives, bilinguals, heritage speakers

Materials, data, and analyses are available online: `https://osf.io/kg7w9`.

2.1 Introduction

According to a Belgian newspaper, young people in the Belgian city of Antwerp speak *Illegaals* “Illegalish”, referring to youth language that is a variety of the Antwerp dialect of Dutch with influences of Arabic, Berber and Turkish (De Preter, 2011). Speakers incorporate words such as *shmetta* “coward” and *wajo* “wow” in their language, which were introduced by young people who speak Arabic, Berber or Turkish as their home language rather than or together with Dutch, the dominant language spoken in Antwerp. The borrowing of lexical items from these home languages (also called heritage languages) is a prominent feature in the language use of Arabic/Berber-Dutch and Turkish-Dutch bilinguals, but presumably the home language affects domains other than the lexicon as well. In the current study, we investigate the effects of heritage languages on the dominant language in the syntactic domain.

When bilinguals listen to or speak a language, both languages are active in the brain (see Kroll & Dussias, 2012; Kroll & Gollan, 2014 for a review on comprehension and production respectively). The fact that the two languages influence each other in proficient bilinguals suggests that the cognitive representations of the two languages are largely shared (cf. Kroll et al., 2015). Indeed, studies with late L2 learners show that proficient L2 learners have shared or connected representations of syntactic structures whenever these structures are similar enough (e.g., Bernolet et al., 2009). It is therefore likely that heritage speakers with a high proficiency in both the heritage language and the dominant language have shared or connected syntactic representations of structures that are similar across the heritage language and the dominant L2 language.

Similar syntactic structures that occur in both the heritage language and the dominant language may nevertheless have different properties in the respective languages, such as relative frequencies and production preferences. The presence or absence of similar structures in the heritage language may affect the syntactic representations of the constructions that need to be acquired in the dominant language. In our study, we report evidence for cross-linguistic influence (i.e., a facilitatory effect on production preferences) as well as cross-linguistic overcorrection (i.e., an inhibitory effect on production preferences). Based on these findings, we hypothesize that the direction of the effect on production preferences of syntactic structures may depend on whether structures are shared or connected. Shared structures may lead to an increase in the production of that structure due to frequent activation of that structure in both languages taken together, whereas connected structures may be produced less due to inhibitory effects resulting from competition between structures during sentence processing. In order to compare the production preferences and structural priming effects for passives in Arabic/Berber-Dutch and Turkish-Dutch heritage speakers to those in Dutch native speakers, we used the structural priming paradigm (Bock, 1986).

Mental representations of syntactic structures

The nature of mental representations of syntactic structures is often studied by means of the structural priming paradigm, exploiting the tendency of speakers to repeat previously processed syntactic structures (Bock, 1986). For instance, when participants are primed with a passive sentence (e.g., *the elephant is treated by the veterinarian*), they are more likely to describe a transitive target item with a passive sentence (*the cheese is being eaten by the mouse*) rather than an active sentence (*the mouse is eating the cheese*) than in an unprimed condition. There are at least two competing accounts of structural priming effects. Pickering and Branigan (1998) assume that structural priming is a short-lived effect caused by residual activation of combinatorial nodes connected to lemmas of verbs and nouns (which contain information on the syntactic structures in which these verbs and nouns can occur), whereas Chang et al. (2006) suggest that priming is a long-lasting effect that occurs due to the error-based, implicit learning of syntactic structures.

Some evidence is more consistent with the residual activation model, such as the lexical boost effect (priming is stronger when the head of the construction is repeated between prime and target, Pickering & Branigan, 1998). Other evidence must be explained from some implicit learning mechanism, especially evidence which points towards more permanent effects of structural priming, such as long-lasting priming (Bock & Griffin, 2000) and effects of verb bias (Bernolet & Hartsuiker, 2010).

More recent explanations of structural priming attempt to integrate the two accounts in a hybrid model, for example Reitter et al. (2011) and Segaert et al. (2011). In brief, these hybrid models assume that structural priming is the result of residual activation of the combinatorial node, which is modulated by a base-level activation of the syntactic structure. The base-level activation of syntactic structures arises due to implicit learning. We will sketch a hybrid model which in essence is similar to what Reitter et al. and Segaert et al. propose, but which is tailored to explaining how (bilingual) production preferences follow from the structural representations of syntactic structures. We assume links between verbs and nodes with syntactic information, and the relative strength of these links is determined through implicit learning (cf. Dell & Chang, 2014). Verbs of different languages may be linked to shared or connected nodes with syntactic information. Production preferences in both monolinguals and bilinguals may follow from the relative strength of the links.

Hybrid model of structural representations

According to Levelt et al.'s (1999) model of speech production, syntactic information is stored in lemmas in the lexical stratum, also called the lemma stratum (cf. Indefrey & Levelt, 2004; Roelofs, 1992, 1997). The lemma stratum consists of a network of lemma nodes containing lexical information that are connected to combinatorial nodes containing syntactic information (Pickering & Branigan, 1998). During speech production, activation spreads through this neural network of nodes. The highest activated lemma is chosen during the stage of lexical selection (cf. Levelt et al.). As a consequence of this activation spreading, the combinatorial nodes to which the selected lemma is connected are activated as well.

To illustrate, the verb *give* is connected to a double object node and a prepositional object node (Pickering & Branigan, 1998). If the double object node receives the highest activation, the phrase *give the dog a bone* would be selected for production, whereas a higher activation of the prepositional object node would lead to the production of the phrase *give a bone to the dog*.

Which node receives the highest activation is partly determined by the strength of the connections between the lemma and the nodes. The strength of these links is determined through implicit learning (either an error-based mechanism, cf. Chang et al., 2006; Dell & Chang, 2014, or an activation-based mechanism, cf. Reitter et al., 2011): through the processing of structures, the relative weight of their representations is strengthened. A relatively stronger connection between a lemma and a combinatorial node means that the lemma has easier access to the grammatical construction. So, when the phrase *give the* *dog a bone* is processed, the lemma *give*, the double object node and the link between the lemma and the combinatorial node will be activated. As a consequence, the connection between the verb give and the double object node is strengthened. Presumably, this also leads to a higher base-level activation or a higher relative weight of the double object node itself, which means that also with other verbs than *give*, the double object node will be more easily activated after processing this structure. Processing a double object phrase thus leads to permanent adjustments to both the verb-specific preferences and the general production preferences of a syntactic structure. As such, more previous experience with the double object construction than with the prepositional object construction leads to a long-term production preference for the double object dative.

Generally, structures with a higher base-level activation are produced more often than structures with a lower base-level activation. In structural priming experiments, production preferences are reflected via the inverse preference effect: less frequent structures show stronger priming effects than more frequent structures (V. S. Ferreira & Bock, 2006). Additionally, Coyle and Kaschak (2008) found that verb bias effects are present in long-term priming, which suggests that production preferences reflect the strengths of the links between verbs and combinatorial nodes. However, the bias of one verb affects the choice for a particular structure with other verbs as well (Bernolet & Hartsuiker, 2010). Hence, verb-specific production preferences may arise through the strength of the connection between the verb and the combinatorial nodes, and there may be a more general production preference for one grammatical structure over the other as a consequence of a higher base-level activation of the combinatorial node of that structure. So, production preferences of syntactic structures seem to be partly verb-specific, and partly independent from verb bias.

Bilingual syntactic representations

What happens to the production preferences of similar syntactic structures in bilingual speakers? Hartsuiker et al. (2004) found that Spanish-English bilingual participants produced more passive sentences in English after a passive prime sentence in Spanish than after a Spanish active prime sentence. These between-language structural priming effects suggest that, assuming the model of Pickering and Branigan (1998), the lemma nodes of Spanish transitive verbs and English transitive verbs are connected to the same combinatorial nodes containing the grammatical information on actives and passives. Hence, combinatorial nodes may not be language-specific and may thus be shared between languages. A consequence of this sharing of combinatorial nodes might be that production preferences are shared between languages as well (Bernolet & Hartsuiker, 2018).

Alternatively, the combinatorial node of a particular structure of one language may be connected to the combinatorial node of that structure of another language. Van Gompel and Arai (2018) argue that only structures that are completely identical in terms of constituent order and hierarchical structure are fully shared, whereas structures that are similar but not identical are connected rather than shared. Between-language priming effects only imply that structures are at least connected. If the same combinatorial node is activated during syntactic processing in both languages, between-language priming effects should be equally strong as within-language priming effects (Hartsuiker & Pickering, 2008), which was indeed found by Kantola and van Gompel (2011). If on the other hand combinatorial nodes are connected rather than shared between languages, between-language priming effects should be weaker than within-language priming effects (at least if one assumes the architecture of Pickering and Branigan [1998] in which multiple verbs within a language share their combinatorial nodes), as priming resulting from the repeated use of one combinatorial node is stronger than priming resulting from co-activated nodes (due to activation loss between input and output nodes). Several studies found stronger within-language priming than between-language priming (e.g., Bernolet et al., 2013; Cai et al., 2011), and a recent simulation model also suggested that this is the case (Khoe et al., 2021).

However, the shared-syntax account and thus the prediction of equally strong within-language priming and between-language priming may only apply to highly proficient L2 learners. Bernolet et al. (2013) found that between-language priming is modulated by proficiency. Structural priming effects between languages seem to become stronger as the L2 proficiency increases, leading to differences in the strength of within- and between-language priming in early learners. Learners might start with item-specific and language-specific (i.e., non-shared) syntactic representations in their L2. Over time, these representations become abstract and shared between the L1 and the L2 (see Bernolet & Hartsuiker, 2018; Hartsuiker & Bernolet, 2017 for a developmental model of the process during which syntactic structures become shared). Therefore only high proficient L2 learners may show abstract structural priming (i.e., priming without lexical overlap) and equally strong between- and within-language priming (i.e., priming based on shared syntactic structures). As such, we may still not expect equally strong between- as within-language priming in studies testing L2 learners who are not highly proficient, even if structures are eventually shared between languages.

Bilingual production preferences

As discussed above, the magnitude of between-language and within-language priming effects is not decisive with regard to the debate on whether combinatorial nodes of syntactic structures are shared or connected between languages in highly proficient bilinguals. Instead, investigating the production preferences of bilingual speakers may inform this debate, as production preferences often differ between languages. Flett et al. (2013) investigated the influence of L1 syntactic preferences on L2 production by testing the dative alternation in late learners of English with Spanish as their L1. Unlike English, Spanish only uses the prepositional object dative. We would therefore expect that the Spanish-English bilinguals would produce a larger proportion of prepositional object constructions in English as well. However, the bilinguals did not produce more prepositional object datives in English than the English-speaking control group, so they did not find an influence of L1 preferences on production in the L2. Flett et al. explained this by arguing that - even in shared structures - the production preferences are language-specific. Nevertheless, they only tested items with verb overlap between prime and target. Consequently, the priming effects may be mainly determined by the strength of the connection between the verb and the combinatorial node rather than on the base-level activation of the combinatorial nodes of syntactic structures. Cross-linguistic influence on production preferences of syntactic structures generalized over verbs should therefore be tested in an experiment without lexical overlap between prime and target, as any effects could then be attributed to the activation of the combinatorial nodes themselves.

In an experiment without lexical overlap between primes and targets, Kootstra and Şahin

(2018) found that Papiamento speakers in the Netherlands use more prepositional object datives than Papiamento speakers in Aruba. The prepositional object dative is much more frequent in Dutch than in Papiamento, since Papiamento has a strong preference for the direct object dative. As Papiamento speakers in the Netherlands are exposed more to Dutch than Papiamento speakers in Aruba, their production preferences of the dative construction in Papiamento seem to be affected by the Dutch production preferences. This increase in the use of the prepositional object dative may be explained by assuming shared combinatorial nodes of syntactic structures. If bilinguals have shared syntactic representations, there is one single combinatorial node for a particular structure in both languages. Exposure to that structure in either language adds to the base-level activation of the combinatorial node of that structure. If one language has a strong preference for one particular structure, the relative weighting of that structure might thus be higher in the other language as well, leading to an increased production of that structure.

In Kootstra and Şahin (2018), the structures under study were equivalent in both languages. If the structures are not exactly similar, they may be connected rather than shared (van Gompel & Arai, 2018). Connected syntactic representations, on the other hand, may sometimes lead to a decreased production of that structure. Kupisch (2014) investigated adjective placement in German-Italian bilinguals, who have either German or Italian as their dominant language. German only has prenominal adjectives. Italian uses postnominal adjectives, but some adjectives can also occur before the noun. The bilinguals who had German as their dominant language did not produce more prenominal adjectives, but rather more postnominal adjectives than the bilinguals dominant in Italian. Kupisch suggests that the bilinguals have three separate syntactic representations: the German prenominal adjective, the Italian prenominal adjective and the Italian postnominal adjective. During sentence selection, there is competition between the three alternatives, and the bilingual speaker needs to inhibit the alternative from the non-target language. As there is larger competition between similar structures (the German prenominal adjective and the Italian prenominal adjective) than between different structures (the German prenominal adjective and the Italian postnominal adjective), the Italian prenominal adjective is inhibited and the Italian postnominal adjective is overused. Anderssen et al. (2018) found the same pattern with prenominal and postnominal possessive structures in heritage speakers of Norwegian with English as their dominant L2. This inhibitory effect is called cross-linguistic overcorrection, and is presumably only found if the relative frequency of the overlapping structures is the opposite between the two languages.

Anderssen and Westergaard (2020) propose that cross-linguistic overcorrection only takes place if one of the languages lacks one of the alternatives available in the other language (e.g., German does not have a postnominal adjective), which they call partial overlap. If both languages have the same syntactic alternatives (which they call total overlap), but differ in the relative frequencies of these alternatives, cross-linguistic influence is supposed to occur. This is also what they found for subject-initial and object-initial clauses in Norwegian-English bilinguals. In English, subject-initial clauses are preferred, whereas Norwegian prefers object-initial clauses. The Norwegian-English bilinguals in their study showed an increased production of the subject-initial clause in Norwegian as an effect of the English production preference for the subject-initial clause.

Inhibitory effects such as cross-linguistic overcorrection are not expected to occur under a shared syntax account, as a shared syntax account does not predict competition between similar structures across languages. More specifically, the developmental model of

Hartsuiker and Bernolet (2017) only predicts cross-linguistic influence, as long as syntactic structures of the two languages share a representation and thus a single combinatorial node. Inhibition presumably arises due to competition between combinatorial nodes. Any inhibitory effects between similar structures may therefore be attributed to separate combinatorial nodes that are connected, rather than shared, and that compete with each other during the selection stage of language production. Such effects are known from word selection: picture naming proceeds faster if the name of the object depicted has been processed recently (e.g., Wheeldon & Monsell, 1992). However, participants are slower to name a pictured object if they were primed with a semantically related word, suggesting that competition takes place between lexical neighbors (Wheeldon & Monsell, 1994).

The two-stage competition model for the production of syntax by Segaert and colleagues (2011, also see Segaert et al., 2014, 2016) assumes that such competition occurs in sentence selection as well. Structural alternatives are assumed to be connected by inhibitory connections, sending lateral inhibition during the selection process. Note that this inhibitory mechanism, which has been proposed to reconcile structural priming effects found in response tendencies and production latencies, describes competition between the nodes of a particular structural alternation within a language rather than competition between the combinatorial nodes of different languages. Nevertheless, the point remains that inhibitory effects can occur when nodes are connected rather than shared.

To sum up, there are two possible ways in which production preferences can differ between languages. First, languages may have the same number of syntactic alternatives, but differ in the relative frequencies of these alternatives (i.e., total overlap). Alternatively, one of the languages may completely lack one of the alternatives (i.e., partial overlap). If syntactic structures are shared, we expect cross-linguistic influence in either situation: a preference for a structure in one language leads to a higher relative frequency of that structure in the other language (Bernolet & Hartsuiker, 2018). If, on the other hand, structures are connected, we may expect different outcomes between the two situations. If languages have the same syntactic alternatives (i.e., total overlap), we may still find cross-linguistic influence (Anderssen & Westergaard, 2020). But if one of the languages does not have all the syntactic alternatives available in the other language (i.e., partial overlap), we may find cross-linguistic overcorrection, that is, a decreased production of the structure that overlaps between the languages.

The passive alternation in Dutch

Both situations may occur in bilingual speakers of Dutch. In Dutch, the passive is formed with the auxiliary verb *worden* and the past participle. The agent is expressed using a prepositional phrase with the preposition *door*. The by-phrase can occur in sentence-final position (example 1) and in sentence-medial position (example 2). Alternatively, the by-phrase may be left out, resulting in the short passive (SP) (example 3).

(1) **PP-final passive**

Het	broodje	word-	t	gegeten	door	de	jongen
The	sandwich	AUX	3SG	eat.PTC	by	the	boy

"The sandwich is being eaten by the boy."

(2) **PP-medial passive**

Het	broodje	word-	t	door	de	jongen	gegeten
The	sandwich	AUX	3SG	by	the	boy	eat.PTC

(3) **Short passive**

Het	broodje	word-	t	gegeten
The	sandwich	AUX	3SG	eat.PTC

Bernolet et al. (2009) argued that the Dutch PP-final passive and the PP-medial passive do not differ significantly from each other in terms of information structure, but only in terms of constituent structure. Furthermore, they showed that the PP-final passive and the PP-medial passive can be primed separately. As such, these two structures presumably have separate combinatorial nodes competing with each other during sentence selection. The PP-final passive is more frequent than the PP-medial passive in Dutch.

In Turkish, the same syntactic alternatives are available as in Dutch (including other word order variations, such as the PP-initial passive, the discussion of which is beyond the scope of this paper). Similar to Dutch, the by-phrase is often omitted in a short passive (Göksel & Kerslake, 2005) (example 6). If the agent of a transitive sentence is overtly expressed and a full passive is produced, the constituent indicating the agent (by means of the postposition *tarafından*) mostly occurs immediately before the verb (Göksel & Kerslake; Ketrez, 2012), which means that on the level of constituent structure, the Turkish passive corresponds to the Dutch PP-medial passive (example 4). Since word order in Turkish is relatively free, Turkish also has a PP-final passive (example 5). So, in the case of the PP-final and PP-medial passive alternation, Dutch and Turkish languages share the syntactic alternatives (i.e., total overlap), but differ in the relative frequencies of these alternatives. The examples are adapted from Öszoy (2009, p. 6).

(4) **PP-medial passive**

araba	soför	tarafından	sür-	ül-	dü
car	driver	by	drive-	PASS-	PAST

"The car was driven by the driver."

(5) **PP-final passive**

araba	sür-	ül-	dü	soför	tarafından
car	drive-	PASS-	PAST	driver	by

(6) **Short passive**

araba	sür-	ül-	dü
car	drive-	PASS-	PAST

Arabic usually does not express the agent, nor does Berber (Gutova, 2013), i.e., most passive sentences lack a by-phrase (cf. Shaqråa, 2007; Badawi et al., 2004; Loutfi, 2015). Example 7 illustrates the passive in Moroccan Arabic. (The Arabic-speaking participants that we tested speak a Moroccan variant, since there is a relatively large group of second and third generation immigrants from Morocco living in Belgium.) In Moroccan Arabic, the morphology of the passive is different from in Modern Standard Arabic. Moroccan

Arabic marks the verb with the prefix *t-*. The formation of the passive in Berber resembles the passive construction in Moroccan Arabic (example 8). So, Dutch, Arabic and Berber all have a short passive, but the PP-final and the PP-medial passive are specific to Dutch (i.e., partial overlap).

(7) **Arabic** (Loutfi, 2015, p. 9)

l-	bab	t-hərrəs
the	door	was broken

"The door was broken." (Moroccan Arabic)

(8) **Berber** (Gutova, 2013, p. 111-112)

aḥuři	i-	ttwa-	ɣars
sheep:EL	3SG:M	PASS-	slaughter:P

"The sheep was slaughtered." (Tarifiyt Berber)

Note that van Gompel and Arai (2018) suggest that identical structures may be fully shared, whereas similar structures may be connected (although there is the theoretical possibility that identical structures are connected as well). In the current study, the syntactic alternatives that are available across the languages are similar rather than identical. Turkish uses postpositions rather than prepositions. In addition, Turkish is an agglutinative language, and passives are marked morphologically rather than syntactically. In Arabic and Berber, the passive is also marked morphologically, and none of the languages uses auxiliary verbs as opposed to Dutch. Therefore the passives in Arabic, Berber and Turkish are presumably predicted to be connected to the Dutch passives rather than shared under their account.

By contrast, the shared-syntax account (Hartsuiker et al., 2004; Hartsuiker & Bernolet, 2017) assumes shared representations for the most similar structure, despite differences in morphology and pragmatics (also see Hartsuiker et al., 2016, referring to priming for genitives between English [*the nun's hat*] and Dutch [*the non haar hoed* "the nun her hat"] [Bernolet et al., 2013]). As far as we know, there are no studies on between-language priming of passives between Arabic/Berber and Dutch and between Turkish and Dutch. However, the small-scale study of Arman Ergin (2019) reports passive priming between the Turkish PP-medial passive and the PP-final passive in English, implying that the structural representation of the morphologically formed Turkish passive activates the syntactically formed passive in English and therefore presumably also in Dutch.

Current study

In the current study, we compare the use of the Dutch passive by Arabic/Berber-Dutch and Turkish-Dutch heritage speakers to the use of the passive by native speakers of Dutch in a structural priming experiment. The production preferences of the Dutch passive may reflect how the base-level activation level of the combinatorial nodes of the Dutch passive is affected by long-term experience with Arabic/Berber and Turkish passives.

If bilinguals have shared syntactic representations (conforming to the shared-syntax account, Hartsuiker et al., 2004; Hartsuiker & Bernolet, 2017), then Turkish-Dutch bilinguals may produce more PP-medial passives than a Dutch group, and Arabic/Berber-Dutch bilinguals may use more short passives than Dutch speakers. If, on the other hand, syntactic structures are connected rather than shared (in line with van Gompel & Arai, 2018), then we still expect that Turkish-Dutch bilinguals use more PP-medial passives than a Dutch group, as Turkish and Dutch have total overlap of the alternatives available for the passive structure. But we would expect a decreased use of short passives in Arabic/Berber speakers due to cross-linguistic overcorrection, since there is only partial overlap between Arabic/Berber and Dutch: the short passive is available in both Arabic/Berber in Dutch, but the PP-final and the PP-medial passive are exclusively available in Dutch.

We chose to test heritage speakers, because Kupisch (2014) and Anderssen and Westergaard (2020) suggest that cross-linguistic overcorrection may only take place in highly proficient heritage speakers. Less proficient speakers or late L2 learners may have more difficulty in inhibiting the other language, leading to cross-linguistic influence even in the case of partial overlap. Since bilingual structural priming studies mostly involve late L2 learners, the current study thus involves an understudied population in the field of bilingual structural priming.

At the same time, most studies on the language of heritage speakers investigate the influence of the dominant L2 on the heritage language rather than vice versa. Any such influence is often attributed to factors such as incomplete L1 acquisition and/or language attrition (see Benmamoun et al., 2013 for a discussion). These factors explain deviant syntactic representations in the heritage language and imply that such an effect will not occur in the other direction, namely that the use of syntactic structures in the dominant L2 language will not be affected by the heritage language. A more recent explanation is the role of differential acquisition, acknowledging the fact that the quantity and the quality of the input of the language during language acquisition is different for heritage speakers than for monolingual speakers, which leads to different outcomes (Kupisch & Rothman, 2016). If it is the input of the language which explains the different use of syntactic structures in the heritage language, given that the input of the dominant L2 is also different for heritage speakers than for monolingual speakers, we may expect bidirectional influences between shared or connected syntactic structures in the heritage language and the dominant language. Therefore the production preferences of heritage speakers may not only be different from those of monolingual speakers in the heritage language, as has been demonstrated previously, but also from monolingual speakers in the dominant L2.

We tested the unprimed production preferences of transitive structures (including the active), immediately followed by a structural priming experiment in which we primed the PP-final passive and the PP-medial passive. In the priming experiment participants were required to start their sentence with the patient in order to avoid active responses. Participants could thus respond with a PP-final passive, a PP-medial passive or a SP. We did not prime the SP (which in natural language is the single passive option in Arabic/Berber, and also a frequent alternative in both Dutch and Turkish), because we believe that it would be unnatural in our picture description task showing two entities. We exploited the structural priming paradigm primarily to elicit the low-frequency PP-medial passive. If it would be the case that one of the bilingual groups strongly disprefers one of the passive alternatives and, consequently, would not use that structure spontaneously, structural priming allows us to see whether the bilinguals nevertheless have an underlying representation of the dispreferred structure. In addition, priming effects reflect production preferences. Due to the inverse preference effect (V. S. Ferreira & Bock, 2006), we expect to find stronger PP-medial priming than PP-final priming in the Dutch group. If the production preferences of the Arabic/Berber-Dutch and Turkish-Dutch bilinguals differ from those of the Dutch group, we may therefore find differences in the relative magnitude of the priming effects as well.

2.2 Method

Participants

We tested 144 participants: 48 participants who are all native speakers of Flemish Dutch, 48 early bilingual speakers of both Flemish Dutch and Arabic or Berber, and 48 participants who are early bilingual speakers of both Flemish Dutch and Turkish. Participants were classified as early bilinguals if they started learning both languages before the age of 6. All participants were aged between 16 and 30, had normal or corrected to normal vision and had no dyslexia. Participants gave their informed consent prior to the experiment and received a gift voucher for their participation in the experiment.

We asked participants to rate their proficiency in Dutch and in their other L1 on a 7-point scale for both language production and language comprehension. Participants reported a high to very high proficiency in both Dutch and the other L1 in active and receptive language use. Even though numerically, they report a bit lower proficiency in their heritage language than in Dutch, their proficiency in both languages is presumably high enough to assume connected or shared representations (cf. the self-rated proficiencies in Hartsuiker et al., 2016, which reports equally strong between-language priming and within-language priming in L1 Dutch-L2 English, L1 Dutch-L2 French bilinguals and L1 Dutch-L2 German bilinguals). To further assess their proficiency in Dutch, participants completed the LexTALE test for Dutch (Lemhöfer & Broersma, 2012). The heritage speakers scored numerically slightly lower on the LexTALE test than the L1 speakers. The LexTALE test measures vocabulary size, which is known to be lower in bilinguals than in monolinguals (e.g., Bialystok & Luk, 2012).

Table 2.1: Participants.

	Dutch (n = 48)	**Arabic/Berber** (n = 48)	**Turkish (n = 48)**
L1	Dutch	Dutch & Arabic (n = 25) Dutch & Berber (n = 21) Dutch, Arabic & Berber (n = 2)	Dutch & Turkish
Gender	6 male, 41 female, 1 other	7 male, 39 female, 2 other	14 male, 33 female, 1 other
Age	16-27, mean 21.43 (2.29)	16-26, mean 19.13 (2.52)	17-29, mean 21.27 (2.62)
Proficiency Dutch	production: 6.98 (0.14) comprehension: 7.00 (0.00)	production: 6.52 (1.31) comprehension: 6.77 (0.47)	production: 6.50 (0.65) comprehension: 6.87 (0.34)
Proficiency other L1	not applicable	production: 5.90 (1.31) comprehension: 5.33 (1.71)	production: 6.17 (0.82) comprehension: 6.60 (0.57)
LexTALE	88.4% (8.1)	78.8% (9.6)	83.4% (7.4)

Note: Standard deviations are indicated in parentheses.

Materials

The materials used were adapted from Bernolet et al. (2009) and included pictures from the International Picture Naming Project (see E. Bates et al., 2003). We constructed three sets of pictures: a pre-experimental baseline set, a target set and a verification set for the priming experiment.

The pre-experimental baseline set consisted of 12 target and 12 filler pictures. Target pictures showed an agent, a patient, and a Dutch transitive verb. There were 3 items for each combination of the agents' and patients' animacy (animate agent/animate patient [AA], animate agent/inanimate patient [AI], inanimate agent/animate patient [IA], inanimate agent/inanimate patient [II]). Filler pictures showed a person or an object and a Dutch intransitive verb. Six of the intransitive verbs were unergative and six were unaccusative.

For the target set, we constructed 36 target and 72 filler pictures. Target items showed an agent and a patient and a Dutch transitive verb. The patient was indicated by means of a red frame. In the pictures for the base prime condition, the verb was omitted from the target item, but the patient was still red-framed. There were 9 target items for each animacy combination (AA, AI, IA, II). Filler items either showed (i) an agent, a patient, and a Dutch transitive with a red frame around the agent; (ii) one object/person and a Dutch intransitive verb, or (iii) two objects/persons without a verb. Of the filler items with an intransitive verb, half of the intransitive verbs involved an unergative verb, whereas the other half used an unaccusative verb. Figure 2.1 shows a typical stimulus item. Figure 2.2 illustrates the three types of filler items.

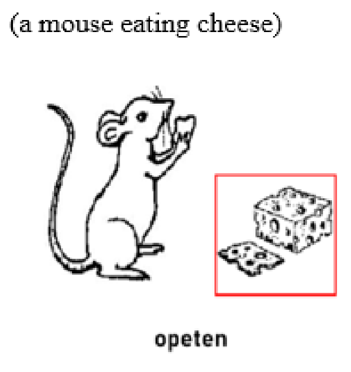


Figure 2.1: Example of a stimulus item.

The target items of both the pre-experimental baseline set and the target set were created in two variants: one with the agent depicted on the left and the patient on the right, and a mirrored variant with the patient depicted on the left and the agent depicted on the right. In previous priming studies with passives, the patient was often depicted on the left to elicit more passive responses (e.g., Bernolet et al., 2009; Bock, 1986). This argument is based on a reading direction from left to right. Since Arabic is read from right to left and this could potentially affect our results, we decided to counterbalance the position of the agent and the patient.

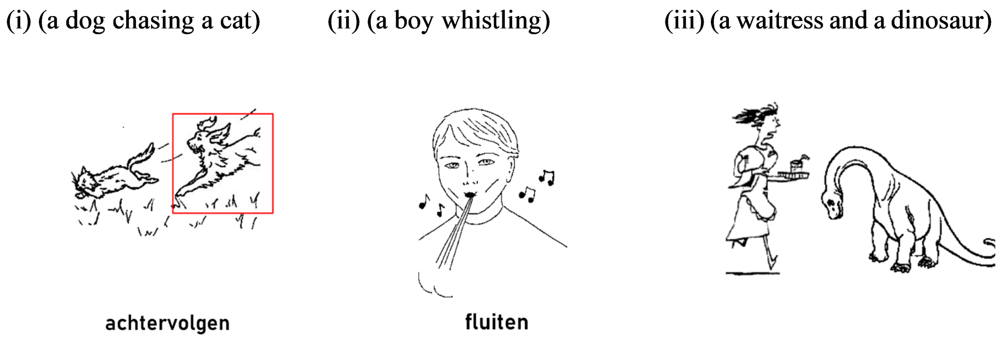


Figure 2.2: Examples of the three types of filler items.

We also constructed a verification set, which included 108 pictures that were similar to the target set. Half of the verification pictures matched the preceding prime sentence and half of the pictures did not match.

In addition to the three sets of pictures, we constructed a set of prime sentences. Similar to the target set, there were 36 critical prime sentences and 72 filler prime sentences. Each critical prime sentence was recorded in three variants, matching the prime conditions (base, PP-final passive, PP-medial passive, see example 9). The sentences were recorded by three female speakers: a speaker of Flemish Dutch, a bilingual Arabic-Dutch speaker, and a bilingual Turkish-Dutch speaker.

(9) (a) **Base condition**

de	dierenarts	en	de	olifant
the	veterinarian	and	the	elephant

(b) **PP-final passive**

De	olifant	wordt	behandeld	door	de	dierenarts.
The	elephant	AUX	treat.PTC	by	the	veterinarian

(c) **PP-medial passive**

De	olifant	wordt	door	de	dierenarts	behandeld.
The	elephant	AUX	by	the	veterinarian	treat.PTC

"The elephant is being treated by the veterinarian."

Design

We designed a pre-experimental baseline task and an experimental task (the priming experiment). For the pre-experimental baseline task, we constructed two lists of target pictures. Both lists started with one filler item and alternated between a target item and a filler item. The pseudo-randomized items were always shown in the same order. Half of the target items displayed the agent on the right and the patient on the left, and vice-versa in the other half of the target items. This was counterbalanced across the two lists, so each item appeared equally often either with the agent or the patient on the right.

As for the priming experiment, each item consisted of a prime sentence, a verification picture and a target picture. There was no lexical overlap between the prime sentences and the elicited target sentences. The priming experiment had a target-filler ratio of 1:2. We pseudo-randomized the order of the items in such a way that there was always at least one filler between two target items, and that the experiment started with three fillers. Three different lists were constructed, such that every item was preceded by a prime from a different prime condition (base, PP-final passive, PP-medial passive) across the lists. Within each list, the prime sentences were presented equally often in the three priming conditions. As for the pre-experimental baseline task, we counterbalanced the position of the agent and the patient in the target pictures. For this purpose, we constructed two variants of each list, which led to a total of six lists.

Procedure

Immediately preceding the priming experiment, we measured the pre-experimental baseline preference for the different transitive alternatives (including the active structure). Participants were told that they would practice with the production part of the priming experiment. They were shown a target picture and were asked to describe this picture using one sentence. These target pictures had no red frame around either the agent or the patient; hence, participants were free to produce either an active or a passive sentence. During the priming experiment, participants would first listen to the prime sentence through headphones. The voice they listened to belonged to a speaker with a similar language background: the Arabic/Berber-Dutch and Turkish-Dutch bilinguals listened to the sentences as recorded by an Arabic-Dutch speaker and Turkish-Dutch speaker respectively, and the Dutch group listened to a speaker of Flemish Dutch. They were then shown a verification picture and were asked to indicate whether this picture matched the preceding sentence by pressing 1 (matching) or 2 (not matching). After pressing one of the keys, the verification picture was replaced by the target picture. Participants were asked to describe this picture using a sentence that started with the figure indicated by the red frame (cf. the color-coded primes of Segaert et al., 2011).

In addition to the experimental task, participants completed a short language questionnaire and did the LexTALE test (a short yes/no-vocabulary test, Lemhöfer & Broersma, 2012). The sessions took place in a quiet room. A session took about 40 minutes. All sessions were recorded with an external audio recorder.

Coding

Pre-experimental baseline

The target responses of the pre-experimental baseline measurement were coded as Active, PP-final passive, SP, or "Other". (Note that no PP-medial passives were produced in the pre-experimental baseline.) For the coding of PP-final passives and SPs, the same criteria were used as for the coding of the priming experiment (see below). A response was coded as Active if the response included a form of the transitive verb, if the subject was an agent and if the object was a patient. An exception was made for active sentences that contained the past participle (mostly sentences with the present perfect, i.e., a form of the auxiliary *hebben* "have" and the past participle, see example 10). These were coded as

"Other", because they are active in terms of information structure, but their morphological complexity is similar to passives in Dutch. Furthermore, "Other" responses included responses in which a conjugated verb was missing or in which a different verb was used, responses in which either the agent or the patient was not mentioned, responses with reflexives (see example 11) and responses of any other structure.

(10) **Present perfect**

de	jongen	heeft	het	broodje	gegeten
the	boy	AUX.3SG	the	sandwich	eat.PTC

"The boy ate/has eaten the sandwich."

(11) **Reflexive**

het	meisje	steek-	t	zich	aan	de	cactus
the	girl	prick-	3 SG	REFL	to	the	cactus

"The girl gets pricked by the cactus."

(Expected response:

het	meisje	word-	t	gestoken	door	de	cactus
the	girl	AUX-	3 SG	prick.PTC	by	the	cactus

"The girl is being pricked by the cactus.")

Priming experiment

Target responses were coded as PP-final passive, PP-medial passive, SP or "Other". A response was coded as a PP-final passive if the response included a subject, an auxiliary (either *worden* or *zijn*), a past participle, and a by-phrase with the preposition *door* following the past participle. A response was coded as a PP-medial passive if the by-phrase preceded the past participle and if the same elements were present as for the PP-final passive responses. A response was coded as an SP if the response contained a subject, an auxiliary and a past participle, and if the agent was not mentioned. Different from the pre-experimental baseline experiment, active responses were coded as "Other". Any other response was coded as "Other" as well. "Other" responses included responses in which a different kind of passive construction was used (PP-initial by-phrase passives, passives with *er* "there" as subject instead of the agent, passives in which a different preposition than *door* was used in the PP mentioning the agent) and responses using any other construction.

Analysis

After the "Other" responses were excluded, the responses fell into the following categories: Actives, PP-final passives, and SPs in the pre-experimental baseline, and PP-final passives, PP-medial passives, and SPs in the priming experiment. Therefore the results needed to be analyzed using a multinomial generalized linear mixed model rather than a binomial model. We used the Markov chain Monte Carlo (MCMC) algorithm, which allows to approximate likelihood estimates over more complex data such as multinomial data.

The MCMC method utilizes a Bayesian framework, which exploits a prior distribution, a likelihood function and a posterior distribution. The prior distribution represents prior beliefs on the parameters, for example on probability. Defining a flat prior means that we believe that the probability will be anywhere between 0 and 1. The likelihood function describes what probability value is most likely given the observed data. The prior distribution and the likelihood function are then combined to determine the posterior distribution. The MCMC algorithm generates random samples to calculate the posterior distribution given the prior and the observed data. For each sample, it evaluates whether the random parameter values (i.e., those provided by the likelihood function) are better than the previously stored ones and if so, updates the parameter values of the posterior distribution, storing how much better the new values are. The reported posterior means are thus approximated through repeated sampling. The effective sample size is a measure of autocorrelation (i.e., the sampled parameter values are very similar to the directly preceding ones). The reported parameter value is more reliable if the effective sample size is closer to our sample size. The *p*-value indicates the probability that the parameter value is larger or smaller than 0.

Similar to a binomial generalized linear effects model, a significant p-value indicates that the posterior mean is significantly higher or lower compared to the reference level. Importantly, as there are three categories instead of two categories within the response variable, a significant effect in one category does however not imply a significant effect in the other category. To illustrate, in a binomial experiment with PP-final passives and PP-medial passives, a significant increase of PP-final passives in a particular condition entails a significant decrease of PP-medial passives. In a multinomial model, this is not the case. A significantly higher posterior mean for PP-final passives means that there are more PP-final passives in that condition compared to the reference level (SPs in our case), but does not say anything about the effect of the increase of PP-final passives on the proportion of PP-medial passives. An increase in the proportion of PP-final passives may go to the expense of both other categories (PP-medial passives and SPs), or may lead to a decrease in only one of the other categories.

We ran our analyses using the *R*-package *MCMCglmm* (Hadfield, 2010). We defined a flat prior following the recommendations of Levshina (2019). For each separate model, we set the number of iterations to 500,000. The burn-in period was set to 60,000 iterations and the thinning interval was 300. These settings led to good model diagnostics, i.e., a good mixture and an autocorrelation of less than 0.1, as recommended by Hadfield (2019).

2.3 Results

Pre-experimental baseline

The Dutch-speaking participants produced 432 Actives (75.0%), 51 PP-final passives (8.9%), 38 SPs (6.6%), and 55 "Others" (9.5%). The Arabic/Berber-Dutch group produced 384 Actives (66.7%), 39 PP-final passives (6.8%), 41 SPs (7.1%), and 112 "Others" (19.4%). The Turkish-Dutch participants produced 406 Actives (70.5%), 69 PP-final passives (12.0%), 28 SPs (4.9%), and 73 "Others" (12.7%). No PP-medial passives were spontaneously produced in any of the language groups. The "Other" responses were disregarded for further analyses. Figure 2.3 shows the production preferences for Actives, PP-final passives and SPs.

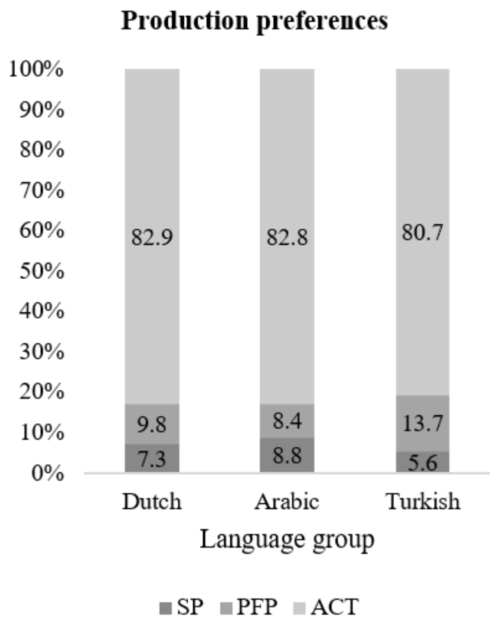


Figure 2.3: Responses in the pre-experimental baseline for each language group (in %).

The pre-experimental production preferences were compared between the different groups by fitting a multinomial generalized linear mixed model to the Active, PP-final passive and SP responses. Language was included as a fixed effect. Random effects were inserted for participants and items. The Active target responses and the Dutch language group were treated as the reference levels. The full model output is provided in the appendix in Table A.2.1. Participants produced significantly more Actives than PP-final passives (post. mean = -4.47 [-6.09, -2.76], $p < .001$) and SPs (post. mean = -4.22 [-6.01, -2.63], $p < .001$). In the pre-experimental baseline, the Arabic/Berber-Dutch and Turkish-Dutch group did not differ significantly from the Dutch group in any of the conditions.

Priming experiment

In the critical items of the priming experiment, participants were forced to start their response with the patient. Therefore any Active responses were considered as "Other" responses here. The Dutch-speaking participants produced 1,319 PP-final passives (76.3%), 147 PP-medial passives (8.5%), 92 SPs (5.3%) and 170 "Others" (9.8%). The Arabic/Berber-Dutch participants produced 1,068 PP-final passives (61.8%), 51 PP-medial passives (3.0%), 251 SPs (14.5%) and 358 "Others" (20.7%).[1] Finally, the Turkish-Dutch

[1]The Arabic/Berber speakers produce notably more "Other" responses than the other two groups. Importantly, the "Other" responses are not ungrammatical and should probably not be attributed to a lower proficiency. The high proportion of "Other" responses in the Arabic/Berber-Dutch group may be because Arabic/Berber-Dutch speakers are less likely to use full passives and consequently, have a general preference for other structures over the passive (including the short passive). For instance, participants used more intransitives (e.g., *de bal rolt weg*

participants produced 1,294 PP-final passives (74.9%), 63 PP-medial passives (3.6%), 175 SPs (10.1%), and 196 "Others" (11.3%). The responses per priming condition are summarized in Figure 2.4.

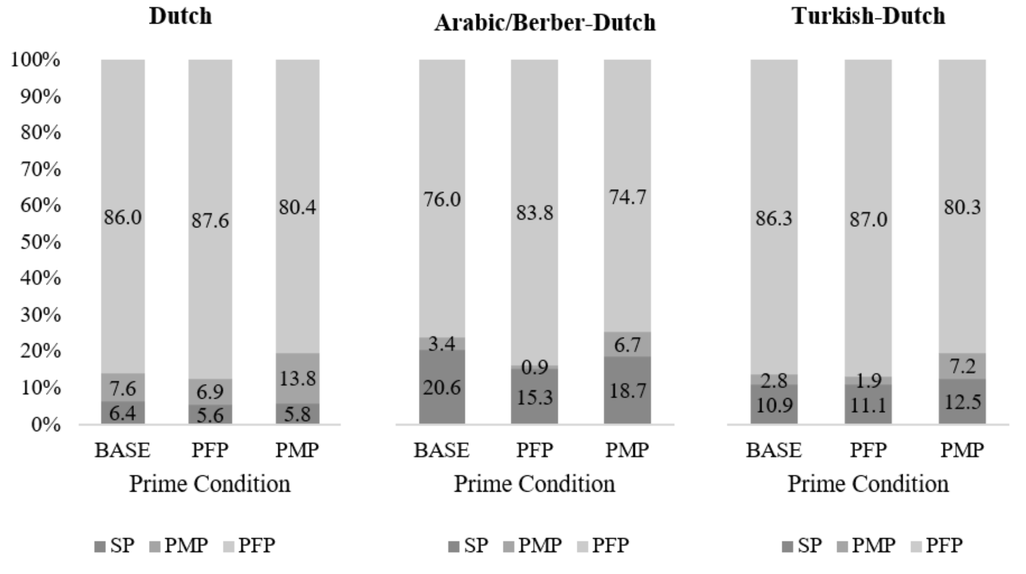


Figure 2.4: Responses per prime condition for each language group (in %).

We compared the production preferences and the priming effects between the three language groups using a multinomial generalized linear mixed model with Prime Condition, Language, and their interaction as fixed effects. We added random effects for participants and items. The reference levels were the SP target responses, the base prime condition, and the Dutch language group. Because of the inverse reading direction in Arabic, we counterbalanced the position of the agent and the patient in the pictures that participants had to describe rather than placing the patient on the left. There were no significant differences in response patterns between the pictures that had the patient on the left and the pictures that depicted the patient on the right, neither in the Arabic/Berber-Dutch group nor in the other language groups. Therefore this variable was not included in the final analyses. The full model output is reported in the appendix in Table A.2.2.

The Dutch participants showed an overall preference for PP-final passives (post. mean = 4.35 [3.47, 5.29], $p < .001$). They produced significantly more PP-medial passives after a PP-medial prime than after a base prime (post. mean = 1.44 [0.55, 2.30], $p < .001$). We did not observe a significant effect of the PP-final prime on the proportion of PP-final passives produced (post. mean = 0.29 [-0.37, 0.95], $p = 0.37$).

Arabic/Berber-Dutch speakers used significantly fewer PP-final passives (post. mean =

"the ball rolls away" instead of *de bal wordt weggerold door de vrouw* "the ball is being rolled away by the woman") and more reflexives (e.g., *het meisje steekt zich aan de cactus* "the girl gets pricked by the cactus" instead of *het meisje wordt gestoken door de cactus* "the girl is being pricked by the cactus").

-1.78 [-2.67, -0.81], $p < .001$) and PP-medial passives (post. mean = -2.40 [-4.19, -0.71], $p < .05$) than Dutch speakers, i.e., significantly more SPs. The Turkish-Dutch group produced fewer PP-medial passives than the Dutch group (post. mean = -2.41 [-4.28, -0.46], $p < .01$), but did not differ from the Dutch group with respect to the number of PP-final responses (post. mean = -0.56 [-1.56, 0.33], $p = .25$). The Arabic/Berber-Dutch and the Turkish-Dutch group did not differ from the Dutch group in terms of PP-medial priming (post. mean = -0.25 [-1.40, 1.04], $p = .67$ and post. mean = 0.39 [-0.98, 1.86], $p = .59$ respectively). The Arabic/Berber-Dutch and the Turkish-Dutch group also did not differ from the Dutch group in terms of PP-final priming, i.e., the number of PP-final passive responses after a PP-final passive primes compared to the amount after base primes (post. mean = 0.13 [-0.65, 1.01], $p = .79$ and post. mean = -0.42 [-1.31, 0.44], $p = .35$ respectively). Nevertheless, the Arabic/Berber-Dutch participants did use fewer PP-medial passives after a PP-final passive prime than the Dutch participants do (post. mean = -1.92 [-3.73, -0.37], $p < .05$).

2.4 Discussion

The pre-experimental baseline test did not show any significant differences between the Dutch, the Arabic/Berber-Dutch, and the Turkish-Dutch group. In the priming experiment, we did find differences in terms of production preferences. The Arabic/Berber-Dutch group produced significantly more SPs than the Dutch participants did. The Turkish-Dutch participants had significantly fewer PP-medial responses than the Dutch group. We found significant PP-medial priming but no PP-final priming in all three language groups. The Arabic/Berber-Dutch participants produced fewer PP-medial passives in the PP-final prime condition than in the base condition. Otherwise, the Arabic/Berber-Dutch and Turkish-Dutch participants did not differ from the Dutch group in terms of priming effects.

In the pre-experimental baseline test, the proportion of actives was more than 80% in each of the three groups. As a consequence of the high proportion of actives, the number of observations of the three passive structures (i.e., PP-final, PP-medial, and SP) is relatively low. Any differences with regard to the different forms of the passive structure are therefore hard to spot. These results confirm the need of our priming experiment in which we targeted the passive structure only.

We predicted that in the case of shared structures between languages, we would find an increased proportion of SPs in the Arabic/Berber-Dutch group and PP-medial passives in the Turkish-Dutch group as compared to the Dutch group. This would follow the developmental model of L2 syntax of Hartsuiker and Bernolet (2017). In the case of connected structures, we expected cross-linguistic overcorrection in the Arabic/Berber-Dutch group, that is, a decreased proportion of the SPs, and cross-linguistic influence in the Turkish-Dutch group, which would mean an increased proportion of the PP-medial passives. Such results would follow the predictions of Anderssen and Westergaard (2020).

Our results are not in line with either accounts. We find cross-linguistic influence in the Arabic/Berber-Dutch group, since the Arabic/Berber-Dutch participants produced significantly more SPs in Dutch than the Dutch participants. We find cross-linguistic overcorrection in the Turkish-Dutch group, as the Turkish-Dutch group produced fewer PP-medial passives than the Dutch group. Since the shared syntax account of Hartsuiker and Bernolet (2017) does not predict any inhibition effects such as cross-linguistic overcorrection, our results suggest that structures may be connected rather than shared, at least under certain circumstances, which may give rise to cross-linguistic overcorrection. Nevertheless, contrary to what Anderssen and Westergaard (2020) argue, partial or total overlap of the syntactic alternatives available between languages does not seem the factor that is decisive of whether there is cross-linguistic influence or cross-linguistic overcorrection. We find cross-linguistic influence in the case of partial overlap and cross-linguistic overcorrection in the case of total overlap, whereas previous studies found the reverse pattern (Anderssen et al., 2018; Anderssen & Westergaard, 2020; Kupisch, 2014).

The production patterns displayed by the Arabic/Berber-Dutch group may still be the consequence of having shared syntactic structures. Van Gompel and Arai (2018) suggest that fully identical structures may be shared between languages, whereas similar but non-identical structures are connected. It may be the case that the Arabic/Berber SP is considered "identical" to the Dutch SP, implying a shared structure between languages, whereas the Turkish and PP-medial passive differ too much from each other to be shared (for instance due to word order differences within constituents: the Dutch by-phrase is formed with a preposition, whereas Turkish uses postpositions). For the Arabic/Berber-Dutch group, we can therefore conclude that the SP is either an instance of shared structures, or that there are connected structures, which would mean that partial overlap of the alternatives available between languages does not always lead to cross-linguistic overcorrection.

As we found inhibition effects in the Turkish-Dutch group, which are not compatible with a shared syntax account, we must assume connected representations in this case. So, contrary to what Anderssen and Westergaard (2020) suggest, cross-linguistic overcorrection can occur when there is total overlap of the available syntactic alternatives between languages, i.e., if language B has a parallel alternative for every structure in language A (for a specific alternation).

During the production of Dutch passives, Turkish-Dutch speakers need to inhibit the combinatorial nodes of the high frequent PP-medial passive and the low frequent PP-final passive in Turkish. There is more co-activation of the high frequent structure than for the low frequent structure (cf. Kupisch, 2014) and thus more lateral inhibition for the PP-medial passive than for the PP-final passive during sentence selection in Dutch. During the competition between the syntactic alternatives, the structure that reaches its activation threshold first, is the structure that will be selected. As a consequence of the larger inhibition for the PP-medial passive than for the PP-final passive, Turkish-Dutch bilinguals produce more PP-final passives than non-bilingual Dutch speakers who are not affected by inhibition.

What determines whether one finds cross-lingusitic overcorrection or cross-linguistic influence? Our study differs from the studies on which the predictions of Anderssen and Westergaard (2020) are based (Anderssen et al., 2018; Kupisch, 2014) mainly in two aspects. First, we tested participants in the dominant L2 rather than in the heritage language. Although both the shared syntax account (Bernolet & Hartsuiker, 2018) and the predictions of Anderssen and Westergaard assume bidirectional influences, it may be the case that language dominance or other extralinguistic factors play a role here. Indeed, Kupisch finds cross-linguistic influence rather than cross-linguistic overcorrection in children, suggesting that age or language awareness may be a factor. Brehmer and Sopata

(2021) also find an effect of age, which interacts with whether bilinguals are simultaneous or sequential bilinguals.

Second, we measured the production preferences in a priming experiment rather than in unprimed conditions such as a production elicitation experiment or a corpus study. When Turkish-Dutch participants are primed with the non-frequent PP-medial passive in Dutch, this may lead to a large prediction error (i.e., the inverse preference effect). A large prediction error leads to relatively high levels of activation. As a consequence, the activation level of Turkish PP-medial passive is temporarily higher, which makes it harder to inhibit. A parallel may be found in the study of Kootstra and Şahin (2018), who found cross-linguistic influence in an unprimed experiment and cross-linguistic overcorrection in a primed experiment.

Turning to the structural priming effects, we found PP-medial passive priming but no PP-final passive priming in the Dutch group. These results confirm our assumptions and are in line with Bernolet et al. (2009): PP-final passives and PP-medial passives can be primed separately. Different from Bernolet et al., we did not find PP-final priming. In their study, participants could describe the pictures either with an active or a passive sentence, whereas in our study participants were forced to use a passive sentence. As a consequence, the proportion of PP-final passives relative to the total of responses is much higher in our study than in Bernolet et al. As the PP-final passive is the preferred passive structure, no PP-final passive priming is observed due to the inverse preference effect. We also explored whether the difference in results can be attributed to the fact that we included SP responses and performed a multinomial analysis, whereas Bernolet et al. coded SP responses as "Others" and fit a binomial model to the results. This is not the case: if we omit the SPs from our analyses, we still do not find significant PP-final priming.

With regard to the priming effects in the bilingual groups, we hypothesized that the proportion of PP-final passives would be lower in the Arabic/Berber-Dutch and Turkish-groups and consequently, that the inverse preference effect would be weaker or absent. However, the Turkish-Dutch group did not differ from the Dutch group with regard to the proportion of PP-final passives produced. Similarly, we did not find PP-final priming, which should most likely be attributed to the inverse preference effect as well.

The Arabic/Berber-Dutch group did produce fewer PP-final passives than the Dutch group. Although we did not find significant PP-final passive priming, the proportion of PP-medial responses was lower in the PP-final prime condition than in the base prime condition in the Arabic/Berber-Dutch group. Crucially, this is not exactly the same as PP-final passive priming, since we are dealing with multinomial responses. In a binomial paradigm, a decrease in one target condition automatically means an increase in the other target condition. In our design, a significant decrease in the proportion of PP-medial responses but no significant increase in the proportion of PP-final responses implies that the proportion of SPs is higher after a PP-final prime than after a base prime. In fact, we also observe a decrease in the number of SPs produced after a PP-final prime, which we interpret as a weakened effect of PP-final passive priming. We thus attribute the absence of PP-final passive priming to the inverse preference effect for all three groups and conclude that Arabic/Berber-Dutch and Turkish-Dutch speakers have similar syntactic representations stored for the PP-final passive in Dutch as the Dutch speakers.

We find PP-medial passive priming in all three language groups. This suggests that Arabic/Berber-Dutch and Turkish-Dutch speakers have representations for the infrequent

PP-medial passive in Dutch that are strong enough to be primed in production, even though both groups did produce fewer PP-medial passives than the Dutch group.

To sum up, our data suggest that at least for the Turkish-Dutch bilinguals, the representations of the PP-final and PP-medial passives are connected between languages rather than shared. Although the passive alternation is an alternation where there is total overlap of the alternatives that are available between languages, we find cross-linguistic overcorrection rather than cross-linguistic influence. More research is needed to understand under which circumstances connected syntactic structures lead to inhibition effects. For instance, it is important to test heritage speakers both in their heritage language and in their dominant L2, and to compare the production preferences of simultaneous and sequential bilinguals in primed and unprimed experiments. Computational modelling of bilingual sentence production with different groups of participants may also contribute to the understanding of cross-linguistic influence and cross-linguistic overcorrection of production preferences in bilingual speakers.

2.5 Conclusion

Our results suggest that for the Dutch PP-final and PP-medial passive structures, production preferences but not priming effects are affected by different preferences in heritage languages. The priming effects suggest that heritage speakers seem to have developed syntactic representations for the uncommon Dutch by-phrase-medial passive that are strong enough to be primed in production. We find an instance of cross-linguistic influence in the Arabic/Berber-Dutch group, which may be either due to a shared representation of the SP between Arabic/Berber and Dutch or the outcome of competition between connected representations, of which the mechanisms are not yet fully understood. As for the Turkish-Dutch group, we find cross-linguistic overcorrection, which can probably be attributed to inhibition effects induced by competition between connected representations of the PP-final and the PP-medial passive between Turkish and Dutch. Further research is needed to understand the mechanisms behind the competition taking place between connected structures in different groups of bilingual speakers.

Part II

Comprehension

Chapter 3

The role of modality-specific processing in structural priming in production and comprehension

Structural priming (encountering a syntactic structure facilitates re-processing of that structure) is a well-established method to investigate syntactic processing in production. Although priming effects in comprehension are less robust than those in production, there seems to be a shared mechanism that drives structural priming in both modalities. Still, modality-specific processing may affect the observed magnitude of structural priming, so that a particular prime may have a different effect on a production target than on a comprehension target. To explore whether there is any role of modality-specific aspects in the magnitude of structural priming, we compared priming from comprehension to production (using a written sentence completion task) to priming from comprehension to comprehension (using a visual world paradigm measuring the reaction times of mouse clicks), keeping the stimuli and the prime procedure constant between experiments. In both a between-participants and a within-participants experiment, we detected priming effects in production but not in comprehension, in spite of a large number of observations. In addition, we found evidence that self-priming played a role in the production experiment. Together, these results suggest that due to modality-specific differences in processing, one may find no or weaker priming in comprehension than in production, despite a common mechanism causing priming.

Keywords: structural priming, production, comprehension, modality-specific processing, self-priming

Materials, data, and analyses are available online: `https://osf.io/4ut5e`.

3.1 Introduction

In a typical conversation, participants seamlessly switch between speaking and listening, constantly needing to produce and comprehend language. To do so, one needs to activate linguistic representations of sounds, words, and sentence structures. It seems to be the case that the same mental representations are activated both during production and comprehension (e.g., Giglio et al., 2022; Indefrey, 2018; Pickering & Garrod, 2007, 2013). Indeed, it has been argued that structural priming effects (i.e., processing a structure facilitates the processing of subsequent instances of that structure) in production and in comprehension rely on the same mechanisms (e.g., Pickering et al., 2013). Nevertheless, there are also modality-specific aspects of processing which may affect structural priming effects. The aim of this study is to assess the impact of modality-specific processing on structural priming effects by comparing structural priming from comprehension to production on the one hand to priming from comprehension to comprehension on the other hand. The findings will contribute to the understanding of the shared and distinct mechanisms that are responsible for structural priming effects in production and comprehension.

Processing in production and in comprehension

Production and comprehension are different tasks. A speaker needs to formulate a sentence by selecting lexical items and syntactic structures in order to express their message (cf. Levelt's model of speech production, 1989; Levelt et al., 1999). In comprehension, on the other hand, a listener receives a sentence as an input, and while the sentence is unfolding, the listener needs to parse that sentence in order to construct its meaning through incremental processing.

Despite the different tasks, production and comprehension processing seem to be integrated (cf. Pickering & Garrod, 2013 for an integrated account). More specifically, in both modalities the representations of syntactic structures need to be activated, and it seems that these representations are shared between modalities (e.g., Indefrey, 2018; Kempen et al., 2012). One source of evidence for this claim comes from neuroimaging studies. For instance, Giglio et al. (2022) argued that production and comprehension engage the same representations, but use them differently. In an fMRI study, they showed that the same brain areas (including the left inferior frontal gyrus [LIFG] and the left medial temporal gyrus [LMTG]) are involved in syntactic processing both in production and in comprehension, as activation in these areas increases upon increasing syntactic complexity. Activation of the LIFG was stronger in production than in comprehension, whereas the LTMG was activated more strongly during comprehension than in production. These differences probably need to be attributed to modality-specific aspects of processing.

More neural evidence for common processing between modalities comes from studies using repetition suppression (Grill-Spector et al., 2006). Activation becomes weaker upon the repetition of stimuli. Segaert et al. (2012, 2013) compared the repetition suppression effects both within modalities (production to production and comprehension to comprehension) and between modalities (production to comprehension and comprehension to production). The experiment showed that the LIFG, LMTG, and bilateral supplementary motor area were involved in adaptation to the repetition of syntactic structure both in the case of production and comprehension. Crucially, in both studies the effects were of the same magnitude within and between modalities, suggesting that production and comprehension exploit shared mechanisms. Additional evidence for this claim comes from behavioral studies exploiting the structural priming paradigm, which we will also use in the current study.

Structural priming

The structural priming paradigm can be used in order to investigate mental representations of syntactic structures. Structural priming refers to "the phenomenon by which processing one utterance facilitates processing of another utterance on the basis of a repeated syntactic structure" (Branigan, 2007, p. 1). Bock (1986) found that participants were more likely to describe a target picture depicting a transitive action with a passive sentence rather than an active sentence if they previously listened to and repeated a passive prime sentence, and vice versa. For example, after a passive prime sentence *The building manager was mugged by a gang of teenagers* participants are more likely to produce *The referee was punched by the fans* instead of *The fans punched the referee* than after an active prime sentence.

This effect has since been demonstrated with several structures in multiple languages and occurs in the absence of the repetition of lexical items (see Pickering & Ferreira, 2008). Importantly, in production, verb overlap between the prime and the target sentence induces stronger priming than a condition without verb overlap, which is called the lexical boost effect (Pickering & Branigan, 1998). In addition, while abstract priming is long-lasting, meaning that the priming effect is present even after ten intervening items (Bock & Griffin, 2000), the priming of items with lexical overlap seems to decay quickly (Hartsuiker & Pickering, 2008).

Structural priming can also be found in comprehension. Note that structural priming in comprehension is usually measured in a different way than in production. In production, structural priming studies usually rely on choice data as being collected by means of a picture description task. Priming in comprehension, on the other hand, is generally measured in terms of reaction times or non-behavioral responses, using online measures such as ERP, eye-tracking, and self-paced reading (an exception is Branigan et al., 2005, who collected choice data in comprehension, as described below).

Arai et al. (2007) were among the first to investigate priming in comprehension, testing the priming of the dative alternation in English in an eye-tracking experiment. Participants were asked to read a direct object (DO) sentence (*The pirate will send the princess the necklace*) or a prepositional object (PO) dative prime sentence (*The pirate will send the necklace to the princess*) aloud. Then they listened to a DO or PO target sentence while looking at a scene with pictures of the agent, recipient, and theme. Participants were either exposed to items with verb overlap between prime and target or presented with items that had a different verb in the prime sentence than in the target sentence. When participants in the verb overlap condition had read a DO prime sentence, there were more and longer anticipatory gazes to the recipient object than after a PO prime sentence. Other studies with a visual world paradigm also found priming in a lexical overlap condition; some also found abstract priming in comprehension (e.g., Q. Chen et al., 2013; Thothathiri & Snedeker, 2008a, 2008b). Priming in comprehension has also been shown in ERPs (e.g., Tooley et al., 2009) and self-paced reading (e.g., Traxler & Tooley, 2008).

Structural priming does not only take place within production and comprehension, but can also be found between modalities. Although Bock (1986) primed from production primes to production targets, most production priming studies have participants read or listen to a prime sentence prior to describing a target picture, thus priming from comprehension to production. In addition, Branigan et al. (2005) found structural priming from production to comprehension, although they primed the interpretation of ambiguous prepositional phrases (e.g., *The policeman prodded the doctor with the gun*) rather than two syntactic alternatives. Recently, Litcofsky and van Hell (2019) observed priming from production to comprehension as well as priming from comprehension to production for active and passive sentences. Crucially, priming from comprehension to production seems to be as strong as priming from production to production (Bock et al., 2007); and similarly, priming from production to comprehension seems to be of the same magnitude as priming from comprehension to comprehension (Branigan et al.). This suggests that the mechanisms behind structural priming are at least partially modality-independent.

Mechanisms behind structural priming

Broadly speaking, there are two competing accounts that attempt to explain the occurrence of structural priming effects: the residual activation model (Pickering & Branigan, 1998) and the implicit learning model (Chang et al., 2006). The residual activation model of Pickering and Branigan assumes that lexical lemma nodes are connected to the combinatorial nodes representing syntactic structures. During the processing of a sentence, activation spreads from the lemma nodes to the combinatorial node of the syntactic structure being processed. This leads to a temporarily higher level of activation of that combinatorial node. As a consequence, it becomes easier to reactivate that syntactic structure. The implicit learning model of Chang et al. states that structural priming is the consequence of a form of error-based learning. The model is implemented in a neural network in which concepts are connected to representational units of structures. A strongly connected syntactic structure is activated more easily than a syntactic structure with weaker connections. Importantly, processing a structure strengthens the connections of a syntactic representation. Therefore, priming of a syntactic structure leads to permanent changes in the neural network. This explains why priming effects can still be found after a number of intervening items (Bock & Griffin, 2000).

Hartsuiker and Pickering (2008) argued that both models account for different aspects of structural priming: the short-lived lexical boost effect may be caused by residual activation or explicit memory (Bernolet et al., 2016), whereas the long-lived abstract structural priming effects may reflect implicit learning (see Reitter et al., 2011; Segaert et al., 2011 for hybrid models integrating both residual activation and implicit learning).

Issues with structural priming in comprehension

Despite the parallels regarding structural priming between production and comprehension, there is in fact no consensus on the exact mechanisms that may play a role in comprehension. In contrast to structural priming in production, the results of structural priming experiments in comprehension vary greatly, especially with regard to the lexical boost effect. Crucially, in Arai et al. (2007), no priming effects were found if the verb of the prime was different from the verb of the target sentence. Similar results were obtained by for example Branigan et al. (2005), Q. Chen et al. (2013), Tooley et al. (2009), and Traxler (2015). It has therefore been suggested that, in contrast to priming in production, lexical overlap might be crucial to priming in comprehension (cf. Ledoux et al., 2007).

However, a growing number of studies do find abstract priming in comprehension (e.g., Arai & Mazuka, 2014; Giavazzi et al., 2018; Kim et al., 2014; Pickering et al., 2013; Thothathiri & Snedeker, 2008a, 2008b; Traxler, 2008; Ziegler & Snedeker, 2019). As such, lexical overlap between prime and target does not seem a prerequisite in order to find structural priming in comprehension. Nevertheless, the exact role of lexical overlap is still very unclear. Pickering et al. (2013) found priming in the absence of lexical overlap, but the effects were weaker than for items with lexical overlap. Their results therefore suggest a mechanism that is similar to the lexical boost effect in production. Fine and Jaeger (2016) and Traxler (2008), on the other hand, observed that the priming effects in an experiment without lexical overlap between prime and target did not differ in magnitude from the effects found in a similar experiment with lexical overlap. This is in sharp contrast with studies reporting priming with lexical overlap but no abstract priming at all.

There is also debate on whether there is a rapid decay of the lexical boost effect, as there is in production. In an eye-tracking experiment, Tooley et al. (2014) still find priming effects after three intervening items in the presence of verb overlap between prime and target. Pickering et al. (2013) also find long-lasting priming in comprehension, regardless of whether the verb is repeated between prime and target. Contrary to the lexical boost effect in production, their results suggest that the lexical boost effect in comprehension does not seem to decay during the lag. However, Fine and Jaeger (2016), who also find evidence for long-lasting priming both with and without verb overlap, argue that the lexical boost effect in comprehension is not persistent, similar to the lexical boost effect in production.

Modality-specific processing

The results of priming studies in comprehension thus seem to be less consistent than those of production studies, and it is not clear what causes these inconsistencies in results. The studies vary greatly in design and the structure tested, but even studies with similar set-ups do not always obtain similar results, such as Arai et al. (2007) and Thothathiri and Snedeker (2008a), who both tested datives with an eye-tracking design. Tooley and Traxler (2010) suggest that this variability in results may be due to the small effect size of structural priming in general (in production as well as comprehension), and techniques commonly used in comprehension, such as eye-tracking or self-paced reading, might be less sensitive than the paradigms used in production studies. Some gain with regard to the consistency of results might therefore be obtained by ensuring sufficient statistical power in future studies (although recommendations on the desired sample size are currently only available for production [Mahowald et al., 2016] and not for comprehension).

In addition to methodological considerations, the variability in results may also lie in more general differences between production and comprehension. There are aspects of processing which are specific to either production or comprehension, and this may affect the outcome of priming. While the priming effects themselves are presumably induced during the processing of the primes (through activating a particular node and/or through updating relative weights by means of implicit learning), the effects may manifest differently across modalities due to modality-specific processing.

There are at least two aspects of processing that may play a role in comprehension, but not in production, namely message predictability and syntactic ambiguity resolution. Whereas constructing a full syntactic representation is always needed in production, listeners may not always need to fully build a representation in comprehension. An incomplete representation may suffice to comprehend a sentence (e.g., F. Ferreira et al., 2002; F. Ferreira & Patson, 2007). Ziegler and Snedeker (2019) therefore argue that "message predictability" determines whether abstract priming is found in comprehension. Especially with behavioral measures, studies in which the target sentences are more predictable (e.g., Arai et al., 2007) seem to find weaker abstract structural priming effects than studies in which the target sentences are unpredictable (e.g., Thothathiri & Snedeker, 2008a). X. Chen et al. (2022) provided experimental evidence for a prediction mechanism in priming in comprehension, showing that the priming effects were modulated by verb bias. There was stronger priming for the DO structure after a DO prime with a PO biased verb than after a DO prime with a DO biased verb, which means that there was stronger priming for items with a larger prediction error.

Furthermore, syntactic ambiguity resolution plays a role in comprehension. Instead of selecting and producing a particular syntactic structure, the listener needs to process a structure that may or may not match the expected structure. There are two influential theories on how listeners deal with this. According to garden-path models (Frazier, 1979; Pickering et al., 2000), listeners start with one particular syntactic analysis of the sentence. When during processing of the sentence this analysis does not longer match the input, syntactic reanalysis takes place. According to constraint-based models, listeners simultaneously consider all possible syntactic analyses (e.g., MacDonald et al., 1994; Spivey & Tanenhaus, 1998). These analyses compete with each other in order to determine the most likely syntactic analysis. This process of syntactic ambiguity resolution may have an effect on the magnitude of structural priming effects. Processing may generally take longer for structures with a low token frequency than for structures that occur more frequently in conversational speech corpora (e.g., Arnon & Snider, 2010), either because syntactic reanalysis needs to take place for the less frequent structure, or because competition between structures is stronger for less frequent structures (due to a higher relative weighting of the more frequent structure) than for more frequent structures. Especially in the case of competition between structures, the less frequent structure may, paradoxically, show inhibition effects when primed. When the activation level of the less frequent structure is increased due to priming, the gap in activation level between the more frequent structure and the less frequent structure may decrease, and this may increase competition (cf. Segaert et al., 2011).

This effect of syntactic ambiguity resolution may clash with the inverse preference effect, which follows from the implicit learning mechanism. The inverse preference effect (V. S. Ferreira & Bock, 2006) means that priming is stronger for the less frequent structure than for the more frequent structure due to larger prediction error and thus stronger learning for less frequent structures. It may therefore be hypothesized that structural priming is weaker in comprehension than in production, since priming of the more frequent structure is weak due to small prediction error, and inhibition during the selection stage may compensate for the priming of the less frequent structure.

Finally, there are also potentially relevant aspects of processing that are specific to production. For instance, self-priming may play a role (Jacobs et al., 2019). Participants usually do not produce an equal number of the two syntactic alternants, but they often have a preference for one of the two structures. When they produce a particular structure, they may prime themselves to use that structure in future trials. In this way, self-priming may have a snowball effect on structural priming. Consequently, structural priming effects may be more pronounced in production than in comprehension, in which self-priming does not play a role.

Comparing structural priming across modalities

As mentioned, previous studies showed that priming from production to production is as strong as priming from comprehension to production (Bock et al., 2007), and that priming from production to comprehension is equally strong as priming from comprehension to comprehension (Branigan et al., 2005). This implies that the effect induced by the prime (caused by residual activation and/or implicit learning) is equal across modalities. In both cases, the modality of the target structure was identical. It has not been investigated yet whether priming from comprehension to production and priming from comprehension to comprehension is similar. If there are modality-specific aspects of processing which have an impact on the observed structural priming effects, such effects would be revealed when manipulating the modality of the target structure rather than the modality of the prime structure. In other words, the structural priming effect induced by the prime may have different effects on the processing of the target structure depending on the modality of the target.

There have been some studies comparing structural priming across modalities with different target structures. Tooley and Bock (2014) directly compared structural priming across modalities, priming transitive and ditransitive structures from production to production and from comprehension to comprehension. Participants would read a prime sentence (aloud in the case of production) word by word, and then perform a distractor task. They were then asked to recall the sentence aloud (for the production trials) or to read the sentence by self-paced reading (for the comprehension trials). This procedure was repeated for the target sentence. For the production trials, it was measured whether participants recalled the structure of the presented target sentence correctly or whether they switched to the alternative structure. Self-paced reading times were obtained from the comprehension trials. The standardized z-scores of the priming effects were of the same magnitude in production and in comprehension, independently of whether there was verb overlap between prime and target.

However, there are several potential issues with the study of Tooley and Bock (2014). Not only was the prime procedure not kept constant between experiments, in order to use similar prime procedures while manipulating modality both for prime and target sentences (i.e., priming from production to production and from comprehension to comprehension), the production items in Tooley and Bock did not involve spontaneous production to the same extent as when using picture description tasks. The study of Litcofsky and van Hell (2019) used a traditional picture description task in order to measure priming from comprehension to production. In addition, they performed an EEG experiment to test production to comprehension. For the production primes, participants were instructed to describe a picture, while being forced to start with either the agent (for active prime sentences) or the patient (for passive prime sentences), as indicated by a green border. In this experiment, all target sentences were in the passive condition, so only passive priming was tested. The N400 component was smaller for primed sentences than for unprimed sentences. It must be noted, however, that all critical items had lexical overlap between prime and target. As a result, the structural priming effects may have manifested themselves in an N400 effect (which is associated with lexical processing), while effects of syntactic processing are usually reflected by a P600 effect.

Both Tooley and Bock (2014) and Litcofsky and van Hell (2019) show that it is a viable approach to compare structural priming when the modality of the target structure differs between experiments, despite the differences in measurement types (choice data in production versus reaction times or ERP data in comprehension). Although it is hard to directly compare the magnitude of the priming effects between modalities, due to the different measure types (cf. Litcofsky & van Hell, 2019), there may be a comparison on a more general level, that is, whether there are observable abstract structural priming effects in production targets as well as in comprehension targets when keeping the items and the prime procedure constant.

Current study

The aim of the current study is to assess the impact of modality-specific aspects on structural priming effects by comparing abstract structural priming effects between a comprehension-production and a comprehension-comprehension structural priming study. This comparison seems to be a missing link with regard to comparing structural priming between and within modalities. Priming between modalities seems to be equally strong as priming within modalities if the target structure has the same modality, implying that the priming effect of the prime sentence is similar regardless of the modality in which it is presented. If modality-specific processes such as syntactic ambiguity resolution and self-priming modulate structural priming, it may be the case that priming with comprehension targets is different from priming with production targets, even when the modality of the prime is kept constant. Since it is very likely that message predictability reduces the magnitude of structural priming effects (Ziegler & Snedeker, 2019), we tried to minimize the message predictability of the target sentences in order to assess the potential role of less investigated effects, and the effect of modality-specific processing on structural priming in general. Nevertheless, it is of course hard to say to what extent listeners needed to build a complete syntactic representation in order to process the sentence.

We measured structural priming in production by means of a picture description task, in which participants had to perform a written sentence completion task in order to describe the picture. In the comprehension trials, we measured reaction times to mouse clicks in a visual world paradigm on pictures displaying the entities mentioned in a sentence that is presented audially. To distinguish between the impact of aspects that are specific to comprehension (that is, syntactic ambiguity resolution) and those that are specific to production (self-priming), we conducted a between-participants as well as a within-participants experiment. In the within-participants experiment, participants were exposed to both production and comprehension trials. As a result, there were fewer production trials in total and thus, there were fewer occasions that could lead to self-priming. Also, due to the comprehension trials, participants would listen to more occasions of their dispreferred structure. Therefore the effect of self-priming, if any, may be weaker in the within-participants than in the between-participants experiment.

Importantly, we tested abstract structural priming, that is, there was no lexical overlap between prime and target. Previous studies that manipulated the modality of the target structure only tested for priming in the presence of lexical overlap between prime and target (Litcofsky & van Hell, 2019; Tooley & Bock, 2014). Nevertheless, lexical overlap is the aspect which is most associated with the inconsistent results in priming in comprehension. Overlap between prime and target may modulate the priming effects both in production and comprehension, but is not considered essential in order to evoke structural priming effects. Crucially, it seems that it is in fact the mechanism behind abstract structural priming which is shared between production and comprehension (Segaert et al., 2013). Abstract structural priming engages the mental representations of syntactic structures, and it seems to be the case that production and comprehension rely on the same mental representations during processing (e.g., Giglio et al., 2022; Indefrey, 2018; Pickering & Garrod, 2007, 2013).

3.2 Experiment 1: Between participants

Method

Participants

We recruited 192 participants on the online platform Prolific.[1] In two cases, the data records reported the same Prolific participant ID number in two separate test sessions, suggesting that there were two repeat-participants. We disregarded the data files of their second session. Thus, we analyzed the data of 190 unique participants (75 male, 114 female, 1 other). All participants were monolingual speakers of any variety of English. They were aged between 18 and 35 (mean: 26.7, *SD* 4.8) and did not report any language and/or literacy problems. They were paid for their participation. Half of the participants were assigned to the production experiment and half of the participants completed the comprehension experiment.

Materials

Sixteen dative verbs were selected. Half of the verbs are used with the preposition *to* and half of the verbs are associated with the preposition *for*. With each verb, two prime items and two target items were created. All of the items were used both as a production item and as a comprehension item. The prime and target sentences were recorded by a female native speaker of English.

For each prime item, we constructed four sentences: a DO dative (*he lent the pirate the jar*), a PO dative (*he lent the jar to the pirate*), a DO baseline (*he saw the pirate and the jar*) and a PO baseline (*he saw the jar and the pirate*). In the DO baseline sentences, the animate noun preceded the inanimate noun. In the PO baseline sentences, the inanimate noun was mentioned before the animate noun. We chose to use two separate baselines to avoid priming of the order of clicking on the animate and inanimate entity (see Procedure). We only used the baseline sentences in the production items. For the comprehension items,

[1]The platform of Prolific is designated to recruit high-quality participants for scientific research. It prevents bots or repeat-participants from participating by requiring a unique IP, phone number and bank account for a participant account.

both the prime items and the target items were only presented in the PO dative and the DO dative conditions. By switching the prime and target sentences between lists, the reaction times of sentences in the prime condition served as the baseline to which we compared the reaction times of the sentences in the target condition (see Design).

We also selected four pictures from the MultiPic databank (Duñabeitia et al., 2018) for each prime item: the two pictures matching the beneficiary and the theme in the dative sentences (e.g., *pirate* and *jar*) and two phonological competitors: an animate object starting with the same phoneme as the theme (e.g., *judge* as the competitor for *jar*), and an inanimate object starting with the same phoneme as the beneficiary (e.g., *pencil* as the competitor for *pirate*). In this way, we minimized the message predictability (cf. Ziegler & Snedeker, 2019). We only selected pictures that had a name agreement of 90 percent or higher, i.e., they were identified with a particular noun in at least 90 percent of the responses in the norming study with British English participants. In this way, we ensured that the intended lexical items and thus phonemes were activated upon exposure to the pictures.

The comprehension target items had the same format as the prime items, displaying four pictures for each item (Figure 3.1). For each critical target item, we selected two pictures matching the beneficiary and the theme in the dative sentences and two phonological competitors. A production target item included a verb prompt, a picture of a person and a picture of an object (Figure 3.2). The pictures were again selected from the normed database MultiPic (Duñabeitia et al., 2018). The verb prompt consisted of the subject *he* and the past simple form of the verb (e.g., *he lent*). The picture of the person had the intended role of beneficiary (e.g., *pirate*) and the picture of the object was intended to be used in the theme role (e.g., *jar*).

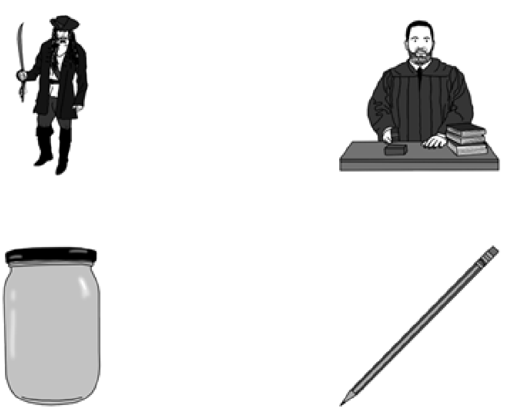


Figure 3.1: Prime item and comprehension target item (*pirate* + *judge* + *jar* + *pencil*).

In addition to the critical items (32 prime items and 32 target items), we constructed 64 filler prime items and 64 filler target items. Filler items included a transitive sentence and a prepositional phrase. Although this structure resembles PO datives, we chose this type of filler items to ensure that participants had to click on two pictures during both critical trials and filler trials. To minimize the expectations of the participants with regard to (the order of) the animacy of the nouns, we ensured that for both the first and the second mentioned noun, it was equally likely to be an animate or an inanimate noun. Therefore

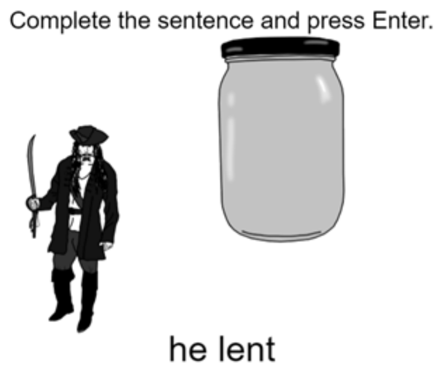


Figure 3.2: Production target item (*he lent* + *pirate* + *jar*).

48 filler sentences included two animate nouns (e.g., *He compared the teacher with the bull*), 48 sentences mentioned two inanimate nouns (e.g., *He left the key in the bus*), 16 sentences mentioned an animate noun and an inanimate noun respectively (e.g., *He met the pilot at the airport*) and finally, 16 sentences mentioned an inanimate noun followed by an animate noun (e.g., *He threw the shoe after the dog*).

For each filler prime/comprehension target item, we selected four pictures too: the two pictures matching the nouns mentioned in the filler sentence, a picture starting with the same phoneme as one of the mentioned nouns (either the first or the second noun) and a non-related picture. The selected pictures always included two animate objects and two inanimate objects. Filler production target items included a transitive verb prompt, a picture of a patient and a picture of an object that was intended to be mentioned in a prepositional phrase, targeting the same type of sentences as the filler prime/comprehension target items.

Design

Each prime item was paired with a target item in order to create a set of 32 critical items. To counterbalance the *to*-verbs and *for*-verbs, half of the prime items with a *to*-verb were paired with an item with a *to*-verb in the set of target items. The other half of the prime items with a *to*-verb were combined with a *for*-verb target item. Similarly, we paired half of the prime items with a *for*-verb with a *for*-verb target item and half of the *for*-verb prime items with a *to*-verb target item. There was no verb or noun overlap between the prime items and the target items.

We constructed eight lists for both the production and the comprehension experiment. Each list started with three filler items. There was always at least one filler item between two critical items. The lists were always displayed in the same order.

For the production experiment, we first created four counterbalanced lists, such that every target item was preceded by a prime from a different prime condition across the lists. In each list the prime items were presented equally often in the four priming conditions. Half of the target items displayed the theme on the left and the beneficiary on the right, and vice-versa in the other half of the target items. We constructed two pseudo-randomized variants of each list in order to counterbalance the position of the theme and the beneficiary, which led to a total of eight lists.

For the comprehension experiment, we started with the construction of two lists, such that every target item occurred in the PO dative condition and in the DO dative condition across the lists. For both lists, we then constructed two versions, to ensure that every target item was preceded by both the PO dative condition and the DO dative condition across the lists. This led to four different combinations of conditions: a PO dative target preceded by a PO dative prime (primed PO condition), a PO dative target preceded by a DO dative prime (antiprimed PO condition), a DO dative target preceded by a DO dative prime (primed DO condition) and a DO dative target preceded by a PO dative prime (antiprimed DO condition). In order to collect baseline, unprimed reaction time measures for each target item, we created another version of each list in which the prime items and the target items of the critical trials were switched. So, for each target sentence we also collected reaction times of that sentence in the unprimed condition. The measures of the prime items thus functioned as a baseline to which we compared the measures of the target items. In total, we constructed eight different lists.

Procedure

The production and the comprehension experiments only differed with regard to the procedure of the target items. The prime procedure was identical between both experiments. Participants were instructed to click on a black dot in the middle of the screen in order to start each trial. They listened to the prime sentence while seeing a display with four pictures. They had to click on the two depicted nouns mentioned in the sentence as fast as possible. After clicking on two of the four pictures, the display was replaced by the target display.

In the production experiment, the target display consisted of two pictures and a verb prompt. Participants were asked to type a completion to the sentence, mentioning the two pictures in their response. A session lasted about 35 minutes.

In the comprehension experiment, the procedure of the target sentences was identical to that of the prime sentences, again starting with clicking on a black dot in the middle of the screen (hence, here the participant could not distinguish between a prime sentence and a target sentence). A session took approximately 25 minutes.

Results

Production

The target responses of the production experiment were coded as DO datives, PO datives, or "Other" responses. Of the 3,008 target responses, 1,001 responses were classified as DO datives (33.3%), 1,194 responses as PO datives (39.7%), and 813 as "Other" (27.0%). We excluded the "Other" responses for further analyses. The responses per priming condition are summarized in Figure 3.3.

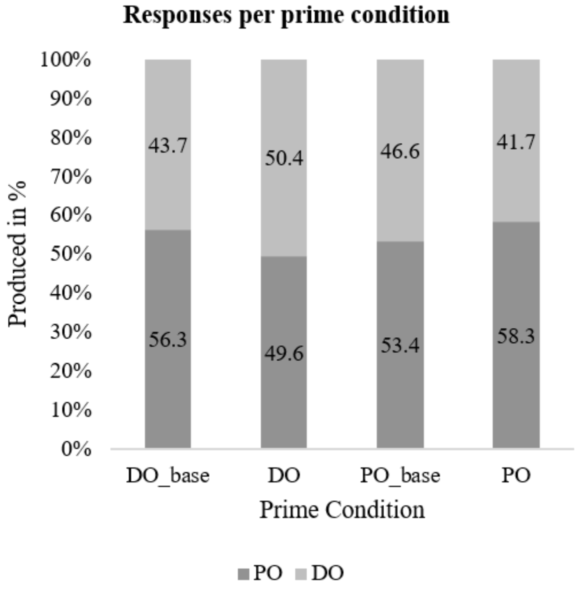


Figure 3.3: Overall production priming effects in Experiment 1.

We analyzed the results of the DO prime condition and the PO prime condition by comparing the proportions of DO responses and PO responses after a DO dative prime to the DO baseline and those after a PO dative prime to the PO baseline.

The target responses were fitted to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer to increase the convergence ability (Powell, 2009). We ran a model with Prime Condition as its fixed factor. We used the PO baseline as the reference level. Conform the maximal random effects structure as proposed by Barr et al. (2013), we started with random slopes for Prime Condition to Participants and Items, and random intercepts for Participants and Items. We simplified the model until there were no convergence or singularity issues. The final model included random intercepts for both Participants and Items, but no random slopes. The model output is reported in the appendix (Table A.3.1).

In order to compare the proportions of DO responses and PO responses after a DO dative prime to the DO baseline and those after a PO dative prime to the PO baseline, we computed post-hoc pairwise comparisons with the package *phia* (De Rosario-Martinez, 2015). The proportion of DO responses was significantly higher after a DO prime than after the DO baseline ($\chi^2(1) = 11.6$, $p < .001$). Similarly, participants tended to produce more PO datives after a PO prime than after a PO baseline item ($\chi^2(1) = 3.3$, $p = 0.07$).

Comprehension

During each prime trial and each target trial, we measured the reaction times of the clicks on the pictures, which led to 6,144 observations. Incorrect clicks (first click: 341 responses, second click: 424 responses) and reaction times more than three standard deviations above the mean (first click: 10.87 seconds, 6 responses, second click: 1.69 seconds, 27 responses) were excluded from further analyses. A total of 5,797 responses were included for analyses of the first trajectory. The analyses of the second trajectory included 5,334 responses. The mean reaction times per structure and per prime type are summarized in Table 3.1.

Table 3.1: Mean reaction times (in seconds) per structure and prime type.

Structure	Type	First click	Second click
DO	base	2.97 (0.78)	0.71 (0.28)
	primed	2.98 (0.73)	0.70 (0.28)
	antiprimed	2.96 (0.70)	0.71 (0.27)
PO	base	2.98 (0.76)	0.72 (0.31)
	primed	2.98 (0.75)	0.72 (0.31)
	antiprimed	2.96 (0.69)	0.72 (0.31)

Note: Standard deviations in parentheses.

The reaction times of the first click were fitted to a linear mixed model (*R*-package *lme4*, D. Bates et al., 2015). We ran a model with Type (base/primed/antiprimed) and Sentence Condition (PO/DO) and their interactions as fixed factors. We started with a maximal random effects structure, as proposed by Barr et al. (2013) and simplified it until convergence by eliminating the random slopes that explained the least variance. The final model included Participant and Item as random intercepts and no random slopes. We used the base level of Type and Sentence Condition PO as the reference level. The model is summarized in Table A.3.2 of the appendix. In order to look at the main effects of priming, we computed post-hoc pairwise comparisons with the package *phia* (De Rosario-Martinez, 2015). The reaction times of a DO sentence did not differ between the baseline condition and the primed condition ($\chi^2(1) = 0.001$, $p = 0.97$). Also, participants responded equally fast to a PO sentence in the baseline condition as in the primed condition ($\chi^2(1) = 0.12$, $p = 0.73$).

We ran a similar model to fit the reaction times of the second click against Type and Sentence Condition and its interactions. We ran a model with a maximal random effects structure, as proposed by Barr et al. (2013). The final model included Participant and Item as random intercepts and no random slopes. As for the model of the first click, we treated the base level of Type and Sentence Condition PO as the reference level. Since the interaction between Type and Sentence Condition was not significant, the reported model is without the interaction. The model is summarized in the appendix in Table A.3.3. Participants responded significantly faster to DO sentences than to PO sentences ($\beta = -0.02$, $SE = 0.006$, $p < .001$). In order to look at the main effects of priming, we computed post-hoc pairwise comparisons with the package *phia* (De Rosario-Martinez, 2015). The reaction times did not differ between the baseline condition and the primed condition ($\chi^2(1) = 0.84$, $p = 0.36$).

3.3 Experiment 2: Within participants

Method

Participants

We recruited another 192 (124 female, 67 male, 1 other, mean age = 27.4 years old, *SD* 4.7) participants on Prolific. They were selected based on the same criteria as for Experiment 1.

Materials

The same materials were used as those of Experiment 1. The items were split in two: half of the items were designated as production trials and half of the items were assigned to be comprehension trials. The critical items were chosen pseudo-randomly, in such a way that all verbs and all conditions were presented equally often in the two modes. As such, each list included 16 production items and 16 comprehension trials, with 4 items for each condition in production (DO dative, PO dative, DO baseline, PO baseline) and 4 items for each combination of conditions in comprehension (DO prime & DO target, PO prime & DO target, PO prime & PO target, DO prime & PO target).

Design

The lists were identical to the lists in Experiment 1. For the production trials, we selected the production items and for the comprehension trials, we used the comprehension items of Experiment 1. So, for this experiment we had eight lists consisting of a mix of production and comprehension trials. In order to collect both production and comprehension data for each item, we created another eight lists in which the modes were reversed: the items that were labelled as production trials were used as comprehension trials, and the items that were designated to be comprehension trials were used as production trials. Hence, we created sixteen mixed lists in total.

Procedure

The procedure was similar to that of Experiment 1. Before each prime trial, participants clicked on a black dot in the middle of the screen in order to start the trial. They listened to the prime sentence while seeing a display with four pictures. They were instructed to click on the two objects mentioned in the sentence as fast as possible. After clicking on two of the four pictures, the display was replaced by the target display, which depended on the modality of the trial. For a comprehension target, the procedure was identical to that of the prime trials. For a production trial, participants saw two pictures and a verb prompt. They were instructed to complete the sentence and type their response. A session took approximately 30 minutes.

Results

Production

We coded the target responses of the production items as DO datives, PO datives, or "Other" responses according to the same criteria as for Experiment 1. Of the 3,072 target responses, 1,065 responses were classified as DO datives (34.7%), 1,324 responses as PO datives (43.1%), and 683 responses as "Other" (22.2%). We excluded the "Other" responses for further analyses. The responses per priming condition are summarized in Figure 3.4.

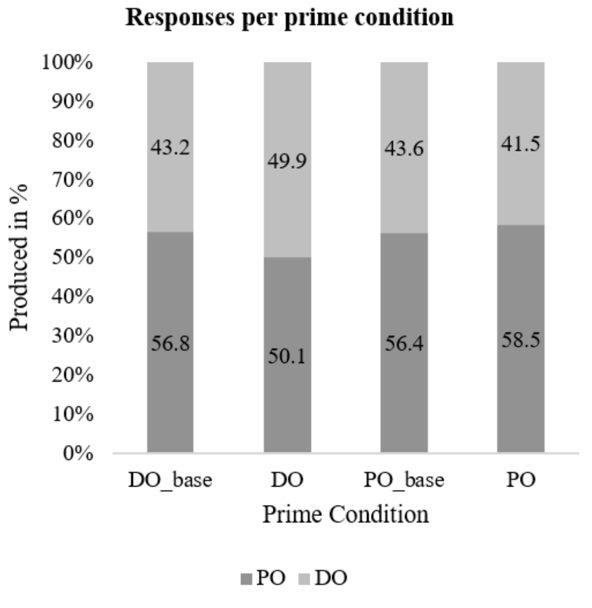


Figure 3.4: Overall production priming effects in Experiment 2.

We analyzed the results in a similar fashion as Experiment 1. We compared the proportions of DO responses and PO responses after a DO dative prime to the DO baseline and those after a PO dative prime to the PO baseline.

The target responses were fitted to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer (Powell, 2009). We ran a model with Prime Condition as a fixed factor. We inserted the PO baseline as the reference level. Conform the maximal random effects structure as proposed by Barr et al. (2013), we added random slopes and random intercepts for Participants and Items. After simplification of the model until there were no issues of convergence or singularity, the final model included random intercepts for Participants and Items, but no random slopes. The final model is reported in the appendix (Table A.3.4).

Post-hoc pairwise comparisons were computed with the package *phia* (De Rosario-Martinez, 2015). According to the post-hoc pairwise comparisons, there was significant DO priming: the proportion of DO responses after a DO dative prime was higher after a

DO prime than after the DO baseline ($\chi^2(1) = 8.5$, $p < .01$). However, participants did not produce more PO datives after a PO prime than after a PO baseline item ($\chi^2(1) = 0.44$, $p = .51$).

Comprehension

For the comprehension trials, we measured the reaction times of the clicks on the pictures. Out of 6,144 observations, we excluded the incorrect clicks (first click: 239, second click: 263) and reaction times higher than three standard deviations (first click; 5.41 seconds, second click: 1.67 seconds) from further analysis. We included 5,837 responses for the analyses of the first click. The analyses of the second trajectory included 5,566 clicks. The mean reaction times per structure and per prime type are summarized in Table 3.2.

Table 3.2: Mean reaction times (in seconds) per structure and prime type.

Structure	Type	First click	Second click
DO	base	3.00 (0.61)	0.73 (0.29)
	primed	2.99 (0.60)	0.71 (0.28)
	antiprimed	3.02 (0.61)	0.71 (0.29)
PO	base	2.99 (0.61)	0.73 (0.30)
	primed	3.00 (0.62)	0.74 (0.31)
	antiprimed	3.00 (0.60)	0.72 (0.30)

Note: Standard deviations in parentheses.

We fitted the reaction times of the first click to a linear mixed model (*R*-package *lme4*, D. Bates et al., 2015). We ran a model with Type (base/primed/antiprimed), Sentence Condition (PO/DO) and its interactions as fixed factors. We started with a maximal random effects structure and simplified it until convergence (Barr et al., 2013). The final model included Participant and Item as random intercepts and no random slopes. The reference level of the model was the base level of Type and Sentence Condition PO. The model is summarized in Table A.3.5 of the appendix. In order to look at the main effects of priming, we computed post-hoc pairwise comparisons with the package *phia* (De Rosario-Martinez, 2015). The reaction times of a DO sentence were not different between the baseline condition and the primed condition ($\chi^2(1) = 0.35$, $p = 0.56$). Also, participants responded equally fast to a PO sentence in the baseline condition as in the primed condition ($\chi^2(1) = 0.17$, $p = 0.68$).

We ran a similar model to fit the reaction times of the second click against Type, Sentence Condition, and their interaction. We ran a model with a maximal random effects structure, as proposed by Barr et al. (2013). The final model included Participant and Item as random intercepts and no random slopes. As for the model of the first click, we treated the base level of Type and Sentence Condition PO as the reference level. Since the interaction between Type and Sentence Condition was not significant, the final model does not include the interaction. The model output is provided in the appendix in Table A.3.6. Participants responded faster to a DO sentence than to a PO sentence ($\beta = -0.01$, $SE = 0.006$, $p < .05$). In order to evaluate the main effects of priming, we computed post-hoc pairwise comparisons with the package *phia* (De Rosario-Martinez, 2015). The reaction times did not differ between the baseline condition and the primed condition ($\chi^2 2(1) = 0.14$, $p = 0.70$).

3.4 Analysis of self-priming

In the production experiment between participants, we found marginally significant priming of the PO structure, whereas in the experiment within participants, we did not find a significant effect of the PO prime on participants' responses. In order to assess to what extent self-priming might have played a role in the different experiments, we performed a joint analysis in which we added a variable for the running count of the structures. So, in order to assess self-priming of the PO, for each trial we counted how many instances of the PO structure the participant had produced so far, and similarly, we counted for each trial how many times participants produced a DO structure up until that point in order to measure self-priming of the DO.

For the DO structure, we fitted the target responses to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer (Powell, 2009) with Prime Condition, Running Count of the DO, and Experiment (between or within participants) and their interactions as fixed factors. The DO baseline was inserted as the reference level for the Prime Condition, and the between-participants experiment was treated as the reference level for Experiment. We added random slopes and random intercepts for Participants and Items, and simplified the model until convergence without singularity issues. The final model did not include any random slopes, and only had a random intercept for Participants. We then removed the non-significant interactions one by one using the *drop1* function of the *jtools* package (Long, 2022). The final model only includes an interaction between the Running Count and Experiment. The model output is reported in the appendix (Table A.3.7).

The results still showed a significant priming effect of DO primes: participants produced more DO structures immediately after a DO prime ($\beta = 0.34$, $SE = 0.09$, $p < .001$). Compared to the reference level, there was only a trend towards self-priming (i.e., the running count of the DO) ($\beta = 0.02$, $SE = 0.01$, $p < .1$). There was a significant interaction between the Running Count and Experiment, showing that the effect of self-priming was larger in the experiment within participants ($\beta = 0.12$, $SE = 0.02$, $p < .001$) than in the experiment between participants.

We performed a similar analysis for the PO structure. In the generalized linear mixed model, we added Prime Condition, Running Count of the PO, and Experiment and their interactions as fixed factors. This time, the PO baseline was inserted as the reference level for Prime Condition. We started from a full model with random slopes and random intercepts for Participants and Items, and simplified the model until there were no convergence and singularity issues. The final model included random intercepts for Participant and Item, and no random slopes. We removed the non-significant interactions one by one by means of the *drop1* function of the *jtools* package (Long, 2022). In the final model, we only included an interaction between the Running Count and Experiment. The model output is reported in Table A.3.8 of the appendix.

The results only showed a marginally significant effect of the PO prime structure (β =

-0.20, $SE = 0.10$, $p < .1$) (which may have been relatively weak due to the NP + PP structure of the fillers). However, the Running Count was significant (β = -0.37, $SE = 0.02$, $p < .001$), showing that participants produced more PO sentences as they have been producing more PO structures. The self-priming effect was stronger in the experiment within participants than in the experiment between participants (β = -0.33, $SE = 0.03$, $p < .001$).

3.5 Discussion

The aim of our study was to study the impact of modality-specific processing on structural priming by comparing priming from comprehension to production on the one hand to priming from comprehension to comprehension on the other hand, keeping the items and the prime procedure constant. We found significant DO priming in the comprehension to production experiment, both between participants and within participants. Despite a large number of observations, we do not find any significant priming effects in the comprehension to comprehension experiment. Our results therefore suggest that production targets may elicit larger structural priming effects than comprehension targets, even if the prime procedure and the critical items are identical.

Neither Experiment 1 nor Experiment 2 showed significant priming effects in comprehension, even though we tried to minimize message predictability (Ziegler & Snedeker, 2019). It might be the case that the design we employed is not sensitive enough to pick up on any structural priming effects. Nevertheless, the method itself is able to detect differences in reaction times, given that overall, participants reacted faster to DO sentences than to PO sentences. Participants may have responded faster to DO sentences than to PO sentences, because for DO sentences they had to click on an animate noun instead of an inanimate noun. Animate nouns may be conceptually more accessible than inanimate nouns (Bock & Warren, 1985). In addition, the difference between DO and PO sentences may partially reflect differences in the duration of the sentence rather than differences in processing speed. The recorded sentences had a duration of 2 to 3 seconds. Despite the fact that participants were instructed to click on the pictures mentioned in the sentence as fast as possible, the average clicking time of 2.97 seconds suggests that participants listened to the entire sentence before clicking on the two pictures. Since DO dative sentences are inherently shorter than sentences with a PO dative, participants may have responded faster to DO dative sentences than to PO dative sentences.

The observation that participants reacted relatively late may also account for the absence of priming effects in comprehension. The time between the moment that participants clicked on the pictures and the moment at which they encountered the disambiguating area may have been too long to detect facilitated processing by repetition of the syntactic structure. Future studies should therefore involve methods which measure effectively during the disambiguating area. These measurements may include eye-tracking or self-paced reading, as has been done in previous studies, or a refined version of the current paradigm, for example by forcing participants to respond faster (although note that it will probably still take longer for a participant to move the mouse than to move their eyes to the relevant picture).

Still, priming effects in production are very robust and are found in a great variety of experimental designs, and this seems to be different for comprehension. Our results therefore do suggest that structural priming in comprehension is harder to detect than priming in production, even when the critical items used and the prime procedure are the same. On the basis of our results, we cannot conclude whether this is due to the design, due to the different measurement types, or due to modality-specific aspects of comprehension. Therefore we suggest that future studies collect both choice data as well as reaction times both for production and comprehension. In this way, the sensitivity of the research method may possibly be disentangled better from the modality-specific effects.

Our aim was to investigate the role of modality-specific processing on structural priming effects in production and in comprehension. Because of the overall absence of structural priming in comprehension, which may be due to the experimental set-up, our results are inconclusive with regard to the role of aspects that are specific to comprehension. We also tapped into the role of modality-specific aspects of production, and more specifically the role of self-priming. We had hypothesized that self-priming may play a larger role in the experiment between participants than in the experiment within participants. In the within-participants experiment, participants were exposed to fewer production trials and consequently, there were fewer occasions that could lead to self-priming. Instead, we found that self-priming is stronger in the within- participants experiment than in the between-participants experiments. The reason for this may be that the immediate prime condition plays a smaller role in the within-participants experiment than in the between-participants experiment due to overall smaller priming effects in the within-participants experiment (which may be caused by unpredictable task-switching). If participants are primed less by the prime condition (and this was especially the case for the PO structure), their personal preference, and thus self-priming, may play a larger role in determining whether they use a DO or a PO.

3.6 Conclusion

Our results show priming from comprehension to production, but no priming from comprehension to comprehension. The absence of priming effects in the latter case may be due to the experimental design, in which participants' reaction came relatively late after the disambiguating area. Because of the complete lack of priming effects in comprehension, we cannot conclude whether the results were weaker with comprehension targets than with production targets due to methodological issues or due to modality-specific aspects of processing. Nevertheless, we found evidence for self-priming in the production experiment, suggesting that modality-specific aspects may indeed modulate structural priming effects. Future studies should collect both choice data as well as reaction times both for production and comprehension in a between participants design in order to further minimize methodological differences between modalities.

Chapter 4

Two sides of the same coin? Comparing structural priming between production and comprehension in choice data and in reaction times

Although structural priming seems to rely on the same mechanisms in production and comprehension, effects are not always consistent between modalities. Methodological differences often result in different data types, namely choice data in production and reaction time data in comprehension. In a structural priming experiment with English ditransitives, we collected choice data and reaction time data in both modalities. The choice data showed priming of the DO and PO dative. The reaction times revealed priming of the PO dative. In production, PO targets were chosen faster after a PO prime than after a baseline prime. In comprehension, DO targets were read slower after a PO prime than after a baseline prime. This result can be explained from competition between alternatives during structure selection. Priming leads to facilitation of the primed structure or inhibition of the opposite structure depending on the relative frequency of structures, which may differ across modalities.

Keywords: structural priming, maze task paradigm, production, comprehension

Materials, data, and analyses are available online: `https://osf.io/c6uyz`.

4.1 Introduction

When people listen to language, they are confronted with the choices that speakers make with regard to formulating their message. Listeners seem to handle this job by making predictions, which implies that they need to adapt when the input does not meet their expectations. These predictions and adaptations are reflected in structural priming effects. Structural priming, which is "the phenomenon by which processing one utterance facilitates processing of another utterance on the basis of a repeated syntactic structure" (Branigan, 2007, p.1), is often used as a method to investigate mental representations of syntactic structures.

Bock (1986) found that participants who were exposed to a passive prime sentence (e.g., *The building manager was mugged by a gang of teenagers*) were more likely to produce a passive sentence (e.g., *The referee was punched by the fans*) rather than an active sentence (e.g., *The fans punched the referee*) when describing a target picture than when they were presented with an active prime sentence. The structural priming effect has proven to be a robust effect, and was found with a range of syntactic structures, in multiple languages and with different populations, including children and L2 learners (see Mahowald et al., 2016 for a meta-analysis). Structural priming has also been observed in comprehension (see Tooley & Traxler, 2010 for a review), for instance in self-paced reading, eye-tracking, or fMRI experiments. However, it remains unclear whether the mechanisms underlying priming in production and comprehension are the same. The priming effects seem to be differentially sensitive to certain effects such as the lexical boost effect (structural priming is stronger when the verb is repeated between prime and target than when the verbs are different [Pickering & Branigan, 1998] and the inverse preference effect [priming effects are stronger for less frequent structures than for more frequent structures, V. S. Ferreira & Bock, 2006]). These differences are difficult to interpret though, because production and comprehension paradigms typically rely on different methods and hence types of data (e.g., production choices and reaction times respectively).

In the current study, we will therefore test a new method to study structural priming in comprehension, namely an adapted version of the maze task paradigm. This paradigm allows us to collect reaction times as well as choice data in comprehension. We compare the results to priming effects in a production experiment, in which we measure choices and reaction times as well by measuring both production choices and onset latencies.

Structural priming in comprehension

Structural priming effects are not only found in production, but also in comprehension. For instance, Arai et al. (2007) measured eye movements during a comprehension experiment with English ditransitives. After reading a prime sentence, participants listened to direct object (DO) sentences (e.g., *The pirate will send the princess the necklace*) and prepositional object (PO) sentences (e.g., *The pirate will send the necklace to the princess*) while looking at a scene depicting the agent, recipient, and the object. Participants looked longer and earlier to the recipient after a DO prime than after a PO prime if there was lexical overlap between prime and target.

It has been argued that structural priming in comprehension relies on the same mechanisms as structural priming in production (Pickering et al., 2013; Tooley & Bock, 2014). Tooley and Bock compared priming effects between production (as observed in a sentence recall task) and comprehension (as measured in a self-paced reading task) by calculating standardized *z*-scores for both modalities and found that the priming effects were of similar magnitude. Some neurological evidence for a common mechanism comes from Segaert et al. (2012, 2013), who showed in an fMRI experiment that the same brain areas are involved during structural priming in the different modalities, and that the effects were of the same magnitude in production and in comprehension.

If the mechanisms behind structural priming are indeed the same for comprehension and for production, we may also expect that factors modulating the magnitude of structural priming effects, such as the lexical boost effect and the inverse preference effect, behave similarly in the two modalities. Indeed, according to Tooley and Traxler (2018) structural priming in comprehension involves a short-lived lexical boost effect (due to residual activation and/or explicit memory of the prime sentence) and long-lasting priming effects (due to implicit learning) too. In addition, X. Chen et al. (2022) found that priming effects in comprehension are modulated by verb bias (as they are in production, cf. Bernolet & Hartsuiker, 2010), which points towards an implicit learning mechanism in comprehension as well.

Nevertheless, results have been very variable in comprehension, especially with regard to lexical overlap. Some studies only find priming effects in comprehension in the presence of lexical overlap between prime and target (e.g., Arai et al., 2007; Branigan et al., 2005; Q. Chen et al., 2013; Tooley et al., 2009; Traxler, 2015). Other studies do observe structural priming effects when there is no lexical overlap between prime and target (e.g., Arai & Mazuka, 2014; Giavazzi et al., 2018; Kim et al., 2014; Pickering et al., 2013; Thothathiri & Snedeker, 2008a, 2008b; Tooley & Bock, 2014; Ziegler & Snedeker, 2019). Importantly, some studies that find structural priming in comprehension both with and without lexical overlap do not find any difference in magnitude, suggesting that structural priming in comprehension is independent of lexical overlap (Fine & Jaeger, 2016; Traxler, 2008). In addition, it has been argued that the lexical boost effect is long-lasting rather than short-lived in comprehension (Pickering et al., 2013; Tooley et al., 2014), whereas Fine and Jaeger (2016) and Tooley and Traxler (2018) find that the lexical boost effect decays during intervening items.

It has been suggested that the results of structural priming experiments in comprehension are much more variable than those in production experiments due to intrinsic differences between production and comprehension (Ziegler & Snedeker, 2019). In comprehension, listeners do not need to build a complete syntactic representation of the sentence, whereas in production, a full representation is always required. As such, priming effects may only be found if the "message predictability" of the target sentences is low enough, so that the listener has to build a full representation of the structure. Alternatively, methodological differences may underlie the more variable results in comprehension. The methods to measure structural priming effects in comprehension are less standardized and perhaps also less sensitive than the methods used in production. Whereas structural priming in production is typically investigated by means of a picture description task, there are several techniques used to measure structural priming in comprehension, mainly self-paced reading, eye-tracking, and electroencephalography. Tooley and Traxler (2018) argue that these methods are less sensitive than the paradigms in production studies. Additionally, for structural priming experiments in production, there is a meta-analysis available, which indicates the sample size needed to obtain sufficient statistical power to establish reliable structural priming effects (Mahowald et al., 2016). The desired sample size for structural priming experiments in comprehension is currently unknown. As a result, comprehension studies are potentially underpowered. This might explain the variability in results with regard to the lexical boost effect, which is an interaction between priming and lexical overlap. If there is a true interaction between priming and lexical overlap in comprehension as there is in production, underpowered studies may either only find statistical evidence for priming in the strongest category, observing priming in the presence of lexical overlap only (e.g., Arai et al., 2007; Branigan et al., 2005; Q. Chen et al., 2013; Tooley et al., 2009; Traxler, 2015), or may not be able to observe a statistically significant interaction, reporting equally strong priming both with and without lexical overlap (e.g., Fine & Jaeger, 2016; Traxler, 2008).

Measurement types

The methodological differences between structural priming experiment in production and comprehension also result in different types of data. Structural priming in production is typically investigated by means of a picture description task. The picture can be described by either of the two alternating structures (e.g., a direct object dative or a prepositional object dative), resulting in (offline) choice data. Structural priming effects in comprehension, on the other hand, are usually measured in terms of online data, such as reaction times. Nevertheless, it is possible to obtain reaction time data in production experiments and choice data in comprehension. For instance, Corley and Scheepers (2002) and Segaert et al. (2011) measured reaction times in addition to choice data in production, whereas Giavazzi et al. (2018) and Kim et al. (2014) collected choice data as well as reaction times in comprehension.

These four studies use a range of methods to obtain both types of data. Corley and Scheepers (2002) investigated the priming of ditransitives in English in a web-based sentence completion experiment. For the prime sentence, participants were asked to complete a given sentence fragment containing a subject, a ditransitive verb and a noun following the verb, which was either animate to elicit a double object dative (e.g., *The bank manager handed the customer*) or inanimate, in order to elicit a prepositional object dative (e.g., *The bank manager handed the cheque*). The target sentence fragment only included a subject and a ditransitive verb (e.g., *The junior surgeon gave*). In addition to the choice data (whether participants produced a double object dative or a prepositional object dative), they also measured the time at which the participant started typing the response. Participants produced more prepositional object datives after a prepositional object prime than after a double object dative prime, and tended to produce more double object datives after a double object dative prime than after a prepositional object prime. Interestingly, participants also started to produce the target sentence completion significantly earlier in the primed condition. Segaert et al. (2011) obtained similar results with active and passive sentences in a spoken production experiment, measuring the speech onset of the target sentence. Importantly, Segaert et al. found that the inverse preference effect was only found in the choice data. Instead, the onset latencies revealed a "preference effect": here, priming was stronger for the more frequent alternative than for the less frequent alternative. This reemphasizes that differences in priming effects between production and comprehension may be related to the nature of the data collected (e.g., choice vs. latency) rather than modality-related differences per se.

In order to collect choice data in comprehension, Kim et al. (2014) investigated the priming of ambiguously attached prepositional phrases (e.g., *The detective noticed* [*the mirror on the wall*] [*with the crack*] / *the mirror on* [*the wall with the crack*]). In a self-paced reading task, participants read a non-ambiguous prime with either a high attachment or a low attachment sentence followed by an ambiguous target sentence. Participants were then asked whether they interpreted the target sentence with high or low attachment. Sentences were more often interpreted with low attachment after a low attachment prime. The choice data were then used to calculate the priming effects in terms of reaction times: Kim et al. compared the reaction times of the target sentences of which participants chose the same interpretation as that of the prime sentence to the reaction times of the target sentences which participants chose to interpret the target sentence differently from the preceding prime sentence and found lower reaction times in the primed condition than in the unprimed condition.

In Kim et al.'s (2014) design, structural priming in comprehension cannot be measured with the same syntactic structures as in production if one wants to obtain choice data, as this design requires the target sentence to be ambiguous. In other words, one sentence structure (e.g., one containing an ambiguously attached prepositional phrase) corresponds to multiple scenes (high attachment or low attachment). By contrast, in a picture description task as used for priming in production, one scene (e.g., a transferring action) corresponds to multiple sentence structures (a double object dative or a prepositional object dative). Giavazzi et al. (2018) took a different approach to obtain choice data in comprehension, using non-ambiguous target sentences. They showed participants pictures with a transitive action, for example a penguin painting a dolphin. In an active sentence, *the penguin* is mentioned first, whereas the passive counterpart starts with *the dolphin*. Participants would listen to a sentence either matching the picture (*The penguin paints the dolphin*) or to a sentence with the reversed action (*The dolphin paints the penguin*). Then they had to choose under time pressure whether the picture matched the sentence. As the entities involved always matched between the picture and the sentence, participants needed to make their decision based on the syntactic structure of the sentence. This procedure was the same for prime sentences and for target sentences. For the target sentences, it was measured whether participants correctly identified the match or mismatch. In addition, their reaction times were measured. Responses were more accurate and faster in the primed condition than in the unprimed condition. With this paradigm, choice data and reaction times can be collected simultaneously. A potential drawback is that the "reversed action" is only available with a limited set of syntactic alternations. For instance, in ditransitive sentences such as *The waitress passes the* [*boxer the cake*] / [*the cake to the boxer*], there is no reverse action, since the object is inanimate and the object is animate. As we will explain below, an alternative way to simultaneously collect choice data and reaction times may be to use an adapted version of the maze task paradigm.

The maze task paradigm

The maze task paradigm was designed as an alternative to self-paced reading (Forster et al., 2009) and is a method to study incremental sentence processing. Similar to self-paced reading, participants read sentences word by word. However, instead of just reading the word and pressing a button in order to proceed, participants complete a forced choice task for each word. For each word of the sentence, participants are presented with a distractor word alongside the correct word (see Figure 4.1). Participants choose which word will be the continuation of the sentence by pressing a button, and their reaction time is measured. It seems that the effects found in experiments exploiting the maze task are generally larger and have a smaller spillover (i.e., effects on the reaction times of words following after the word for which effects are predicted) than effects found with self-paced reading (N. Witzel et al., 2012) and with eye-tracking (J. Witzel & Forster, 2014). There are different variants of the maze task, according to the type of distractor words used in the task. Distractor words may be pseudowords (lexicality maze), an ungrammatical continuation of the sentence (grammaticality maze), or a statistically unlikely word (automated maze). The latter was developed by Boyce et al. (2020), and it automatically generates distractor words based on a machine-learning language model. It is essentially a variant of the grammaticality maze task, as it uses existing words; often the generated distractor words will be ungrammatical in the context of the critical sentence as well.

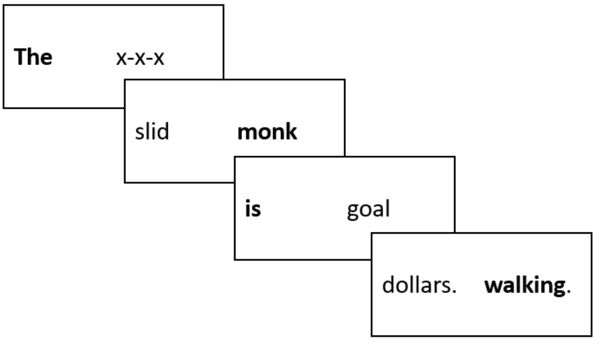


Figure 4.1: Example of maze task paradigm (with distractor words generated in an automated way).

Current study

In our study, we investigated whether the (automated) maze task paradigm is a suitable task to detect structural priming in comprehension. As far as we know, we are the first to exploit the maze task paradigm in a structural priming experiment. The maze task paradigm has many practical advantages over methods such as eye-tracking, EEG, and fMRI, as it is less expensive, requires less materials and is easier to implement in a web-based experiment (cf. Boyce et al., 2020). In addition, it may be more sensitive than self-paced reading and as such, may be able to detect structural priming effects that are hard to find with self-paced reading. Our main reason to use the maze task paradigm, however, was that an adapted version of the maze task allows us to simultaneously collect choice data and reaction time data.

In our comprehension experiment, participants read prime and target sentences. The prime sentences were either DO or a PO sentences, or a baseline sentence (cf. Bernolet et al., 2009 for the use of a baseline as a reference level relative to which priming effects are measured). We measured the reaction times to DO and PO target sentences after the different prime conditions. In addition, in the free-choice version, participants chose whether the target sentence they read was a DO or a PO sentence. In order to make a choice between two syntactic alternatives, participants choose between two plausible words, each of which implies a different continuation of the sentence, corresponding to one of the syntactic alternatives. That is, PO and DO sentences have an identical beginning of the sentence and only start to deviate from each other halfway the sentence (e.g., *The waitress passes the* [**boxer** *the cake*] / [**cake** *to the boxer*]). At this critical point, DO sentences have an animate noun to indicate the recipient (*boxer*), whereas PO sentences mention an inanimate noun, which is the object (*cake*). Participants chose between *boxer* and *cake* to determine whether they will read a DO or a PO sentence. For all other words of the sentence, they chose between the correct word and a distractor word.[1]

[1]It may be possible that this choice between two plausible continuations of the sentence draws the attention of participants and reveals the purpose of the experiment. Therefore only half of the participants were assigned to the free-choice version, and we inserted the version (no-choice vs. free-choice) as a fixed effect in our analyses. In none of the analyses there was an interaction between the priming effect and the version. In addition, after the experiment participants were questioned on their thoughts on the purpose of the experiment. The majority

We expected the participants to more often choose to read a DO target after a DO prime than after a baseline prime, and similarly, that they more often choose a PO target after a PO prime as compared to a baseline prime. In addition, we expected lower reaction times for a DO target after a DO prime than after a baseline prime and lower reaction times for a PO target after a PO prime than after a baseline prime. Choices may be more sensitive to priming than reaction times. Therefore the effects, if any, may be larger in the choice data than in the reaction time data.

The second aim of our study was to investigate whether any observed priming effects in comprehension are similar to the effects found in production. Corley and Scheepers (2002), Giavazzi et al. (2018), Kim et al. (2014), and Segaert et al. (2011) all suggest that choice data and timing data correlate within a modality. That is, the studies find evidence for structural priming with either measure, but they are limited to either production or comprehension. We aimed to test whether choice data and timing data are similar between modalities, so, whether the priming effects obtained by choice data or reaction times in comprehension are similar to the priming effects found in production. In this way, we might be better able to attribute any differences in priming effects (for instance with regard to the lexical boost effect) between the two modalities to either differences in measurement types or differential processing.

For this purpose, we also tested the critical items of the comprehension experiment in a production experiment. After reading a prime sentence, participants were asked to describe a picture by typing a completion to a sentence prompt. The prompt would include the sentence until the critical point (e.g., *The waitress passes the*). We measured the onset latencies of the critical word, which is the reaction time of the first keystroke. Participants would either write [*boxer the cake*] or [*cake to the boxer*], thus choosing to produce either a DO or a PO sentence. If priming effects are similar between production and comprehension, the effects in the choice data and the reaction time data should be similar across modalities. This hypothesis includes the lexical boost effect: we expected that any observed priming effects will be stronger if there is verb overlap between prime and target than if the verb of the prime sentence is different from the verb of the target sentence.

4.2 Experiment 1: Comprehension experiment

Method

Participants

For the four different versions of our experiment (with and without a choice between structures in the target sentences, and with and without verb overlap between prime and target), we recruited 96 participants each. In total, we tested 384 participants (216 female, 157 male, 11 other) with American English as their L1 on the online platform of Prolific.

of participants thought the experiment was on lexical effects, such as measuring their vocabulary size. Therefore it seems that the choice was not very obvious for participants.

Participants were aged between 18 and 83[2] (mean: 38.8, *SD*: 13.9) and did not report any literacy or language problems. They were paid for their participation in the experiment.

Materials and design

We constructed 36 prime-target pairs with ditransitive sentences. We created prime sentences in three prime conditions: a baseline condition, which was an intransitive sentence (e.g., *The monk is laughing*), a double object (DO) dative sentence (e.g., *The waitress passes the boxer the cake*), and a prepositional object (PO) dative sentence (e.g., *The waitress passes the cake to the boxer*). The target sentences came in two variants, namely in the DO condition and in the PO condition. In addition, we created 144 filler sentences and 4 practice sentences, which were intransitive sentences of the same format as the baseline sentences, and transitive sentences (e.g., *The witch tickles the dancer*).

For each word of the sentence, participants would need to decide between the correct word and a distractor word (i.e., a word that is an implausible continuation of the sentence). For both the prime and the target sentences and for the fillers, we created distractor words using the automated MAZE tool (Boyce et al., 2020), which is available online (`https://vboyce.github.io/Maze/`). Each sentence started with *The* on the left and *x-x-x* on the right. Other than that, we distributed the correct words and distractor words in such a way that half of the correct words appeared on the left of the screen and half of the correct words appeared on the right of the screen.

We created four different versions. In two versions, the verb of the prime sentence was identical to the verb of the target sentence. In the other two versions, there was no verb overlap between prime and target. In addition, we manipulated whether participants would be able to choose a syntactic structure while reading the target sentences. In the version with a choice, participants had to choose between an animate noun (for a DO sentence) and an inanimate noun (for a PO sentence) on the critical point. To illustrate, for the sentence *The waitress passes the* [*boxer the cake*] / [*cake to the boxer*], participants in the free-choice condition had to choose between *boxer* and *cake*. The sentence would then continue as either a DO or a PO, according to their choice. In the version without a choice, participants would read a DO target sentence in half of the critical items and a PO target sentence in the other half of critical sentences.

For the two versions (with and without verb overlap) with a choice between DO and PO target sentences, we constructed three lists each, so that every item was preceded by a prime from a different prime condition across the lists. Within each list, we presented the prime sentences equally often in the three priming conditions. For the two versions without a choice, we created six lists each, such that every item appeared as a DO and as a PO after the three prime conditions across the lists. Participants were pseudo-randomly assigned to one of the lists; we tested 96 participants per version.

The experiment was programmed in PsychoPy (Peirce et al., 2019) and hosted on its online platform Pavlovia.

[2]The tested group has a relatively wide age range, which may have led to more variability in the data than in other experiments that are otherwise comparable. We added random intercepts for participants in our statistical models to minimize the impact of interspeaker variability on our statistical results.

Procedure

During the priming experiment, participants would read the prime and target sentences word by word. For each word, two words were presented on the screen: the correct word and the distractor word. To choose for the word on the left of the screen, participants had to press the key A, while they had to press the key L in order to choose for the word on the right of the screen. If they chose the distractor word instead of the correct word, they were prompted to try again, until they chose the correct word. This is different from most previous studies using the maze paradigm, where sentences would be cut off after an error. We chose to use the forgiving maze paradigm (cf. Boyce et al., 2020), because we wanted to ensure that participants would read the entire prime sentence before reading the target sentence to preserve the balance of the three prime conditions. Immediately after the priming experiment, participants performed the LexTALE test, which is a short yes/no-vocabulary test (Lemhöfer & Broersma, 2012), and completed a survey on their demographic background (gender, age, L2 knowledge). We used this information to confirm whether participants met the selection criteria. A session took about 30 minutes.

Coding

We registered the timing of the key press for the choice of each word. In addition, for the participants assigned to the choice version, we recorded their choice for either the DO or the PO for each target sentence. To calculate the priming effect, we compared the proportion of DO or PO responses after a DO or PO prime respectively to the proportion of DO or PO responses after a baseline prime.

For the analysis of the reaction times, we first removed the outliers for each word, defined as reaction times deviating more than three times the standard deviation from the mean. We measured the reaction times for the disambiguating word at the critical point (which is the animate or the inanimate noun) as well as for each word of the spillover area, which is the remainder of the sentence. The spillover area comprised two words for DO sentences (e.g., *The waitress passes the* **boxer** [*the cake*]) and three words for PO sentences (e.g., *The waitress passes the* **cake** [*to the boxer*]). The reason for this way of measuring priming effects is that structural priming effects in comprehension have been reported in the disambiguating area as well as in the spillover region (e.g., Tooley et al., 2014; Wei et al., 2016). Since the data are not normally distributed, we log-transformed all the reaction times (log10(x)), after which the Q-Q plot of the plotted reaction times showed a more normal distribution. We calculated the priming effect by comparing the reaction times after a DO or PO prime to the reaction times after a baseline prime.

Results

Choice data

With regard to the choice data, we collected a total of 6,912 observations. Participants chose to read the PO structure in 3,484 of the cases (50.4%) and the DO structure in 3,428 of the items (49.6%). Figure 4.2 shows the choice for a DO or a PO structure per prime condition.

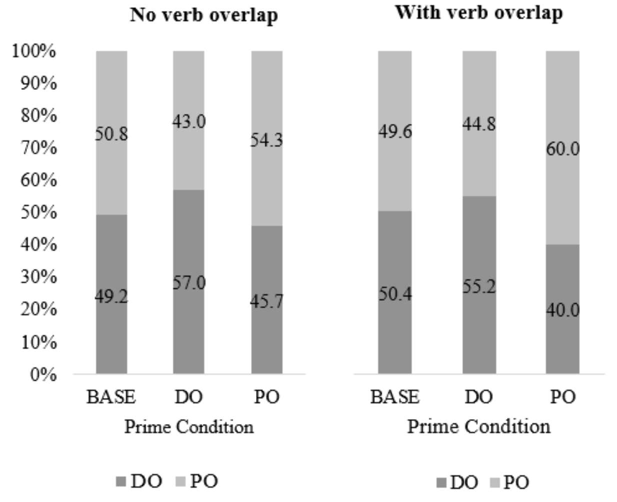


Figure 4.2: Choice for target structure per prime condition.

We fitted the responses to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer in order to increase convergeability (Powell, 2009). We ran a model with Prime Condition (baseline/DO/PO) and Verb Overlap (no/yes) and their interaction as fixed factors. The baseline prime condition and no verb overlap served as the reference level. We started with the maximal random effects structure as proposed by (Barr et al., 2013), adding random slopes and random intercepts for Participants and Items. At each step, we simplified the model until the model converged without any singularity issues. The final model included random intercepts for Participants and Items and no random slopes. We calculated the exponentiated estimate using the *summ* function of the *jtools* package (Long, 2022). The model output is summarized in the appendix (Table A.4.1).

The participants chose significantly more often to read a DO after a DO prime (57.0% in the condition without verb overlap, 55.2% in the condition with verb overlap) than after a baseline prime (49.2% in the condition without verb overlap, 50.4% in the condition with verb overlap) (β = -0.35, SE = 0.09, p < .001), indicating DO priming. With regard to DO priming, there was no significant difference between the condition with and without verb overlap. In the condition without verb overlap between prime and target, there was only a trend towards more PO responses after a PO prime (54.3%) than after a baseline prime (50.8%) (β = 0.16, SE = 0.09, p < .1), but in the condition with verb overlap, there were significantly more PO responses after a PO prime (60.0%) as compared to a baseline prime (49.6%) (β = 0.31, SE = 0.12, p < .05).

Reaction time data

For the reaction time data, we included a total of 13,824 observations (before exclusion of any outliers, which were calculated for the disambiguating word and each spillover word separately). Table 4.1 and Table 4.2 show the mean reaction times per prime condition for the disambiguating word and each words in the spillover region of the DO and PO target sentences respectively. Figure 4.3 and 4.4 plot the reaction times (in ms) per prime condition, separately for the condition with and without verb overlap.

Table 4.1: Mean reaction times (in ms) of DO targets per prime condition.

Prime Condition	Disambiguating word	Spillover word 1	Spillover word 2
BASE	919 (394)	754 (351)	794 (262)
DO	917 (378)	744 (324)	800 (263)
PO	933 (388)	767 (328)	813 (269)

Note: Standard deviations in parentheses.

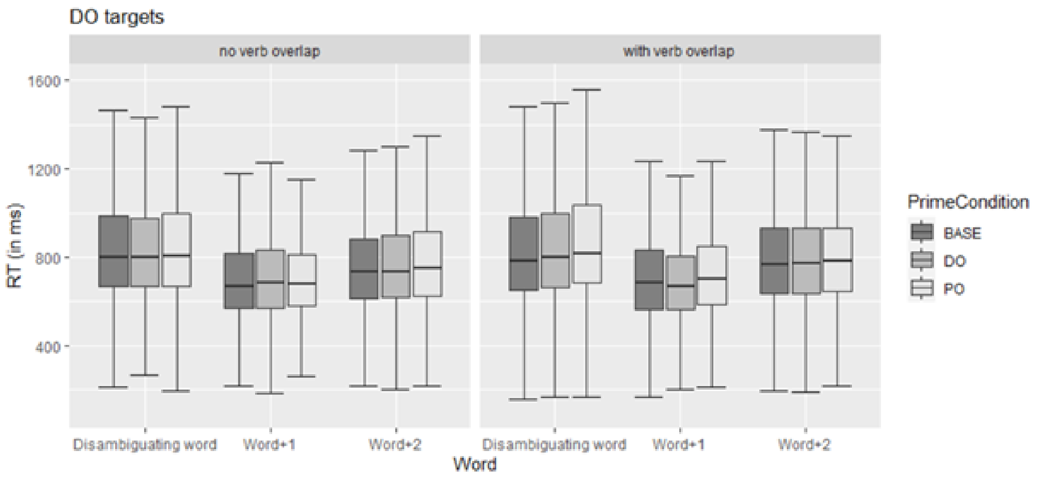


Figure 4.3: Reaction times per prime condition (DO targets) in ms.

We fitted the log-transformed reaction times of each word from the critical point to a linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with Target Condition (DO/PO), Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), Choice (no/yes), and their interactions as its fixed factors. For each fixed effect, the level mentioned first here served as the reference level. We ran separate analyses for the words of the DO targets (6,884 observations) and for PO targets (6,940 observations), since we are interested in the priming effects within each target structure. For each model, we started from the full model with a maximal random effects structure (Barr et al., 2013) and simplified it until convergence. In addition, we removed the interactions which did not contribute significantly to the model in order to interpret the main effects. Full model outputs are reported in the appendix (Table A.4.2-A.4.8).

The participants were slower to respond to a DO target after a PO prime than after a baseline prime. This finding was nearly significant in the disambiguating area ($\beta = 0.018$, $SE = 0.009$, $p < .1$) and was statistically significant in both the first ($\beta = 0.026$, $SE = 0.012$, $p < .05$) and the second word of the spillover area ($\beta = 0.017$, $SE = 0.009$, $p < .05$).

Table 4.2: Mean reaction times (in ms) of PO targets per prime condition.

Prime Condition	Disambiguating word	Spillover word 1	Spillover word 2	Spillover word 3
BASE	902 (379)	689 (264)	683 (217)	788 (262)
DO	903 (379)	688 (270)	685 (218)	797 (271)
PO	899 (367)	675 (254)	684 (209)	803 (267)

Note: Standard deviations in parentheses.

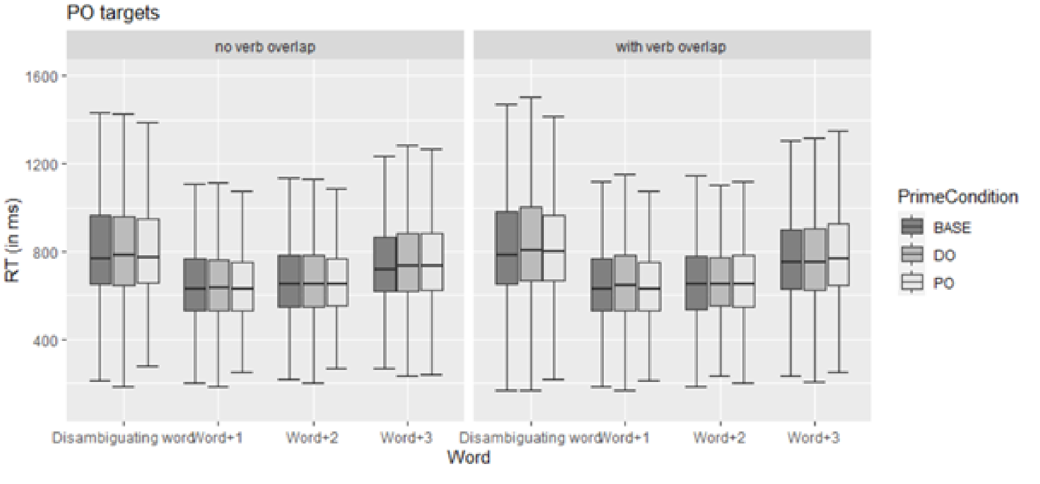


Figure 4.4: Reaction times per prime condition (PO targets) in ms.

We did not find a significant effect of prime structure on reaction times to PO targets in the disambiguating word. In the spillover region, there was a tendency towards faster reaction times to a PO target after a PO prime than after a baseline prime ($\beta = -0.013$, $SE = 0.008$, $p < .1$). This facilitatory tendency was compensated in the final word of the PO target sentence, where participants were significantly slower after a PO prime than after a baseline prime ($\beta = 0.019$, $SE = 0.007$, $p < .01$).

Whether participants had a choice which target structure to read, affected the reaction times to the last word of the target sentence. Participants who chose to read a DO target, had significantly faster reaction times than participants who were exposed to a DO target in the no-choice version ($\beta = -0.062$, $SE = 0.022$, $p < .01$). Similarly, participants who chose for a PO target, were faster to read the final word of the sentence than participants reading a PO target in the no-choice version ($\beta = -0.072$, $SE = 0.023$, $p < .01$). The choice condition never interacted with any of the priming effects. Reaction times were not affected by verb overlap.

4.3 Experiment 2: Production experiment

Method

Participants

We recruited an additional 192 participants (117 female, 73 male, 2 other) with American English as their L1 on Prolific. Participants were aged between 18 and 80 (mean: 41.9, *SD*: 14.7) and did not report any literacy or language problems. They received monetary compensation.

Materials and design

We used the same 36 prime-target pairs with ditransitive sentences and 144 transitive and intransitive fillers as in Experiment 1. Again, we had prime sentences in three prime conditions: a baseline condition, which was an intransitive sentence, a direct object (DO) dative sentence, and a prepositional object (PO) dative sentence.

For each target sentence, we selected a corresponding picture, using the same pictures as in Hartsuiker and Pickering (2008). In addition, we created a prompt for each target, so that the recorded time of writing onset would be at the critical point. For instance, the prompt could be *The waitress passes the*, which could be completed by [*boxer the cake*] or [*cake to the boxer*]. For the transitive fillers, the prompt included the subject and the main verb (e.g., *The witch tickles* [*the dancer*]). As for the intransitive fillers, only the subject was given (e.g., *The monk* [*is laughing*].

We added a verification task to the prime sentences. For this purpose, we selected a picture for each prime sentence. Half of the pictures corresponded to the prime sentence, whereas the other half of the pictures did not match the action described by the prime sentences.

We created two versions: one with verb overlap between prime and target, and one without lexical overlap. For both versions, we constructed three lists, so that every item was preceded by a prime from a different prime condition across the lists. Within each list, we presented the prime sentences equally often in the three priming conditions. Participants were pseudo-randomly assigned to one of the lists; we tested 96 participants per version.

Identical to the comprehension experiment, the production experiment was run using PsychoPy/Pavlovia (Peirce et al., 2019).

Procedure

For each trial of the priming experiment, participants were first presented with the prime sentence. After reading the prime sentence, they would press Space. The prime sentence was then replaced by a verification picture and participants had to decide whether the picture matched the sentence they just read. Then the target picture was shown with the verb written below it. In addition, they would see a text box with the prompt. Participants were asked to write a completion to the sentence.

As in Experiment 1, participants completed the LexTALE test and a survey on their demographic background immediately after the priming experiment. A session took about 40 minutes in total.

Coding

We coded the typed responses as a DO dative, a PO dative, or as an "Other" response. "Other" responses included responses in which only one of the animate entities was mentioned (e.g., *The nun showed the hat*).

We recorded the timing of the first key press, that is, when participants started to type their response. For the analysis of the reaction times, we removed the outliers (reaction times deviating more than three times the standard deviation from the mean). Since the data were not normally distributed, we log-transformed the reaction times (log10(x)). We checked the Q-Q-plot to certify that the log-transformed reaction times had a more normal distribution. We calculated the priming effect by comparing the reaction times after a DO or PO prime to the reaction times after a baseline prime.

Results

Choice data

We collected a total of 6,912 observations. Participants' responses were coded as a PO dative in 4,590 of the cases (66.4%) and as a DO dative in 2,039 (29.5%) of the cases. 283 responses were coded as "Other" (4.09%). We excluded the "Other" responses for further analyses. Figure 4.5 shows the choice for a DO or a PO structure per prime condition, with and without verb overlap.

We fitted the responses to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer in order to increase convergeability (Powell, 2009). We ran a model with Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), and their interaction as fixed factors. The baseline prime condition and no verb overlap served as the reference level. We started with the maximal random effects structure as proposed by Barr et al. (2013), adding random slopes and random intercepts for Participants and Items. We simplified the model until it converged without any singularity issues. The final model included random intercepts for Participants and Items and a random slope of the factor Verb Overlap to both Participants and Items. The model output is summarized in the appendix (Table A.4.9).

The significant intercept indicates that participants produced significantly more PO datives than DO datives in the baseline condition ($\beta = 0.79$, $SE = 0.29$, $p < .01$). After a baseline prime, participants in the condition with verb overlap produced more PO datives ($\beta = 1.33$, $SE = 0.34$, $p < .001$). Participants produced more PO datives after a PO prime

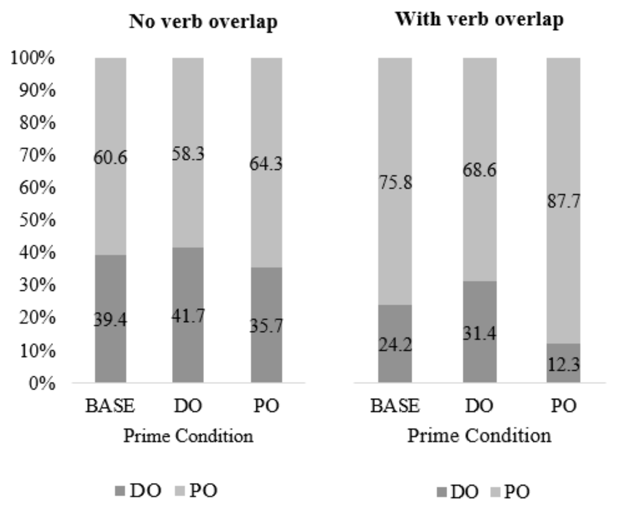


Figure 4.5: Produced structure per prime condition.

than after a baseline prime in the condition without verb overlap (β = 0.28, *SE* = 0.012, $p <$ .05), and this priming effect was even stronger in the condition with verb overlap (β = 1.13, *SE* = 0.19, $p <$ .001). As for the DO targets, in the condition without verb overlap, participants did not produce more DO datives after a DO prime than after a baseline prime (p = 0.13). However, the proportion of DO datives was significantly higher after a DO prime than after a baseline prime in the condition with verb overlap between prime and target (β = -0.47, *SE* = 0.17, $p <$ .01), indicating DO priming.

Reaction time data

As for the reaction time data, we included a total of 6,568 observations, excluding the 283 "Other" responses and 39 outliers. Table 4.3 shows the mean reaction times per prime condition. In Figure 4.6, the mean reaction times are plotted for both the experiment with and the experiment without overlap.

Table 4.3: Mean onset latencies (in ms) per prime and target condition.

Target Condition	DO target	PO target
BASE	3989 (2459)	3485 (2474)
DO	4095 (2705)	3565 (2563)
PO	4074 (2566)	3506 (2.410)

Note: Standard deviations in parentheses.

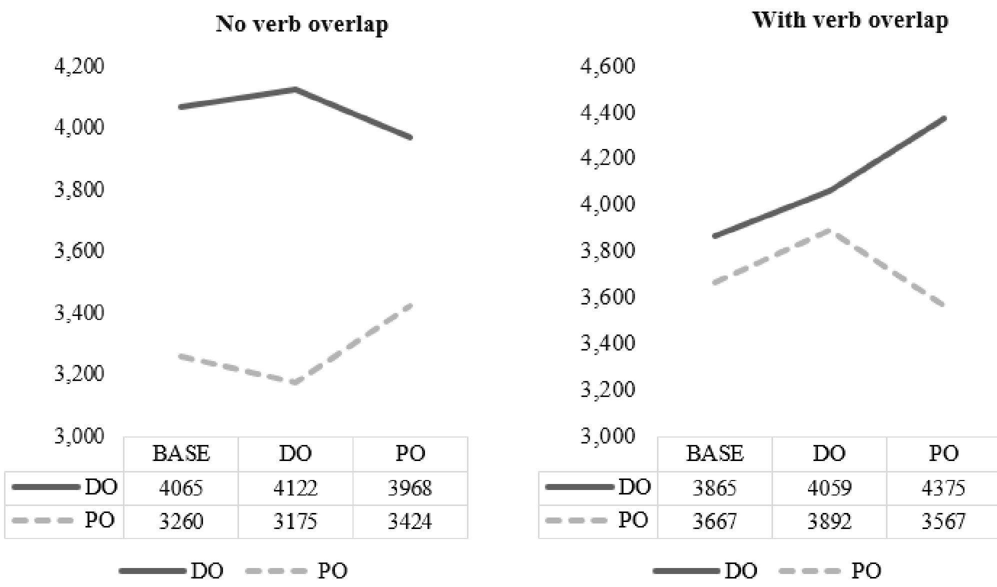


Figure 4.6: Mean onset latencies (in ms) per prime and target condition, with/without verb overlap.

We fitted the log-transformed onset latencies to a linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with Target Condition (DO/PO), Prime Condition (baseline/DO/PO), Verb Overlap (no/yes) and their interactions as its fixed factors. We started from the full model with a maximal random effects structure (Barr et al., 2013) and simplified it until the model converged and had no singularity issues. The final model included random intercepts for Participant and Item and no random slopes. The model output is reported in the appendix (Table A.4.10).

The reaction times were significantly lower for PO sentences than for DO sentences (β = -0.196, SE = 0.028, $p < .001$). In addition, reaction times to PO target sentences were higher in the condition with verb overlap than in the condition without verb overlap (β = 0.103, SE = 0.041, $p < .05$). Most importantly, participants reacted significantly faster to a PO target sentence after a PO prime than after a baseline prime when there was verb overlap between prime and target (β = -0.118, SE = 0.059, $p < .05$).

4.4 Discussion

In the comprehension experiment, participants chose significantly more often to read a DO structure after a DO prime than after a baseline prime, and this effect was not affected by lexical overlap between prime and target. In contrast, the effect of PO priming was only significant if there was lexical overlap between prime and target. As for the reaction times, we found that participants responded significantly slower to a DO target after a PO prime sentence than after a baseline prime sentence, regardless of lexical overlap. In addition, participants were faster to process a PO target after a PO prime sentence than after a baseline prime sentence in the verb overlap condition. In the production experiment, we observed largely the opposite pattern. Participants produced more PO sentences after a PO prime than after a baseline prime, and this priming effect was stronger if there was lexical overlap between prime and target. Participants also used a DO sentence more often after a DO prime than after a baseline prime, but only in the lexical overlap condition. In the reaction time data, we found that participants were faster to start producing their PO sentences after a PO prime than after a baseline prime, but only if there was verb overlap between prime and target.

With regard to our first research question, our study demonstrates that the maze task paradigm is able to detect structural priming effects in comprehension. In addition, providing participants with a choice between the two alternatives does not significantly affect the results compared to a design in which participants read a particular structure for each target item. Therefore, the maze task seems to be a suitable paradigm to collect choice data and reaction times simultaneously in a structural priming experiment. We also found some evidence for the lexical boost effect, given that PO priming was only found in the presence of lexical overlap between prime and target. This shows that the maze task is not only able to detect structural priming, but is also sensitive enough to detect certain interactions (such as lexical overlap) with the priming effect.

Our second research question was how these priming effects in comprehension relate to priming effects in production. For this purpose, we distinguished between measurement type (choice data vs. reaction times) and modality (production vs. comprehension). Although we do not find significant effects in all conditions, we do find evidence for structural priming effects in the choice data as well as in the reaction time data in both modalities, and these structural priming effects can be lexically boosted in both production and comprehension.

More specifically, in the choice data, we found lexically-mediated priming of both the DO and the PO structure in production and in comprehension. In production, the lexical boost effect could directly be observed in the PO condition: there was significant PO priming in the condition without lexical overlap, and priming was significantly stronger in the condition with lexical overlap. Although there was no direct evidence for the lexical boost effect in the comprehension data, it can still be inferred that the effect occurs. In the DO condition, there was a significant priming effect in the absence of lexical overlap, suggesting that lexical overlap is not a prerequisite to find structural priming. In the PO condition, priming was only significant in the presence of lexical overlap, suggesting that priming is detected more easily when there is lexical overlap between prime and target. Together, these results suggest that there is abstract structural priming in comprehension which can be magnified by the lexical boost effect.

Note that in our study, we compared the priming effects against a baseline reference level, which allowed us to assess DO priming and PO priming separately. This is different from most structural priming studies, in which priming is measured relative to the opposite structure. It is common for studies using a baseline reference level to not observe priming for all priming conditions (e.g., Bernolet et al., 2009; Segaert et al., 2011). The relative proportions of the two structures in the baseline condition are often in between the relative proportions found in the primed conditions. As the structural priming effect is relatively small, the comparison of the individual prime conditions against the baseline reference level may not always turn out to be significant, especially when interactions (i.e., with and without lexical overlap) are involved. Numerically, in both the production and the comprehension choice data, we found an increase in the use of the structure that is primed, and this increase was stronger in the condition with lexical overlap than in the condition without lexical overlap. The choice data thus suggest that the role of lexical overlap is similar for production and for comprehension.

Turning to the reaction time data, the different findings between production and comprehension (that is, facilitation of the PO structure after a PO prime in production, and inhibition of the DO structure after a PO prime in comprehension) may be explained by the two-stage competition model by Segaert et al. (2011). This model aimed to explain their findings for priming of the active and the passive structure in production. In the choice data, they found an effect of inverse preference, that is, stronger priming for the less frequent passive structure. In the reaction time data, on the other hand, they found facilitation of the primed active structure. In other words, they found a preference effect: the more frequent active structure was primed more with regard to the onset latencies than the less frequent passive structure. They explained their results in terms of their two-stage competition model, which has a selection stage and a planning stage.

According to the two-stage competition model (Segaert et al., 2011), grammatical structures are represented in nodes that are associated with a particular activation level. The base level of activation is higher for more frequent structures than for less frequent structures. During the selection stage, the two alternating structures compete with each other. Semantic and pragmatic factors, but also structural priming, may temporarily affect the activation level of a particular structure. When the activation level of one of the structures reaches the selection threshold, that structure is selected and passed on to the planning stage.

When the more frequent structure is primed, in the case of Segaert et al. (2011) the active structure, the gap in activation level between the active and the passive structure increases, due to residual activation. As a result, the selection time for the active structure becomes shorter. By contrast, when the less frequent structure is primed, that is, the passive structure, the activation level of the passive increases, and this decreases the gap in activation level between the active and the passive. Consequently, it takes longer to select the passive structure. This delay is compensated during the planning stage, which proceeds faster for the primed structure.

As such, during the selection stage, the relative frequency of the alternating structures is crucial in whether priming of a particular structure leads to facilitation or inhibition effects. In the choice data, we found different production preferences in the production data and the comprehension data. In the production choice data, there was an overall preference for the PO structure (similar to other production studies with L1 speakers of English, e.g., Corley & Scheepers, 2002; Pickering & Branigan, 1998). In the comprehension choice data, there was no clear preference for either the DO or the PO structure. The strong preference for the PO structure in production may be due to the target pictures depicting the ditransitive action. In these pictures, the object is always in between the agent and the recipient, as is the case in the PO structure. At the same time, participants may have chosen more often for the DO structure in the comprehension experiment, as the choice concerned a choice between an animate and an inanimate noun, and animate nouns may be conceptually more accessible than inanimate nouns (cf. Bock & Warren, 1985), which enhances the choice for the DO structure.

The reaction time data of the production experiment are in accordance with the two-stage selection model (Segaert et al., 2011). We found that participants start producing a PO target response faster after a PO prime than after a baseline prime. We do not find any priming effects for the DO structure. As the PO structure is the most frequent structure in production, there is facilitation both during the selection stage and the planning stage. The lack of priming of the DO structure is caused by inhibition as the result of priming of the least frequent structure during the selection stage, which is compensated during the planning stage. Note that we only find PO priming if there is lexical overlap between prime and target. Segaert et al. did not find an interaction with lexical overlap for the onset latencies; they only found a lexical boost effect in their choice data (but they tested fewer participants than we did, and measured the onset latencies in spoken production instead of written production). Our data show that the onset latencies are sensitive towards verb overlap as well.

As for the reaction time data of the comprehension experiment, we observed that participants were slower to respond to a DO target after a PO prime. This may be caused by inhibition of the PO structure, suggesting that during the comprehension experiment, the PO structure had a lower activation level than the DO structure (due to the accessibility of animate nouns, favoring the DO structure). Crucially, in Segaert et al. (2011), no direct evidence of longer selection times for the less frequent structure was found, and they argue that this is because of compensation during the planning stage, which proceeds faster for the primed structure. In comprehension, there is no planning stage, and consequently, no compensation for the inhibitory effect (although there may occur priming during follow-up processes that are specific to comprehension, such as thematic role mapping). Indeed, we seem to find direct evidence for a longer selection time for a DO target after a PO prime. As such, in the reaction times of the comprehension experiment, we did not find a facilitatory effect of the structure with the highest activation level, but we observed an inhibitory effect of the structure with the lowest activation level.

It is important to note that the competition model of Segaert et al. (2011) is a model for sentence production and not for sentence comprehension. Nevertheless, the selection stage of the model is congruent with constraint-based models of sentence processing (e.g., MacDonald et al., 1994; cf. Humphreys & Gennari, 2014 for neuroimaging evidence). Constraint-based models assume that competition between alternatives occurs in order to resolve syntactic ambiguity. Alternatively, the inhibitory effect of the PO prime on the DO target may be explained by garden path theories (e.g., Frazier, 1979; Pickering et al., 2000, which assume that syntactic ambiguity is resolved by reanalysis. These theories assume that one starts with processing the "default structure", and when the input does not match the expected structure, reanalysis needs to take place in order to process the alternative structure. In this case, one may expect longer processing of the less expected structure. Presumably, this effect is strongest for the least frequent structure after priming of the most frequent structure.

4.5 Conclusion

Our results suggest that it is fruitful to simultaneously collect choice data and reaction time data in production as well as comprehension, in order to compare the two modalities using more equal measurement types. Structural priming effects were very comparable between production and comprehension in the choice data. Priming of a particular structure leads to an increase in the preference for that structure compared to a baseline reference level. In addition, the priming effects are sensitive to the lexical boost effect. Concerning the reaction time data, we observed facilitated processing of a primed structure in production, but inhibited processing of the alternative structure after priming of a structure in comprehension.

This result fits with the two-stage competition model of Segaert et al. (2011), in which there is competition between alternatives during the selection stage. The relative frequency of structures determines whether priming leads to facilitated processing of the primed structure or rather inhibition of the opposite structure. Inhibition effects are more likely to be observed in comprehension than in production, because in production, facilitated processing of the primed structure during the planning stage compensates for the inhibition during the selection stage. Thus, our results suggest that that while observed priming effects may be different between modalities, the underlying mechanism of structural priming seems to be the same.

A better understanding of the relationship between priming in production and in comprehension may aid future studies on the nature of structural representations across modalities in children, L2 speakers or people with language problems such as aphasia, for whom participating in a comprehension experiment may be more feasible than in a production experiment. More research is needed to gain a better understanding of how relative frequency interacts with the polarity (i.e., facilitation or inhibition) and the magnitude (i.e., (inverse) preference effects) of structural priming effects. Future studies should investigate priming effects in comprehension with structures displaying a larger difference in their relative frequencies to further assess the role of facilitation and inhibition during the processing of primed structures. Our study shows that the maze task paradigm may be a valuable tool to investigate such structural priming effects in comprehension.

Chapter 5

The development of L2 syntactic representations in production and comprehension

Previous research has shown that the magnitude of structural priming in the L2 is modulated by proficiency (Bernolet et al., 2013). Abstract structural priming increases with increasing proficiency, whereas the magnitude of lexically-dependent priming decreases during development. The developmental account of L2 syntax proposed by Hartsuiker and Bernolet (2017) explains these proficiency effects in terms of a transition from item-specific representations to abstract structural representations. Arguably, the developmental account applies to both production and comprehension. In our study, we investigated whether the L2 proficiency effects on structural priming as observed in the choice data of Bernolet et al. (2013) in production are also found in the onset latencies of the produced responses, and whether similar proficiency effects can be observed in structural priming in comprehension for both choices and reaction times. Although proficiency indeed modulated priming in the production onset latencies as well as in comprehension, the observed polarity of the correlation between priming and proficiency was not consistent. The data suggest a hybrid developmental account, which assumes an interplay between explicit memory, residual activation and implicit learning, which in turn determine the magnitude of priming at different stages of L2 development.

Keywords: structural priming, production, comprehension, L2 syntax, proficiency

Materials, data, and analyses are available online: `https://osf.io/3wnvk`.

5.1 Introduction

Many learners of a foreign language have the experience of talking to a native speaker but failing to understand their reply. Conversely, especially for languages related to their L1, L2 speakers may be able to understand the L2, but may not be capable to speak the language themselves. This suggests that distinct language skills are required for production and for comprehension. At the same time, when learning a language, the modalities are also intertwined: the learned skills in one domain contribute to the acquisition of skills in the other domain. For instance, Swain (1985) argues that L2 learners require "comprehensible input" as well as "comprehensible output" in order to become proficient in their L2 and, more specifically, to acquire the grammar of the new language.

Nevertheless, L2 development is often studied separately in production and in comprehension. In comprehension, the development of L2 syntactic processing may be investigated by means of ERP experiments measuring responses to syntactic violations or fMRI studies observing neural activation during syntactic rule learning (see Kotz, 2009 for a review). In production, the development of L2 syntactic representations is often investigated by means of structural priming. Structural priming is "the phenomenon by which processing one utterance facilitates processing of another utterance on the basis of a repeated syntactic structure" (Branigan, 2007, p.1), and was developed as an experimental paradigm by Bock (1986), who found that participants were more likely to describe a target picture showing a transitive action with a passive sentence instead of an active sentence if they were exposed to a passive prime sentence than after an active prime sentence. For example, when participants were primed with the passive prime sentence *The building manager was mugged by a gang of teenagers*, they were more likely to produce *The church is being struck by lightning* instead of *Lightning is striking the church* than after an active prime sentence. Although structural priming is stronger if the verb is repeated between prime and target (the so-called lexical boost effect, Pickering & Branigan, 1998), the effect is found in the absence of lexical overlap between prime and target, implying that it taps into abstract representations of syntactic structures.

Crucially, structural priming also takes place between languages. For instance, Hartsuiker et al. (2004) found that participants produced more passive sentences in English after a Spanish passive prime sentence than after a Spanish active prime sentence. It therefore seems to be the case that representations of similar syntactic structures are shared between languages, at least for proficient L2 speakers. Bernolet et al. (2013) found that there is a positive correlation between the magnitude of abstract structural priming and L2 proficiency, whereas there is a negative correlation between the magnitude of the lexical boost effect and L2 proficiency. More specifically, more proficient L2 speakers of English with Dutch as their L1 were primed stronger for genitives (e.g., *the doll of the boy* and *the boy's doll*) than less proficient L2 speakers, at least for critical items which had no overlap in head nouns between prime and target. The less proficient speakers were primed stronger in a within-L2 priming experiment for items with lexical overlap, as they rely on item-specific syntactic representations. These findings suggest that L2 learners start with separate, language-specific structural representations, which become shared between the L1 and the L2 over time. This transition from item-specific representations to abstract structural representations has been captured in the developmental account of L2 syntax (Bernolet & Hartsuiker, 2018; Hartsuiker & Bernolet, 2017). Importantly, the study of Bernolet et al. (2013) is a production study, and as such, the developmental account of L2 syntax is primarily based on findings in production.

Although structural priming has also been found in comprehension (e.g., Arai et al., 2007; Thothathiri & Snedeker, 2008a, 2008b; Traxler, 2008), it is not clear whether the same proficiency effects are found in L2 comprehension as in L2 production. The aim of the current study is to investigate whether the developmental account of L2 syntax also applies to comprehension. More specifically, we test whether priming of ditransitive structures in L2 comprehension is affected by the same proficiency effects as in L2 production. Ditransitive events can be described by either a direct object dative (DO), e.g., *The waitress passes the boxer the cake*, or a prepositional object dative (PO), e.g., *The waitress passes the cake to the boxer*. We performed both a comprehension and a production experiment with Chinese-English and Dutch-English bilinguals, in which we investigated the interaction between structural priming of ditransitives and L2 proficiency in English. As Dutch-English bilinguals are on average more proficient in English than Chinese-English bilinguals, testing these two groups allowed us to investigate a large range of proficiencies.

Structural priming in L2 comprehension

Although the developmental account of L2 syntax is based on structural priming in production, it may also be applicable to comprehension. It has been suggested that syntactic representations are shared between production and comprehension (e.g., Giglio et al., 2022; Indefrey, 2018; Kempen et al., 2012; Pickering & Garrod, 2007), at least in L1 speakers. If syntactic representations are shared between languages, and if the same representations are exploited during syntactic processing for both production and comprehension, we may expect priming between languages in comprehension as well. Kidd et al. (2015) indeed observed cross-linguistic priming in comprehension. They investigated priming of subject and object relative clauses in bilingual speakers of English and German. In English, there is a word order difference between subject relative clauses (*the woman that kisses the girl*) and object relative clauses (*the woman that the girl kisses*), while in German, there is no difference between subject and object relative clauses in the surface structure (*die Frau, die das Mädchen küsst*). They found that German relative clauses were interpreted more often as an object relative clause when they were preceded by an English object relative clause prime (if the verb was repeated between prime and target), implying cross-linguistic priming in comprehension. Similar results were obtained by Hsieh (2017), priming from Chinese to English. Here, cross-linguistic priming was also found in the absence of verb overlap between prime and target, suggesting that abstract structural representations are shared between languages in comprehension as well.

In addition, if syntactic representations are shared between modalities, then the development of L2 syntax may progress similarly in production and in comprehension. In that case, we may expect similar proficiency effects when testing structural priming in comprehension as in production. That is, abstract structural priming may be stronger in more proficient participants than in less proficient participants, whereas lexically-dependent priming may be stronger in less proficient participants than in more proficient participants (cf. Bernolet et al., 2013). To our knowledge, no direct evidence for proficiency effects in L2 comprehension has been found in priming studies. Wei et al. (2018) investigated whether the magnitude of within-language structural priming of English relative clauses was affected by proficiency in L2 speakers of English with Chinese as their L1, but did not find any effect, and they suggest that this may be because their participant group was too homogeneous in terms of their proficiency in English.

Nevertheless, Hwang et al. (2018) provided evidence for shared representations between the L1 and L2 which become stronger as proficiency increases in a non-priming study. Participants with Korean as their L1 were shown pictures depicting a causative event together with either an active (e.g., *Jen fixed a computer*) or a causative (e.g., *Jen had her computer fixed*) sentence in English, their L2. They were instructed to decide whether the sentence described the event accurately. In Korean, active structures are used for both causative and transitive events. Participants who were more proficient in English, were more likely to incorrectly accept an active sentence as an accurate description for the causative event than less proficient participants, implying that cross-linguistic influence increases as proficiency increases. As such, it seems that syntactic structures become shared over time, predicting proficiency effects in structural priming in comprehension too.

Similar to Bernolet et al. (2013), the effect of lexical overlap between prime and target on the magnitude of structural priming may depend on proficiency. Indeed, there have been some variable results on the lexical boost effect in comprehension. Wei et al. (2018) only found significant structural priming in comprehension in L2 speakers in the presence of lexical overlap between prime and target, whereas Wei et al. (2019) did observe abstract structural priming. They found within-language priming of English reduced relative clauses in a self-paced reading experiment with L2 speakers of English who had Chinese as their L1. Participants were faster to read a reduced relative clause sentence after a reduced relative clause prime sentence than after a main clause prime sentence, both in the condition with and without verb overlap between prime and target. One reason for the necessity of lexical overlap in order to observe structural priming in some studies but not in other, may be the role of proficiency.

However, explaining differences in results between modalities with regard to the lexical boost effect in terms of L2 proficiency is complicated, since similar inconsistencies in terms of the lexical boost effect have been found in priming studies in L1 comprehension. First, it has been suggested that the magnitude of priming is generally weaker in comprehension than in production (Ziegler & Snedeker, 2019). In contrast to production, listeners may not need to build a full syntactic representation of a sentence in comprehension, which may lead to less residual activation than in production. Second, some authors argue that the lexical boost effect behaves differently in the two modalities. This is primarily due to inconsistent results in comprehension. Some studies only find structural priming in comprehension if there is verb overlap between prime and target (e.g., Arai et al., 2007; Branigan et al., 2005; Q. Chen et al., 2013; Tooley et al., 2009; Traxler, 2015). Others find equally strong abstract structural priming as lexically-dependent priming (Fine & Jaeger, 2016; Traxler, 2008). In addition, it has been argued that the lexical boost effect may be long-lasting in comprehension (Pickering et al., 2013; Tooley et al., 2014), while it decays quickly in production (Hartsuiker & Pickering, 2008). On the other hand, Traxler (2008) found evidence for stronger priming in comprehension if there is lexical overlap, and Fine and Jaeger (2016) and Tooley and Traxler (2018) found that the lexical boost effect decays during intervening items, suggesting that the lexical boost effect may work similar between production and comprehension.

Comparing structural priming between modalities

Given the important role of the magnitude of structural priming and the lexical boost effect in the developmental account of L2 syntax on the one hand, and the inconsistencies between production and comprehension on the other hand, it is challenging to investigate to what extent the developmental account as outlined for production, also applies to comprehension. In other words, one needs to disentangle any modality-specific differences (which are also observed in L1 priming) from the proficiency effects associated with the L2 development of syntactic structures. There may also be modality-specific effects that only play a role in L2 priming, depending on the proficiency level. For instance, especially in beginning L2 learners it might be the case that participants avoid the more difficult structure in production, while they may show priming in comprehension, revealing a mental representation of the unused structure. As such, in order to contribute to the understanding the relationship between production and comprehension during the development of L2 syntactic structures, we need to compare structural priming effects between production and comprehension in L2 speakers of various proficiency levels. Such a comparison may enable us to assess differences in priming between production and comprehension at different stages of L2 development.

When comparing structural priming between production and comprehension, in addition to any differences induced by modality-specific processing, one needs to take into account methodological differences. Structural priming studies in production often exploit a picture description task, in which a picture can be described with two alternating structures, such as a DO or a PO dative structure. This task results in (binomial) choice data. In comprehension studies, structural priming is often measured by means of online data, such as reaction times. This may especially be problematic when investigating how the magnitude of structural priming is affected by proficiency, as the magnitude of priming effects may be affected by the measurement type even within modality. For instance, in their production study, Segaert et al. (2011) found stronger passive priming than active priming in the choice data, whereas active priming was stronger than passive priming in the initiation times of responding.

To remedy for this difference in measurement types, in an earlier study with L1 speakers of English (see Chapter 4), we compared structural priming effects between modalities within the same measurement type. That is, we collected choice data and reaction time data in both production and in comprehension. In production, participants performed a picture naming task, and we collected their choice for either the PO or the DO dative. In addition, we measured their onset latencies, that is, at what time they started to produce their sentence (cf. Segaert et al., 2011). In comprehension, we used an adapted version of the maze task paradigm, in which participants read sentences word by word, for each word choosing between the correct word and a distractor word (Boyce et al., 2020; Forster et al., 2009). In the adapted version, participants chose between reading a PO or a DO dative. For the choice data, we found lexically-boosted priming in both production and comprehension. In the reaction time data, there was a difference between modalities. In production, PO target sentences were produced faster after a PO prime than after a baseline prime if there was verb overlap between prime and target. In comprehension, participants were slower to read a DO target after a PO prime than after a baseline prime, regardless of whether there was lexical overlap. In other words, structural priming may not only have a facilitatory effect on the processing of the primed structure, but also an inhibitory effect on the alternative structure, especially in reaction times.

Two-stage competition model

The contrasting results in the reaction times (namely a facilitatory effect in the PO structure in production and an inhibitory effect of PO primes on DO targets in comprehension) as observed in Chapter 4 may be explained by the two-stage competition model for structural priming effects (Segaert et al., 2011, 2014, 2016). This model is a hybrid model of structural priming, attributing the priming effects to both residual activation of recently activated nodes representing syntactic structures (Pickering & Branigan, 1998) and an error-based implicit learning mechanism, meaning that encountering a structure leads to long-term changes in the relative weights of structures (Chang et al., 2006). In the two-stage competition model, syntactic structures are represented as nodes, which have a particular base level of activation. This base level of activation is updated by implicit learning and is higher for more frequent structures than for less frequent structures. Language processing takes place in two distinct stages: a selection stage and, at least for production, a planning stage. During the selection stage, there is competition between alternating structures (such as the DO and the PO). A structure is selected when the activation level of that structure reaches its selection threshold. The amount of competition depends on the relative activation levels of the alternating structures. The activation level may be affected by semantic, pragmatic, and contextual factors, as well as by structural priming. When the more frequent structure is primed, residual activation is added to the already higher base level of activation of that structure. As a result, the gap in activation level between the two competing structures increases, which decreases the competition and thus the selection time of that particular structure. This may be why PO target sentences were produced faster after a PO prime than after a baseline prime in Chapter 4, as the PO structure was the most frequent structure in production.

When on the other hand the less frequent structure is primed, the difference between the activation levels of the most frequent and the less frequent structure decreases due to residual activation of the latter. As a result, there is more competition between the two structures, which leads to longer selection times. In production, this delay is compensated by priming during the planning stage, which resulted in a null effect for the least frequent structure in Segaert et al. (2011). But in comprehension, there presumably is no planning stage, meaning that inhibition effects may be observed (although there may theoretically also be priming during comprehension-specific follow-up processing stages). In Chapter 4, participants were slower to read a DO target after a PO prime than after a baseline prime, suggesting that priming of the PO structure (the least frequent structure in comprehension) led to inhibition in comprehension.

As such, the relative frequency of the alternating structures seems to play an important role in the magnitude of structural priming and even the polarity of structural priming. Note that the relative frequency also affects the magnitude of priming in choice data, in which the less frequent structure is often primed stronger than the more frequent structure (the so-called inverse preference effect, V. S. Ferreira & Bock, 2006). Typically, the relative frequency of alternating structures may be different for L2 speakers than for L1 speakers, as the consequence of linguistic transfer from the L1, and may also differ between L2 speakers at different proficiency levels due to the differential amount of linguistic transfer (cf. Hwang et al., 2018) and input in the L2 (cf. Sijyeniyo et al., 2023).

Importantly, it is not necessarily the case that we find facilitation in production and inhibition in comprehension. For instance, in Wei et al. (2019), reduced relative clause sentences were read faster after a reduced relative clause prime sentence than after a main clause prime sentence. For the most frequent alternative, facilitation effects may be found in both production and in comprehension. Inhibition effects are predicted to only be found for the least frequent alternative. Due to compensation of inhibition effects during the planning stage in production, inhibition effects may be more evident in comprehension than in production, although there may also occur priming during follow-up processing in comprehension. Also, the two-stage competition model (Segaert et al., 2011) is not explicit on what happens if the target structure is different from the primed structure. Especially in comprehension, participants may be exposed to the alternative structure after priming with a particular structure. But also in production, participants still produce the non-primed structure, while there may still be an effect of the prime on the target structure visible in the onset latencies. Following the reasoning from Segaert et al., as the competition between structures is affected by the prime structure, the effects may be independent from the target structure. So, the inhibition effects induced by the least frequent structure may be found both if the structure is repeated between prime and target, and if the structure is different between prime and target. However, it is possible that the compensatory priming during the selection stage in production and perhaps during other follow-up processes (also in comprehension) only occurs if the prime and the target structure are the same.

Taken together, in order to determine whether the developmental account of L2 syntax applies to both production and comprehension, several aspects need to be taken into account. First, the lexical boost effect is not only involved in proficiency effects, but may also lead to differential results between production and comprehension in general. It is therefore crucial to not only test for proficiency effects in comprehension, but to compare these results to a similar experiment in production and more specifically, to compare between modalities within measurement types (choice data and reaction times). Second, the magnitude and even the polarity of structural priming is not only affected by proficiency and/or modality, but also by the relative frequency of the alternating structures. It may therefore be the case that different effects are found for production than for comprehension, while the underlying mechanism may be the same. A way to take into consideration the role of relative frequency when interpreting any differences in terms of proficiency and/or modality-specific effects, is to test multiple groups of L2 speakers with different L1s, in order to validate any observed effects in a replication design.

Current study

In the current study, we will investigate to what extent the developmental account of L2 syntax (Hartsuiker & Bernolet, 2017), while being developed for production, applies to comprehension as well. Our first aim is to investigate whether we can replicate the proficiency effects found in Bernolet et al. (2013) in production and more specifically, whether these proficiency effects are found both in the choice data and in the reaction time data. The production study of Bernolet et al. only collected choice data. To our knowledge, there are no studies investigating the onset latencies in production in L2 speakers. For this purpose, we performed two production experiments with ditransitives with L2 speakers of English who have either Chinese (Experiment 1A) or Dutch (Experiment 1B) as their L1, collecting both the production choices and the onset latencies. Both Chinese and Dutch have a similar alternation between the PO and the DO structure as English, and in both languages L1 speakers displayed a preference for the PO structure in previous production experiments (cf. Cai et al., 2012 [Chinese]; Bernolet & Hartsuiker, 2010 [Dutch]). We expect to replicate the findings of Bernolet et al. (2013). More specifically, we predict stronger abstract structural priming in more proficient L2 speakers than in less proficient learners, and stronger priming in items with lexical overlap between prime and target in less proficient L2 speakers than in more proficient L2 speakers.

The second aim of our study is to explore whether any proficiency effects in production can also be observed in comprehension, both in choice data and in reaction time data. In other words, is there a correlation between L2 proficiency and the magnitude of priming in comprehension? More specifically, do we find stronger abstract structural priming in more proficient L2 speakers than in less proficient learners on the one hand, and stronger priming in items with lexical overlap between prime and target in less proficient L2 speakers than in more proficient L2 speakers? In two comprehension experiments, we test structural priming effects of ditransitive structures in L2 speakers of English who have either Chinese (Experiment 2A) or Dutch (Experiment 2B) as their L1. If structural priming in production and comprehension rely on the same underlying mechanisms, we expect similar priming effects in the comprehension experiments and in the production experiments. In addition, if the developmental account of L2 syntax applies to both production and comprehension, any observed interaction between priming and L2 proficiency may also be similar between the modalities.

5.2 Experiment 1A: L2 production (L1 Chinese)

Method

Participants

For the first L2 production experiment, we tested 96 participants in the verb overlap condition and 96 participants in the condition without verb overlap. For this purpose, we recruited 192 participants (130 female, 46 male, 16 other) of any level of L2 English with Chinese as their L1 on WeChat. Participants were aged between 15 and 30 (mean: 21.6, *SD*: 2.5) and did not report any literacy or language problems. They were paid for their participation.

Materials and design

We used the same materials as in Chapter 4. There were 36 prime-target pairs with ditransitive sentences. The prime sentences came in three conditions: a baseline condition with intransitive sentences (e.g., *The monk is laughing*), the DO condition (e.g., *The waitress passes the boxer the cake*) and the PO condition (e.g., *The waitress passes the cake to the boxer*). We also used 144 filler sentences and 4 practice sentences, which were either intransitive sentences (similar to the baseline condition) or transitive sentences (e.g., *The witch tickles the dancer*).

Each prime sentence was paired with a verification picture. Half of the pictures displayed the action described by the prime sentence, whereas the other half of the pictures did not match the event from the prime sentence. Each target sentence was associated with a corresponding picture and a prompt. The pictures were re-used from Hartsuiker and Pickering (2008). The prompt included the target sentence until the critical point, so that the recorded time of the writing onset would be at the critical point. The first character of the noun was always different between the theme and the recipient. To illustrate, the prompt could be *The waitress passes the*, which could be completed by [*boxer the cake*] or [*cake to the boxer*]. The transitive fillers had the subject and the main verb as the prompt (e.g., *The witch tickles* [*the dancer*]). As for the intransitive fillers, only the subject was provided (e.g., *The monk* [*is laughing*]).

Procedure

During each trial of the priming experiment, participants read the prime sentence and pressed Space to continue. They were then presented with the verification picture. Participants were instructed to decide whether the picture corresponded to the prime sentence. Upon their decision, the verification picture was replaced by the target picture with the verb below it. In addition, there was a text box containing the verb prompt. Participants were asked to write a completion to the sentence, describing the target picture.

After the priming experiment, participants performed the LexTALE test, which is a short yes/no-vocabulary test (Lemhöfer & Broersma, 2012) and completed a survey in which we asked for their self-rated proficiency of English and for additional background information. A session took about 45 minutes.

Coding

The typed responses were coded as a DO dative, PO dative, or an "Other" response. "Other" responses were ungrammatical responses or responses in which only one of the animate entities was mentioned (e.g., *The waitress passes the cake*). We measured the priming effect by comparing the proportion of DO or PO responses after a DO or PO prime to the proportion of DO or PO responses after a baseline prime.

In addition to the written responses, we recorded the timing of the first key press, which was the time at which participants started to type their response after the start of the exposure to the target picture. We removed any outliers (reaction times deviating more than three times the standard deviation from the mean) before the analyses. Since the data were not normally distributed, we log-transformed the reaction times (log10(x)), after which the Q-Q plot displayed a more normal distribution. We calculated the priming effect by comparing the reaction times after a DO or PO prime to the reaction times after a baseline prime.

Results

Proficiency

Participants had an average LexTALE score of 58.4% (*SD* 12.1), ranging from 37.5% to 97.5%. On a scale of 0 (very low) to 10 (very high), they rated their proficiency in speaking at 4.4 (*SD* 1.8, range 1-8), in listening at 4.6 (*SD* 1.9, range 1-9), in reading at 6.1 (*SD* 1.8, range 1-10) and in writing at 5.1 (*SD* 1.9, range 0-9). We also calculated their overall self-rated proficiency by averaging the self-rated proficiency scores of the four skills for each participant, which was 5.1 (*SD* 1.5), ranging from 1 to 8. The average self-rated proficiency had a more normal distribution than the LexTALE scores, according to the Q-Q plots, which is why we decided to only include self-rated proficiency in the reported analyses (although we did check whether the results remained the same if we inserted LexTALE scores instead of self-rated proficiency scores in the model, which was the case).

Choice data

We collected 6,912 observations in total. Participants produced a PO dative in 3,968 (57.4%) of the target sentences and a DO dative in 2,094 (30.3%) of the target sentences. 850 responses (12.3%) were coded as "Other" responses. We excluded the "Other" responses for further analyses. Figure 5.1 shows the proportion of DO and PO responses per prime condition, with and without verb overlap.

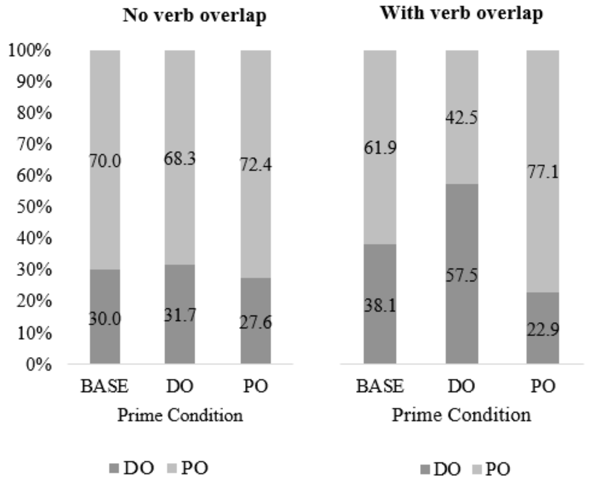


Figure 5.1: Produced structure per prime condition (Chinese group).

We fitted the responses to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer in order to increase convergeability (Powell, 2009). We inserted Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), self-rated proficiency scores, and the interactions as fixed effects. The baseline prime condition and the no verb overlap condition were inserted as the reference level. We also included random intercepts for Participant and Item, and random slopes of the fixed effects to Participant and Item. We started from the full model with a maximal random effects structure (Barr et al., 2013) and simplified it until convergence. The final model included random intercepts for Participant and Item and random slopes of Prime Condition to Participant and to Item and of Verb overlap to Item. As Proficiency was not involved in any significant interaction, the displayed model output only includes the interaction between Prime Condition and Verb Overlap. The model output summary is reported in the appendix (Table A.5.1).

Participants were more likely to produce a PO dative than a DO dative in the baseline condition, as indicated by the significant intercept (β = 1.791, *SE* = 0.377, p < .001). Participants descriptively used more PO datives after a PO prime than after a baseline prime in the condition without verb overlap, but this tendency did not reach conventional levels of significance (β = 0.363, *SE* = 0.212, p < .1). The participants produced significantly more PO datives after a PO prime than after a baseline prime in the presence of lexical overlap (β = 1.011, *SE* = 0.275, p < .001). Participants also produced more DO datives after a DO prime than after a baseline prime if there was verb overlap between prime and target (β = -1.287, *SE* = 0.227, p < .001). Thus, we found lexically-dependent priming of both the PO and the DO dative. The proportions of PO and DO datives were not significantly affected by Proficiency, nor did they interact with the priming effects.

Reaction time data

We included 5,847 observations in the analyses for the reaction time data, excluding 850 "Other" responses and 215 outliers. In Table 5.1, the mean reaction times per prime condition are displayed. Figure 5.2 shows the mean reaction times separately for the conditions with and without verb overlap.

Table 5.1: Mean onset latencies (in ms) per prime and target condition (Chinese group).

Target Condition	DO target	PO target
BASE	4850 (6262)	5225 (6342)
DO	4696 (5165)	5863 (7112)
PO	4719 (6383)	5187 (6587)

Note: Standard deviations in parentheses.

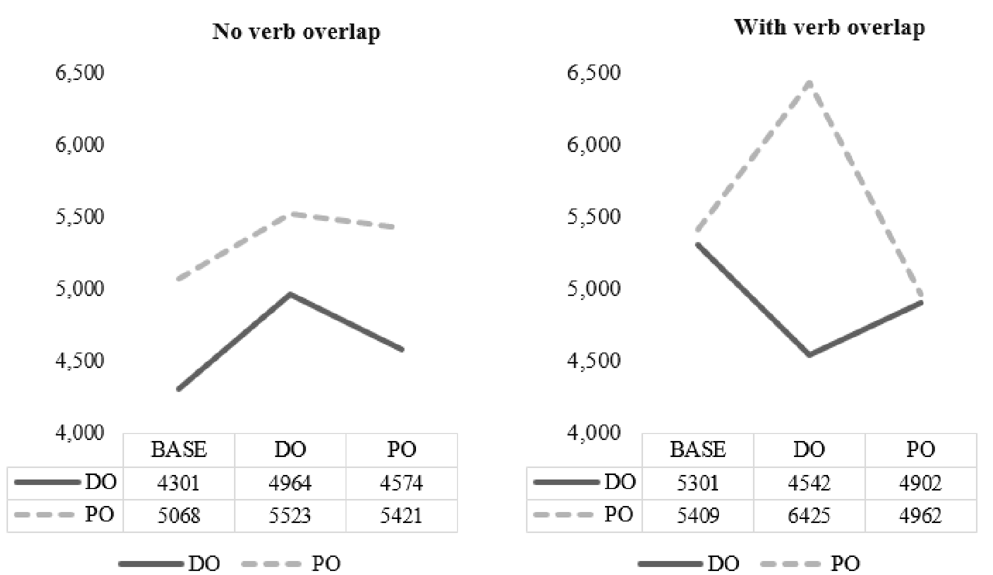


Figure 5.2: Mean onset latencies (in ms) per prime and target condition (Chinese group).

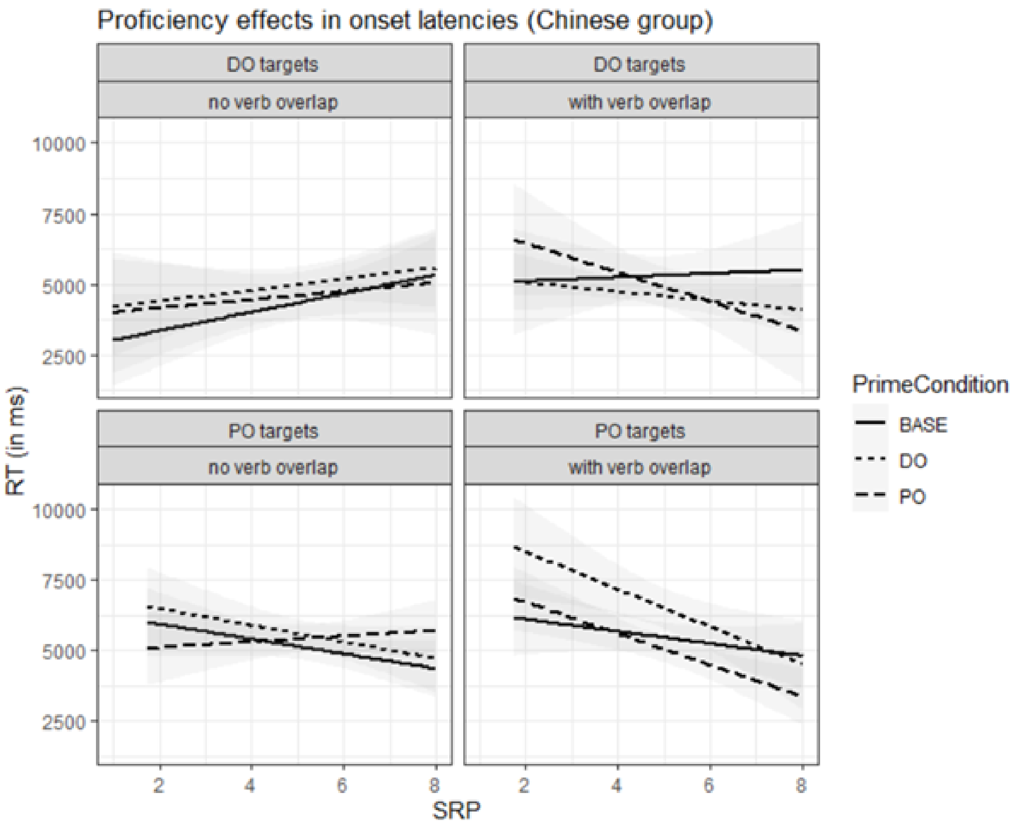


Figure 5.3: Proficiency effects in onset latencies (Chinese group).

The log-transformed reaction times were fitted to a linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with Target Condition (DO/PO), Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), the self-rated proficiency scores and the interactions as its fixed factors. The DO target condition, the baseline prime condition and the no verb overlap condition were the reference level. We started from the full model with a maximal random effects structure (Barr et al., 2013) and simplified it until the model converged and had no singularity issues. The final model included random intercepts for Participant and Item and a random slope of Target Condition for Participant. Using the *drop1*-function, we checked whether the four-way interaction contributed significantly to our model. As this was not the case, we removed it from our model. The model output is summarized in the appendix (Table A.5.2). The interactions between proficiency, priming, and verb overlap are plotted in Figure 5.3.

Reaction times to DO targets were lower after a DO prime than after a baseline prime when there was verb overlap between prime and target (β = -0.150, SE = 0.063, $p < .05$). In addition, participants were significantly slower to produce a PO target after a DO prime than after a baseline prime in the verb overlap condition (β = 0.221, SE = 0.081, $p < .01$). More proficient participants tended to be slower to respond to a PO target after a PO prime than after a baseline prime (β = 0.074, SE = 0.043, $p < .1$). In short, our results showed lexically-dependent facilitatory priming of the DO dative and lexically-dependent inhibitory priming of DO primes on PO targets. There was some indication of an inhibitory priming effect of PO primes on PO targets for more proficient participants.

5.3 Experiment 1B: L2 production (L1 Dutch)

Method

Participants

We recruited an additional 192 participants (110 female, 77 male, 5 other) with Dutch as their L1 and any level of L2 English on the online platform Prolific. Participants were aged between 18 and 66 (mean 30.0, *SD* 11.2). They were paid for their participation.

Materials, design, procedure, coding

The materials, design, procedure and coding were identical to Experiment 1A.

Results

Proficiency

The average LexTALE score over participants was 88.0% (*SD* 8.3), within a range from 60% to 100%. On a scale of 0 (very low) to 10 (very high), participants rated their proficiency in speaking at 7.9 (*SD* 1.3, range 3-10), in listening at 8.9 (*SD* 1.0, range 5-10), in reading at 8.9 (*SD* 1.1, range 5-10) and in writing at 7.8 (*SD* 1.3, range 4-10). We also calculated the average self-rated proficiency scores over the four skills for each participant, which was 8.4 (*SD* 1.0), ranging from 5 to 10. Similar to the other experiments, we included self-rated proficiency as a measure of proficiency in our data analyses.

Choice data

We collected a total of 6,912 observations. Participants produced 4,741 (68.6%) PO datives and 2,074 (30.0%) DO datives. In addition, 97 (1.4%) of the responses were coded as "Other". We excluded the "Other" responses for further analyses. Figure 5.4 shows the proportion of DO and PO responses per prime condition, with and without verb overlap.

The responses were fitted to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer in order to increase convergeability (Powell, 2009). We inserted Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), self-rated proficiency scores, and the interactions as fixed effects. The baseline prime condition and the condition without verb overlap were treated as the reference level. We also included random intercepts for Participant and Item, and random slopes of the fixed effects for

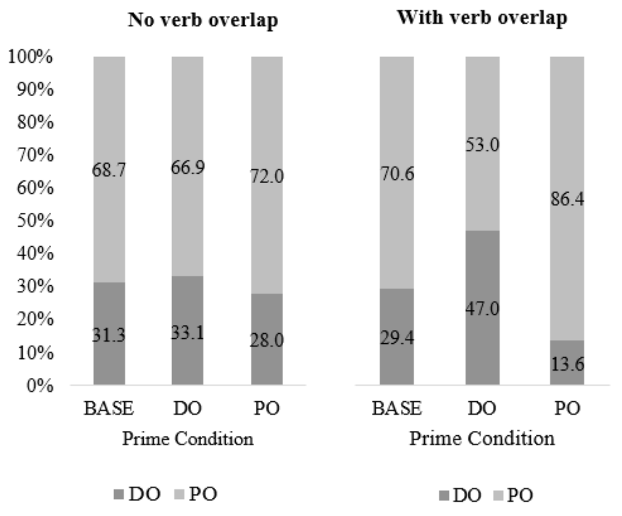


Figure 5.4: Produced structure per prime condition (Dutch group).

Participant and Item. We started from the full model with a maximal random effects structure (Barr et al., 2013) and simplified it until convergence. The final model included random intercepts for Participant and Item and no random slopes. The model output is reported in the appendix (Table A.5.3).

Participants were more likely to produce a PO dative than a DO dative in the baseline condition, as shown by the significant intercept (β = 1.501, SE = 0.286, $p < .001$). The proportion of DO responses after a DO prime was higher than after a baseline prime when there was verb overlap between prime and target (β = -1.140, SE = 0.169, $p < .001$). Participants produced more PO datives after a PO prime than after a baseline prime in the condition without verb overlap (β = 0.330, SE = 0.123, $p < .01$), and this increase was stronger with lexical overlap between prime and target (β = 1.262, SE = 0.186, $p < .001$). The increase in the production of PO datives after a PO prime compared to a baseline prime in the condition with verb overlap tended to be smaller for more proficient participants than for less proficient participants (β = -0.322, SE = 0.173, $p < .1$). Hence, we observed lexically-dependent priming of the DO dative, and lexically-boosted priming of the PO dative. The PO priming effect in the verb overlap condition tended to decrease with increasing proficiency.

Reaction time data

We included 6,688 observations in the analyses for the reaction time data, excluding 97 "Other" responses and 127 outliers. Table 5.2 shows the mean reaction times per prime condition. Figure 5.5 displays the mean reaction times separately for the conditions with and without verb overlap.

Table 5.2: Mean onset latencies (in ms) per prime and target condition (Dutch group).

Target Condition	DO target	PO target
BASE	3355 (1528)	2847 (1463)
DO	3444 (1577)	2952 (1625)
PO	3517 (1677)	2832 (1491)

Note: Standard deviations in parentheses.

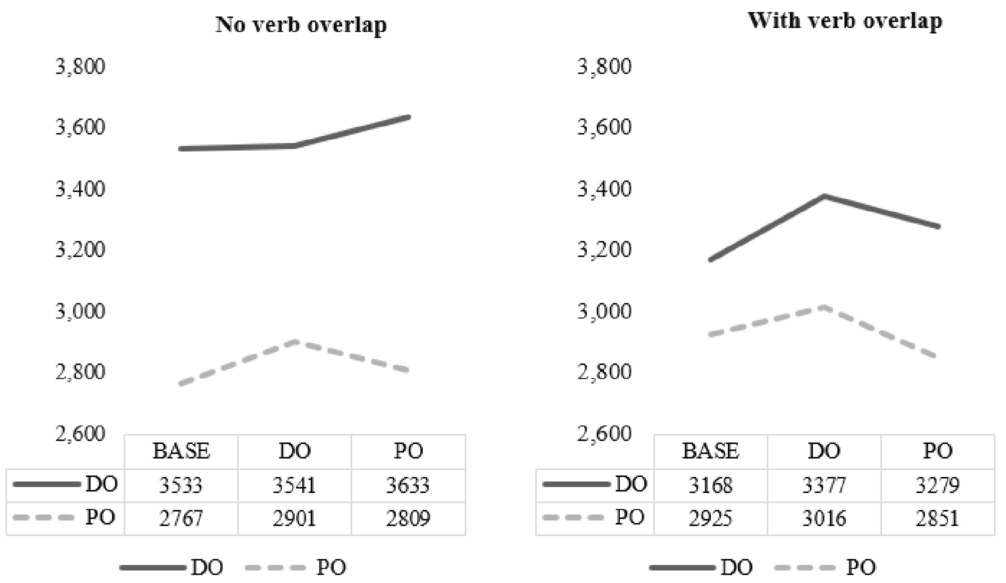


Figure 5.5: Mean onset latencies (in ms) per prime and target condition (Dutch group).

The log-transformed reaction times were fitted to a linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with Target Condition (DO/PO), Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), the self-rated proficiency scores, and the interactions as its fixed factors. The DO target condition, the baseline prime condition and the condition without verb overlap were the reference level. We started from the full model with a maximal random effects structure (Barr et al., 2013) and simplified it until the model converged and had no singularity issues. The final model included random intercepts for Participant and Item and a random slope of Target Condition to Participant. Using the *drop1*-function, we checked whether the four-way interaction contributed significantly to our model. As this was not the case, we removed it from our model. The model output is summarized in the appendix (Table A.5.4); we only mention the most relevant significant results here. We found a main effect of target structure, but the priming effects were only found in interaction with proficiency. These interactions are plotted in Figure 5.6. Priming did not significantly interact with verb overlap between prime and target.

Less proficient participants were faster to respond to a DO target after a DO prime than after a baseline prime, and this effect decreased with increasing proficiency (β = 0.050, *SE*

= 0.021, $p < .05$). In addition, proficiency significantly affected the priming effect of the DO structure on PO targets ($\beta = -0.085$, $SE = 0.022$, $p < .001$). Less proficient participants were slower to produce a PO structure after a DO prime than after a baseline prime, and this effect decreased upon increasing proficiency. To summarize, we found a facilitatory priming effect of DO primes on DO targets as well as an inhibitory effect of DO primes on PO targets, and both effects were independent of lexical overlap and only present in the less proficient participants.

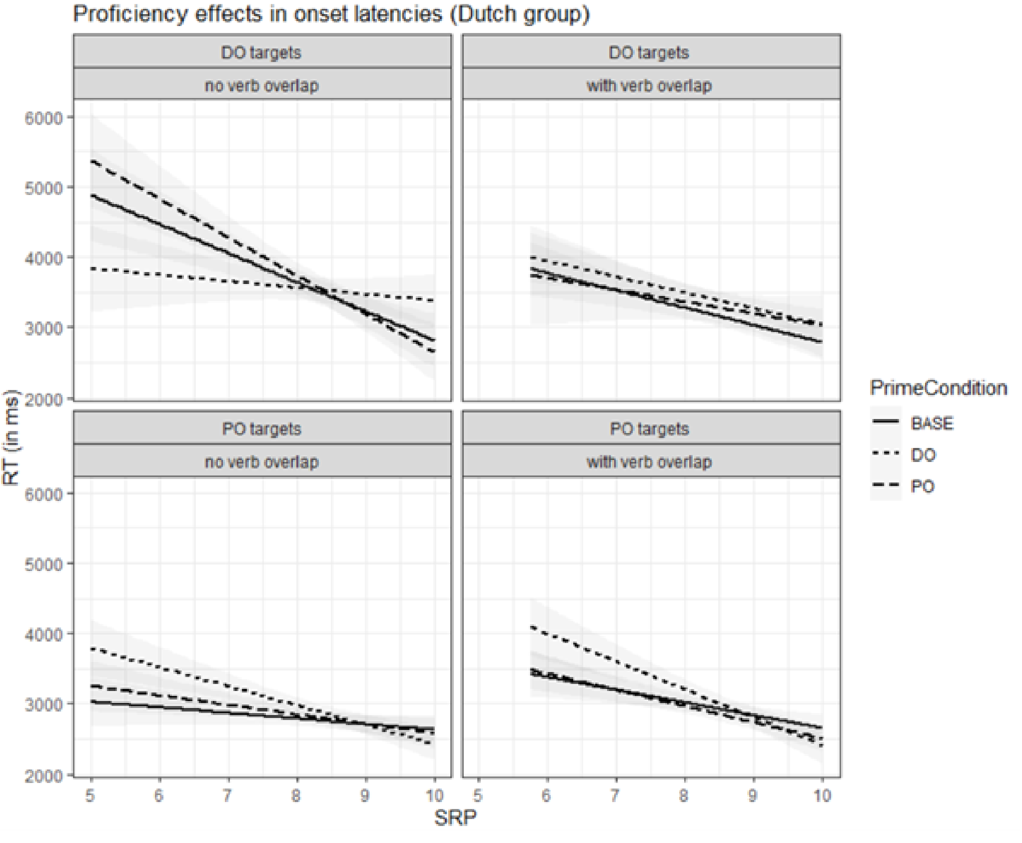


Figure 5.6: Proficiency effects in onset latencies (Dutch group).

5.4 Discussion Experiment 1: L2 production

In line with previous studies, we found a preference for the PO structure in both the Chinese-English and the Dutch-English group in the choice data. In both groups, we found priming of the DO and the PO structure in the presence of verb overlap between prime and target. There was also abstract structural priming of the PO structure, which was nearly significant in the Chinese-English group and reached statistical significance in the Dutch group. We did not find proficiency effects in the Chinese group. In the Dutch group, lexically-dependent priming tended to be stronger for less proficient participants than for more proficient participants.

We also found evidence for priming of the initiation times of producing a response. In the Chinese-English group, the onset latencies of DO targets were lower after a DO prime than after a baseline prime, pointing towards a facilitatory effect of the prime structure on processing. Similar to Chapter 4 and in line with the two-stage competition model (Segaert et al., 2011), we also found an inhibitory effect of priming. Participants were significantly slower to initiate a PO target after a DO prime than after a baseline prime in the verb overlap condition. Similar to the Chinese-English group, the less proficient Dutch-English participants were faster to initiate a DO target and slower to produce a PO target after a DO prime than after a baseline prime. Different from the Chinese-English group, in the Dutch-English group, the effects are present regardless of verb overlap.

The first aim of our study was to investigate whether proficiency effects on priming in production can also be found in the onset latencies. We replicated the proficiency effects in the production choices from Bernolet et al. (2013). In the Dutch group, the priming effect of the PO structure in the presence of verb overlap is stronger for less proficient participants than for more proficient participants. This is in line with the proficiency effects observed in Bernolet et al., suggesting that less proficient participants rely more on item-specific representations. We did not find evidence for stronger abstract priming in more proficient speakers than in less proficient speakers within the bilingual groups. However, abstract structural priming is numerically stronger in the more proficient Dutch-English group than in the less proficient Chinese-English group, consistent with Bernolet et al. [1] Also note that proficiency effects on abstract priming within the L2 may be weaker for ditransitives than for genitives (which was the structural alternation tested by Bernolet et al.). Reanalyzing the data of Schoonbaert et al. (2007), Hartsuiker and Bernolet (2017) argued that in ditransitives, explicit memory also plays a role during abstract structural priming due to similar meanings of different ditransitive verbs. This may explain the lack of a proficiency effect within the bilingual groups. Less proficient and more proficient participants show abstract structural priming of a comparable magnitude, but the locus of priming may be different: in the less proficient participants, priming may be induced by explicit memory, while in the high proficient participants, priming may be due to residual activation of abstract structural representations.

In the onset latencies, we found priming effects in less proficient participants in the Dutch-English group, and the effects disappeared upon increasing proficiency. Crucially, this proficiency effect was independent of lexical overlap, which contrasts with Bernolet et al. (2013). A possible reason may be that according to the two-stage competition model, both residual activation of structural representations and implicit learning play a role during priming, while the predicted proficiency effects follow from a residual activation account. According to an implicit learning account, abstract structural priming may in fact be stronger for less proficient participants than for more proficient participants due to larger prediction error as a consequence of less experience with the L2. The contradictory effect of the interaction between proficiency and residual activation on the one hand and the interaction between proficiency and prediction error on the other hand may be reflected differently between the two measurement types.

[1]Note that the self-rated proficiencies should probably not be compared between the groups, because they might be affected by cultural values. Nevertheless, the objective LexTALE scores also show that on average, the Dutch-English group is more proficient than the Chinese-English group.

5.5 Experiment 2A: L2 comprehension (L1 Chinese)

Method

Participants

We tested 96 participants in each of the four between-participants conditions in our experiment (no-choice vs. free-choice in the target sentences, with vs. without verb overlap between prime and target). We therefore recruited an additional 384 participants (203 male, 176 female, 5 other) with Mandarin Chinese as their L1 of any level of L2 English via WeChat. Participants were aged between 16 and 84[2] (mean: 23.0, *SD*: 7.5). They received monetary compensation.

Paradigm

Similar to Chapter 4, we used the maze task paradigm as a method to measure structural priming in comprehension. The maze task is an alternative method to self-paced reading for studying incremental sentence processing. It has been found that the maze task leads to larger effects and a smaller spillover (effects on words after the critical region) than self-paced reading (N. Witzel et al., 2012) and eye-tracking (J. Witzel & Forster, 2014). In addition, it allows us to collect choice data and reaction time data simultaneously. Participants read sentences word by word. For each word of the sentence, participants choose between a distractor word and the correct continuation of the sentence by pressing a button.

There are several variants of the maze task paradigm. We used the automated maze task (Boyce et al., 2020), which automatically generates distractor words based on a machine-learning model. In addition, we used an adapted version for structural priming studies, as we did in our previous study with L1 speakers (Chapter 4). In this version, participants were provided with a choice between the two syntactic alternatives under investigation. In other words, participants choose whether they read a DO or a PO target sentence. PO and DO sentences have an identical beginning of the sentence and only start to deviate from each other halfway the sentence (e.g., *The waitress passes the* [**boxer** *the cake*] / [**cake** *to the boxer*]). At this critical point, DO sentences have an animate noun to indicate the recipient (*boxer*), whereas PO sentences mention an inanimate noun, which is the object (*cake*). Participants choose between *boxer* and *cake* to determine the continuation of the sentence. For the other words of the sentence, they choose between the correct word and a distractor word. This free-choice variant of the maze task paradigm is illustrated in Figure 5.7.

[2]Although we recruited the participants for the production experiment on the same platform and according to the same criteria as for the comprehension experiment, the tested age range is considerably larger, but the mean age and *SD* are comparable between the experiments.

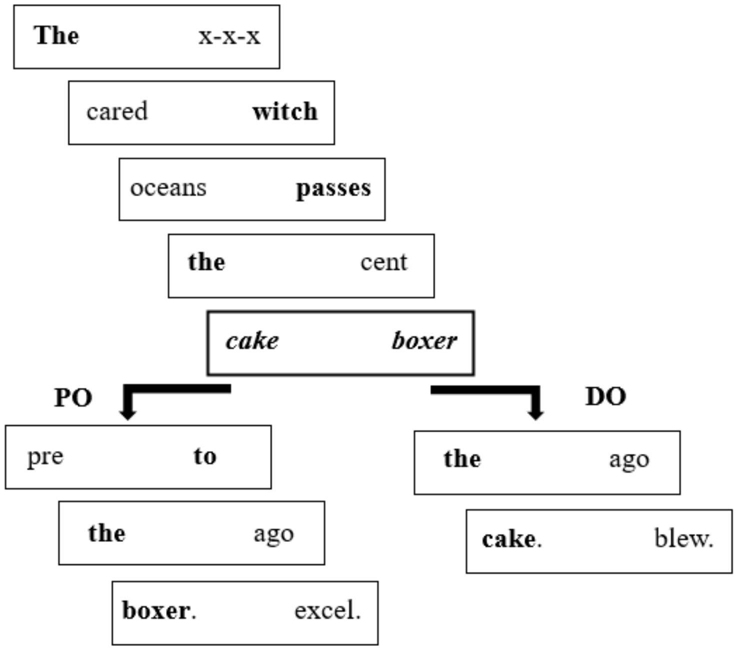


Figure 5.7: Free-choice version of the automated maze task paradigm.

Materials and design

The materials and the design were identical to Chapter 4. The 36 prime-target pairs and the 144 filler sentences were identical to those in Experiment 1. For the target sentences, we had a DO and a PO variant. For every word of each sentence, we had distractor words created by the automated MAZE tool (`https://vboyce.github.io/Maze/`, made available by Boyce et al., 2020). Each sentence started with *The* on the left and *x-x-x* on the right. The remainder of the words and distractors were distributed evenly, so that half of the words appeared on the left on the screen and half of the words were on the right side of the screen.

We had four different versions in order to create a 2x2 factorial design, manipulating verb overlap and target structure choice. In the verb overlap condition, there was verb overlap between the prime and the target sentence. In the condition without verb overlap, the verb was always different between prime and target. In addition, we had a "free-choice" condition and a "no-choice" condition. In the free-choice condition, participants chose between an animate noun (for a DO sentence) and an inanimate noun (for a PO sentence). To illustrate, for the sentence *The waitress passes the* [*boxer the cake*] / [*cake to the boxer*], participants in the free-choice condition had to choose between *boxer* and *cake*. The sentence would then continue as either a DO or a PO, depending on their choice. In the no-choice condition, half of the target sentences appeared as a DO sentence and half of the target sentences were PO sentences.

The two free-choice versions (with and without verb overlap) had three lists each, to ensure that each target item was preceded by all three prime conditions across the lists. Within each list, there was an even distribution of the prime conditions. For the two no-choice versions, we had six lists, so that every target item appeared as a DO and as a PO after the three prime conditions across the lists.

We programmed the experiment in PsychoPy (Peirce et al., 2019) and ran it using its online platform Pavlovia. Identical to the production experiments, immediately after the experiment participants completed the LexTALE test (Lemhöfer & Broersma, 2012) and the survey asking for their self-rated proficiency. In this survey, we also had an open question on what they thought was the purpose of the experiment, which especially served to monitor whether the choice between two plausible words in the free-choice version revealed the aim of the experiment.

Procedure

During the priming experiment, participants would read the prime and target sentences word by word. For each word of the sentence, participants were presented with the correct word and the distractor word. They pressed the key A to choose the word on the left of the screen or the key L to choose the word on the right of the screen. If they chose the distractor word instead of the correct word, they were asked to try again, until they chose the correct word (the forgiving maze paradigm, cf. Boyce et al., 2020). In this way, we were certain that participants would always read the entire prime sentence before reading the target sentence, preserving the balance of the three prime conditions. On average, a session took about 35 minutes.

Coding

For each word, the timing of the key press was stored. We removed any outliers, which were defined as reaction times deviating more than three times the standard deviation from the mean. We measured the reaction times for the disambiguating word (either the animate or the inanimate noun, which is the critical word at which the DO and the PO structure start to deviate) as well as for each word in the spillover area, which is the remainder of the sentence. The spillover area comprised two words for DO sentences (e.g., *The waitress passes the* **boxer** [*the cake*]) and three words for PO sentences (e.g., *The waitress passes the* **cake** [*to the boxer*]). The reason for having these different measures is that structural priming effects in comprehension have been reported in the disambiguating area as well as in the spillover region (e.g., Tooley & Bock, 2014; Wei et al., 2016). Since the data were not normally distributed, we log-transformed all the reaction times (log10(x)), after which the Q-Q plot of the plotted reaction times showed a more normal distribution.

For the target sentences in the free-choice condition, we registered the choice for either the DO or the PO structure in addition to the reaction times. We calculated the priming effect by comparing the proportion of DO or PO choices after a DO or PO prime respectively to the proportion of DO or PO choices after a baseline prime.

Results

Participants had an average LexTALE score of 56.9% (*SD* 12.1), ranging from 36.25% to 96.25%. On a scale of 0 (very low) to 10 (very high), they rated their proficiency in speaking at 5.2 (*SD* 2.0, range 0-10), in listening at 5.5 (*SD* 2.2, range 0-10), in reading at 6.5 (*SD* 1.8, range 0-10), and in writing at 5.9 (*SD* 2.0, range 0-10). We also calculated their overall self-rated proficiency by averaging the self-rated proficiency scores of the four skills for each participant, which was 5.8 (*SD* 1.8), ranging from 0.5 to 10.

Choice data

We collected 6,912 observations in the free-choice condition. Participants chose to read a PO structure for 4,054 (58.7%) of the target sentences and chose to read a DO target sentence in 2,858 (41.3%) of the cases. Figure 5.8 shows the choice for a DO or a PO structure per prime condition.

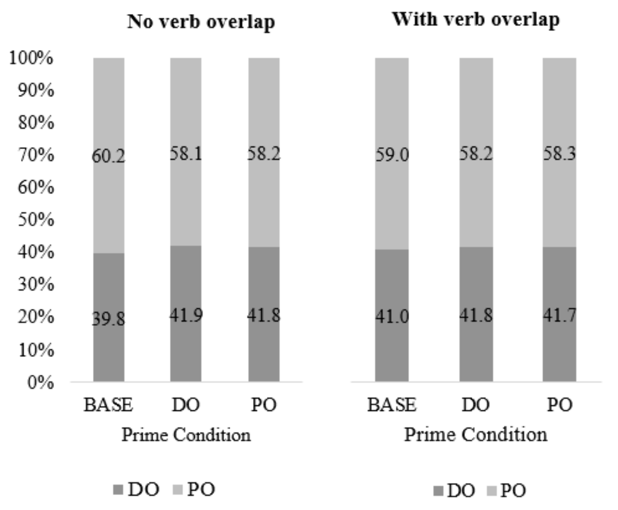


Figure 5.8: Chosen structure per prime condition (Chinese group).

We fitted the responses to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer in order to increase convergeability (Powell, 2009). Our model included Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), and self-rated proficiency and their interactions as fixed factors. The baseline prime condition and the condition without verb overlap were inserted as the reference level. We scaled and centered the self-rated proficiency scores to increase convergeability. We started with the maximal random effects structure conforming to Barr et al. (2013), adding random slopes and random intercepts for Participants and Items. We simplified the model until it converged without any singularity issues. The final model included random intercepts for Participants and Items and no random slopes. Using the *drop1*-function, we checked whether the three-way interaction contributed significantly to our model. As this was not the case, we removed it from our model. The model output is reported in the appendix (Table A.5.5). More proficient participants responded faster than less proficient participants (β = -0.188, SE = 0.073, $p < .05$) in the condition without verb overlap, and there were no significant priming effects or interactions between proficiency and priming.

Reaction time data

Taking together the free-choice condition and the no-choice condition, we had 13,752 observations for the reaction time data before exclusion of the outliers. We analyzed the results separately for DO targets (6,278 observations) and for PO targets (7,474 observations), as we are interested in the priming effects within each target structure. In addition, a practical reason for running separate analyses for each structure is that PO target sentences have one extra word in the spillover area.

For each target structure and for each word (i.e., the disambiguating word and the spillover words), we fitted a linear mixed model (*R*-package *lme4*, D. Bates et al., 2015). Our model included Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), and self-rated proficiency and their interactions as fixed factors. The baseline prime condition and the condition without verb overlap were included as the reference level. We scaled and centered the self-rated proficiency scores to increase convergeability. We started with the maximal random effects structure conforming to Barr et al. (2013), adding random slopes and random intercepts for Participants and Items. We simplified the model until it converged without any singularity issues. In addition, for each model we removed the non-significant interactions in order to interpret the main effects. Full model outputs are provided in the appendix (Table A.5.6-A.5.12).

DO targets

Table 5.3 shows the mean reaction times per prime condition for the disambiguating word of the DO target sentences and the two words in the spillover region. Figure 5.9 plots the reaction times (in ms) per prime condition, separately for the condition with and without verb overlap. The interaction between proficiency and the priming effects is plotted in Figure 5.10.

Table 5.3: Mean onset reaction times (in ms) per prime condition for DO targets (Chinese group).

Prime Condition	Disambiguating word	Spillover word 1	Spillover word 2
BASE	791 (831)	593 (792)	759 (1013)
DO	787 (823)	558 (526)	748 (1156)
PO	808 (926)	593 (803)	748 (964)

Note: Standard deviations in parentheses.

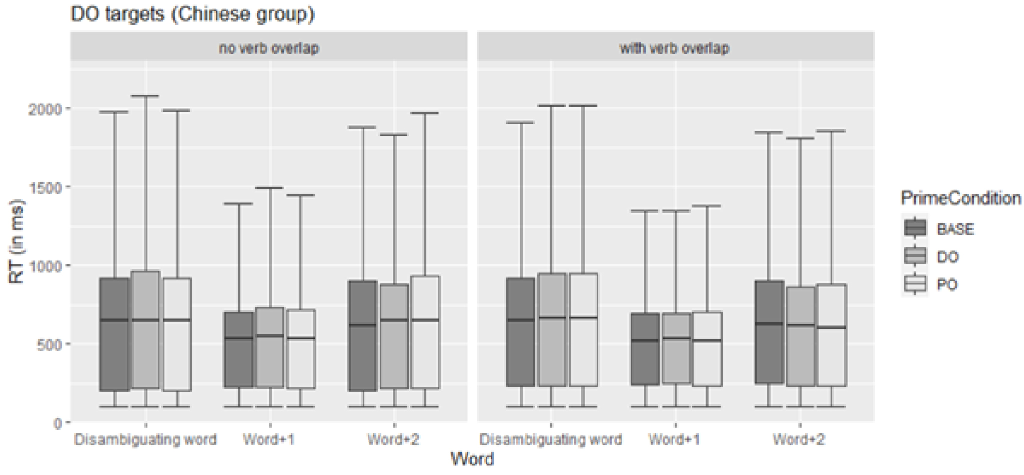


Figure 5.9: Box-plotted reaction times for DO targets per prime condition (Chinese group).

In the first word of the spillover area, participants in the condition with verb overlap between prime and target had higher reaction times to a DO target after a PO prime ($\beta = 0.050$, $SE = 0.024$, $p < .05$ than after a baseline prime). Although numerically the effect was only present in the more proficient participants and not in the less proficient participants, it did not interact significantly with proficiency. In the disambiguating area, there was a significant interaction between the DO prime condition, verb overlap, and proficiency. Less proficient participants were faster to respond to a DO target after a DO prime than after a baseline prime in the condition with verb overlap. This effect disappeared with increasing proficiency ($\beta = 0.081$, $SE = 0.028$, $p < .01$). In brief, we found lexically-dependent priming of the DO dative and a lexically-dependent inhibitory priming effect of PO primes on DO targets. The first effect decreased significantly upon increasing proficiency.

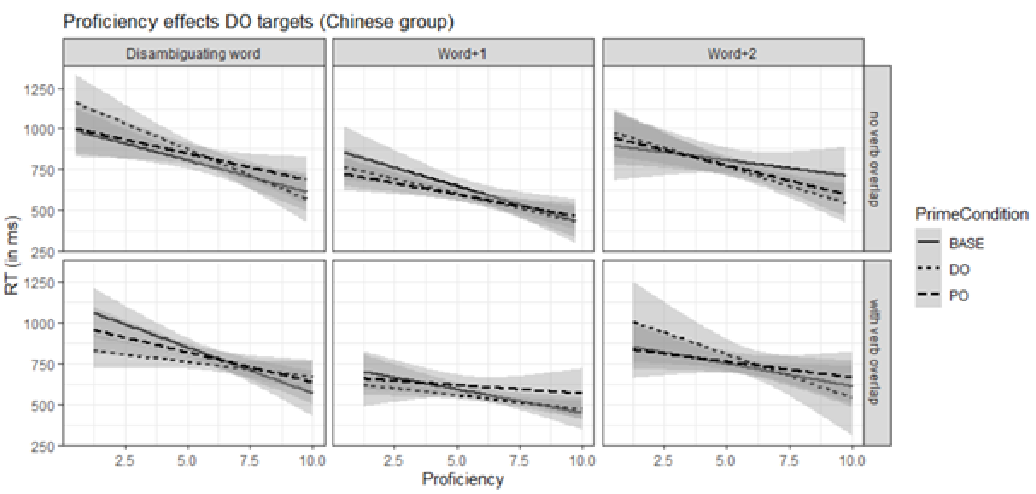


Figure 5.10: Proficiency effects in reaction times on DO targets (Chinese group).

PO targets

Table 5.4 shows the mean reaction times per prime condition for the disambiguating word of the DO target sentences and the two words in the spillover region. Figure 5.11 plots the reaction times (in ms) per prime condition, separately for the condition with and without verb overlap. The proficiency effects on the reaction times and their interactions with the prime condition are plotted in Figure 5.12.

Table 5.4: Mean onset reaction times (in ms) per prime condition for PO targets (Chinese group).

Prime Condition	Disambiguating word	Spillover word 1	Spillover word 2	Spillover word 3
BASE	801 (700)	572 (559)	583 (564)	831 (930)
DO	848 (900)	588 (641)	597 (585)	825 (873)
PO	831 (837)	578 (820)	599 (787)	845 (1014)

Note: Standard deviations in parentheses.

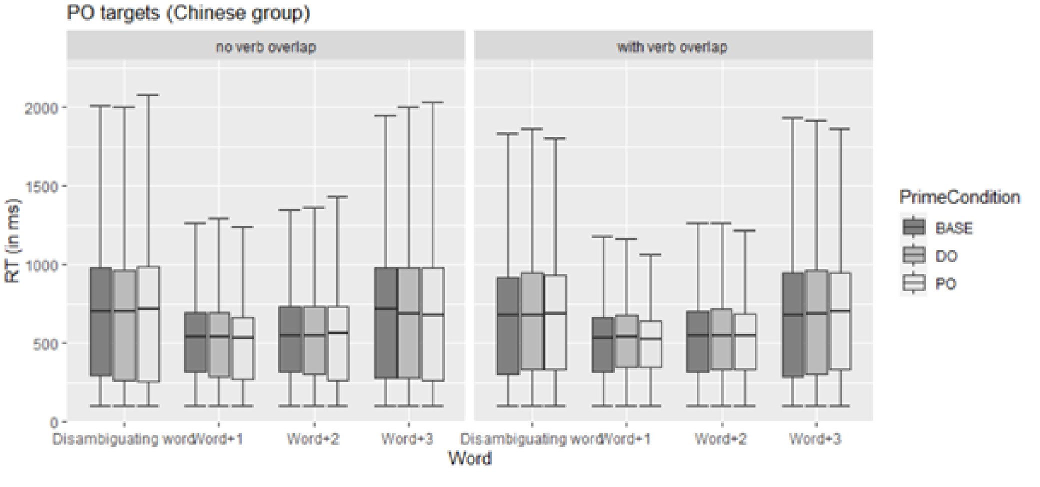


Figure 5.11: Box-plotted reaction times for PO targets per prime condition (Chinese group).

In the first word of the spillover area, participants displayed a tendency towards higher reaction times after a DO prime than after a baseline prime, but in the verb overlap condition only (β = -0.040, *SE* = 0.021, $p < .1$). In the second word of the spillover region, this effect was statistically significant, regardless of verb overlap or proficiency (β = 0.022, *SE* = 0.010, $p < .05$). In the final word of the PO target sentence, participants showed significantly higher reaction times after a PO prime than after a baseline prime in the condition with verb overlap (β = 0.055, *SE* = 0.025, $p < .05$). In the disambiguating area, participants who were less proficient tended to be slower to respond to a PO target after a DO prime than after a baseline prime if there was verb overlap between prime and target. This effect disappeared upon increasing proficiency (β = -0.044, *SE* = 0.024, $p < .1$). In short, we found a lexically-dependent inhibitory effect of DO primes on PO targets, and this effect tended to decrease upon increasing proficiency.

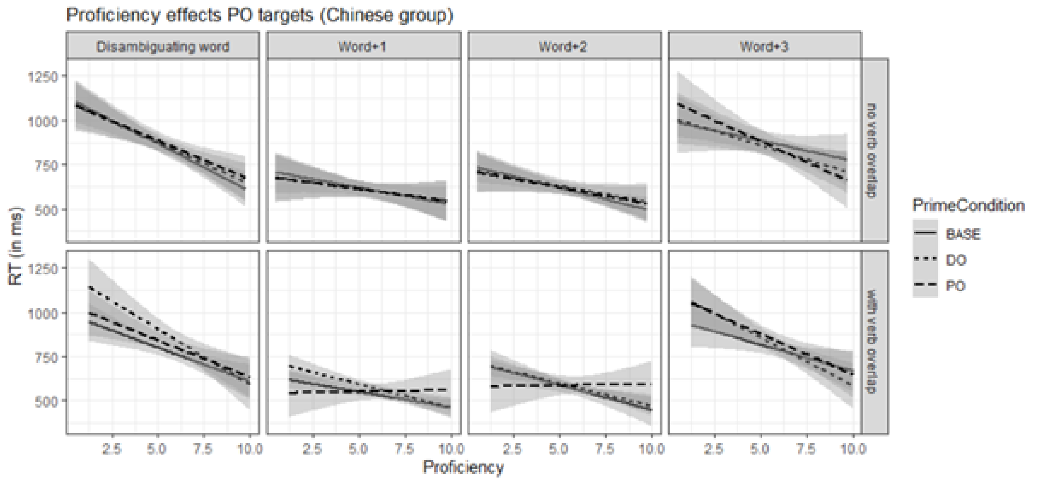


Figure 5.12: Proficiency effects in reaction times on PO targets (Chinese group).

5.6 Experiment 2B: L2 comprehension (L1 Dutch)

Method

Participants

We recruited 384 participants (271 female, 104 male, 9 other) with Dutch as their L1 and any level of L2 English on public social media groups. They were aged between 13 and 65 (mean 23.5, *SD* 6.2). Participants received monetary compensation for their participation.

Materials, design, procedure, coding

The materials, design, procedure, and coding were identical to Experiment 2A.

Results

Proficiency

Participants had an average LexTALE score of 80.0% (*SD* 11.8) within a range of 32.5% to 100%. On a scale of 0 (very low) to 10 (very high), they rated their proficiency in speaking at 7.2 (*SD* 1.3, range 3-10), in listening at 8.3 (*SD* 1.2, range 4-10), in reading at 8.3 (*SD* 1.2, range 5-10), and in writing at 7.1 (*SD* 1.5, range 2-10). The average self-rated proficiency scores over the four skills for each participant was 7.7 (*SD* 1.1), ranging from 4 to 10. Similar to the other experiments, we included self-rated proficiency as a measure of proficiency in our data analyses.

Choice data

We collected 6,912 observations in the free-choice condition. Participants chose to read a PO structure for 3,474 (50.3%) of the target sentences and chose to read a DO target sentence in 3,438 (49.7%) of the cases. Figure 5.13 shows the choice for a DO or a PO structure per prime condition.

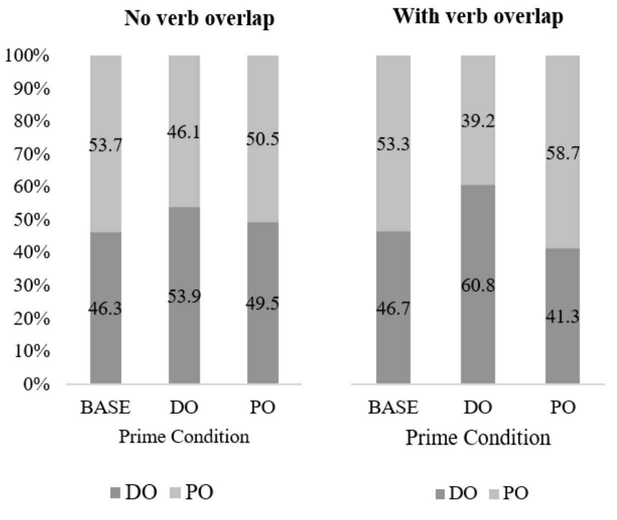


Figure 5.13: Chosen structure per prime condition (Dutch group).

We fitted the responses to a generalized linear mixed model (*R*-package *lme4*, D. Bates et al., 2015) with a BOBYQA optimizer in order to increase convergeability (Powell, 2009). Our model included Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), and self-rated proficiency and their interactions as fixed factors. The baseline prime condition and the condition without verb overlap were inserted as the reference level. We scaled and centered the self-rated proficiency scores to increase convergeability. We started with the maximal random effects structure conforming to Barr et al. (2013), adding random slopes and random intercepts for Participants and Items. We simplified the model until it converged without any singularity issues. The final model included random intercepts for Participants and Items and a random slope of the factors Prime Condition to Participants. The model output is summarized in the appendix (Table A.5.13).

Participants chose more DO targets after a DO prime than after a baseline prime in the condition without verb overlap (β = -0.35, SE = 0.10, $p < .001$), and this increase was significantly stronger if there was lexical overlap between prime and target (β = -0.30, SE = 0.14, $p < .05$). More proficient participants were more likely to choose a DO target after a DO prime if there was verb overlap than less proficient participants (β = -0.27, SE = 0.13, $p < .05$). Finally, participants were more likely to choose a PO target after a PO prime than after a baseline prime in the presence of verb overlap (β = 0.38, SE = 0.13, $p < .01$). In sum, we found lexically-boosted priming of the DO, and this effect increased upon increasing proficiency in the verb overlap condition. In addition, we found lexically-dependent PO priming.

Reaction time data

Taking together the free-choice condition and the no-choice condition, we had 13,752 observations for the reaction time data before exclusion of the outliers. Similar to Experiment 1, we analyzed the results separately for DO targets (6,894 observations) and for PO targets (6,930 observations). Per target structure and for each word, we fitted a linear mixed model (*R*-package *lme4*, D. Bates et al., 2015). Our model included Prime Condition (baseline/DO/PO), Verb Overlap (no/yes), and self-rated proficiency and their interactions as fixed factors. The baseline prime condition and the condition without verb overlap were inserted as the reference level. We scaled and centered the self-rated proficiency scores to increase convergeability. We started with the maximal random effects structure conforming to Barr et al. (2013), adding random slopes and random intercepts for Participants and Items. We simplified the model until it converged without any singularity issues. In addition, for each model we removed the non-significant interactions in order to interpret the main effects. Full model outputs are provided in the appendix (Table A.5.14-A.5.20).

DO targets

Table 5.5 shows the mean reaction times per prime condition for the disambiguating word of the DO target sentences and the two words in the spillover region. Figure 5.14 plots the reaction times (in ms) per prime condition, separately for the condition with and without verb overlap. The influence of proficiency on the priming effects is plotted in Figure 5.15.

Table 5.5: Mean onset reaction times (in ms) per prime condition for DO targets (Dutch group)

Prime Condition	Disambiguating word	Spillover word 1	Spillover word 2
BASE	930 (371)	719 (223)	788 (230)
DO	903 (336)	710 (212)	801 (235)
PO	936 (360)	730 (227)	803 (232)

Note: Standard deviations in parentheses.

In the disambiguating area, participants were faster to read a DO target after a DO prime (β = -0.025, *SE* = 0.009, $p < .01$). In the first word of the spillover area, participants responded significantly slower to a DO target after a PO prime than after a baseline prime (β = 0.018, *SE* = 0.006, $p < .05$). There were no significant interactions between the fixed effects in the disambiguating area. However, in the second word of the spillover area, we found a three-way interaction between the DO prime condition, verb overlap, and proficiency. More proficient participants were significantly faster to read the final word of a DO target after a DO prime than after a baseline prime in the presence of verb overlap than less proficient participants (β = -0.034, *SE* = 0.016, $p < .05$). In sum, we found lexically-independent facilitatory priming of the DO dative. This effect increased upon increasing proficiency, but only in the verb overlap condition. There was some evidence for lexically-independent inhibition of PO primes on DO targets in the spillover area.

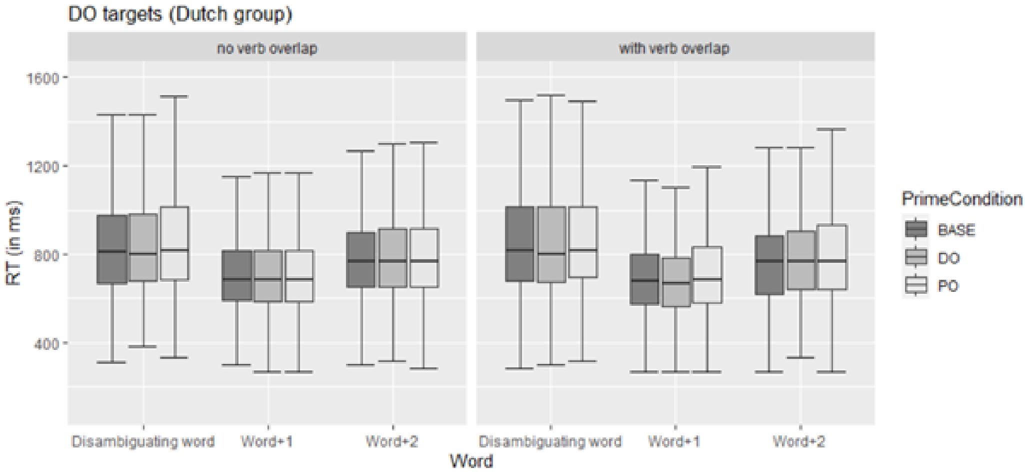


Figure 5.14: Box-plotted reaction times for DO targets per prime condition (Dutch group).

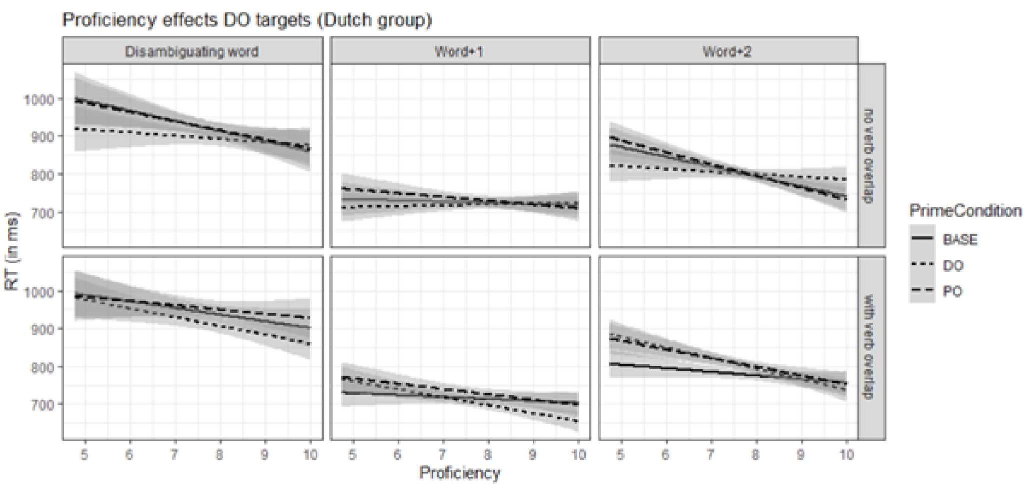


Figure 5.15: Proficiency effects in reaction times on DO targets (Dutch group).

PO targets

Table 5.6 shows the mean reaction times per prime condition for the disambiguating word of the PO target sentences and the two words in the spillover region. Figure 5.16 plots the reaction times (in ms) per prime condition, separately for the condition with and without verb overlap.

In the disambiguating area, PO targets were read slower after a DO prime than after a baseline prime if there was verb overlap between prime and target ($\beta = 0.044$, $SE = 0.019$, $p < .05$). In the first word of the spillover area, participants were slower to respond to a PO target after a DO prime than after a baseline prime ($\beta = 0.021$, $SE = 0.007$, $p < .001$). In the final word of the PO target sentence, participants tended to respond slower after a PO prime than after a baseline prime in the verb overlap condition ($\beta = 0.027$, $SE = 0.014$, $p <$

.1). To sum up, there was a lexically-dependent inhibitory priming effect of the DO prime on PO targets. This priming effect did not interact with proficiency (see Figure 5.17).

Table 5.6: Mean onset reaction times (in ms) per prime condition for PO targets (Dutch group).

Prime Condition	Disambiguating word	Spillover word 1	Spillover word 2	Spillover word 3
BASE	906 (358)	643 (194)	686 (193)	764 (221)
DO	921 (363)	654 (196)	684 (194)	771 (220)
PO	917 (365)	635 (171)	683 (180)	779 (217)

Note: Standard deviations in parentheses.

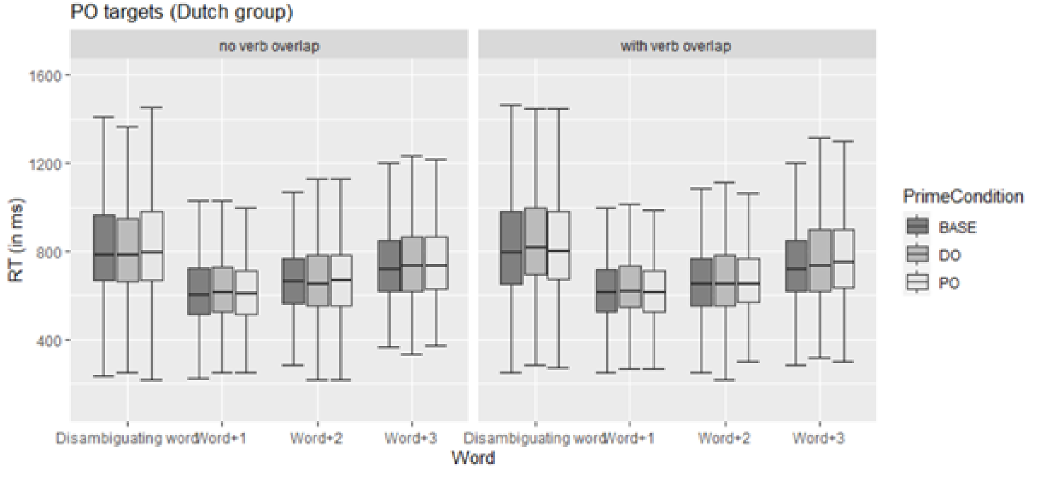


Figure 5.16: Box-plotted reaction times for PO targets per prime condition (Dutch group).

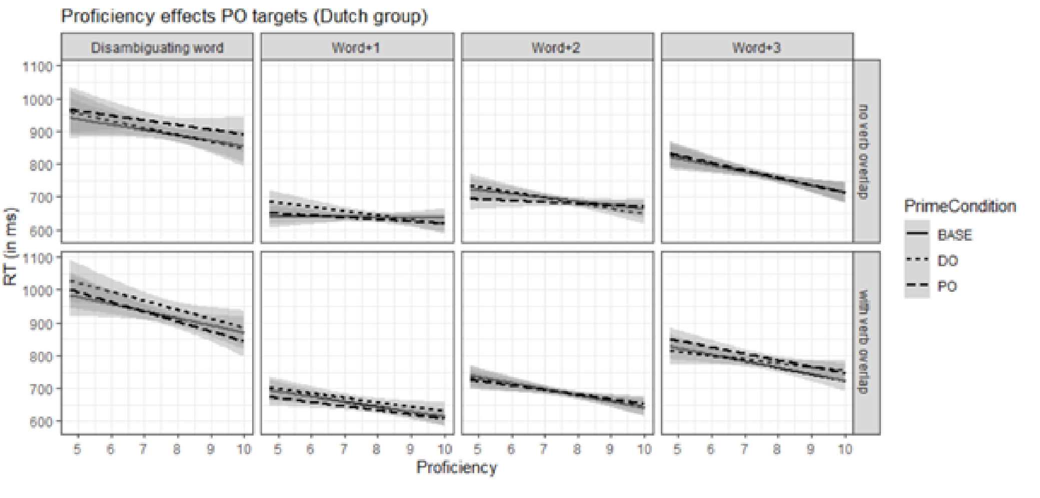


Figure 5.17: Proficiency effects in reaction times on PO targets (Dutch group).

5.7 Discussion Experiment 2: L2 comprehension

Our first aim was to investigate whether structural priming effects in comprehension interact with proficiency. In the choice data of the comprehension experiment, we did not find any significant priming effects in the Chinese-English group. Dutch-English bilinguals chose more often to read a PO target after a PO prime than after a baseline prime with verb overlap. In addition, they more often chose a DO target after a DO prime than after a baseline prime, and this effect was stronger in the condition with verb overlap than in the condition without verb overlap. In the condition with verb overlap, DO priming increased upon increasing proficiency.

In general, more proficient participants have lower reaction times than less proficient participants. In the Chinese-English group, the DO prime structure had a facilitatory effect on processing of DO targets in the verb overlap condition, and this effect was stronger for less proficient participants than for more proficient participants. DO targets were processed more slowly after a PO prime with verb overlap, but this effect was significant in the spillover region only. We also found an inhibitory effect in PO targets: PO targets were processed slower after a DO prime. In the disambiguating area, this effect was only significant in the condition with verb overlap, and it was stronger for less proficient speakers than for more proficient speakers.

In the Dutch-English group, we also found faster processing of DO targets after a DO prime than after a baseline prime on the disambiguating word. Although this effect was independent of verb overlap, the effect was stronger for more proficient participants than for less proficient participants in the verb overlap condition. In the spillover region, participants were slower to respond to a DO target after a PO prime. We observed an inhibitory effect of the DO prime on PO targets in both the disambiguating word and the spillover area if there was verb overlap between prime and target.

All in all, there is clear evidence for structural priming in comprehension, and priming can be observed in both the choice data and the reaction times. We found evidence for facilitatory priming as well as an inhibitory effect of priming on the alternative structure. The observed patterns and the magnitude of priming do not only depend on whether there is lexical overlap between prime and target, but also on the bilingual group, which differ in proficiency. Also within the bilingual groups, the structural priming effects interact with proficiency.

The second aim of the current study was to test whether the proficiency effects interact with structural priming in comprehension as described in the developmental account for L2 syntax (Hartsuiker & Bernolet, 2017), that is, in a similar fashion to production. The findings show that proficiency modulated the magnitude of priming in comprehension as well. In the reaction times of the Chinese-English group, we found stronger priming in the presence of verb overlap for less proficient speakers than for more proficient speakers. In the choice data of the Dutch-English group, we found stronger lexically-dependent priming for more proficient participants than for less proficient participants. The latter finding contradicts the predictions of the developmental account, as that account predicts stronger lexically-dependent priming for less proficient participants than for more proficient participants. Nevertheless, similar results were obtained in Chapter 1, where we found that more proficient participants produced more passives rather than actives after a prime with lexical overlap than less proficient participants. One reason for this discrepancy may be related to the fact that modulations by proficiency could be non-linear, especially when taking into account data of very low proficient participants. Importantly, Hartsuiker and Bernolet (2017, also see Bernolet & Hartsuiker, 2018) argued that low proficient L2 learners show lexically-dependent priming due to explicit memory (a copy-edit strategy) instead of residual activation. A learner needs to have some representation of the structure in order to have residual activation of the structure. As such, the proficiency effects on priming may have an inverse U-shape. Medium proficient L2 learners may display stronger lexically-dependent priming than low proficient L2 learners due to the absence of residual activation in the latter group. At the same time, lexically-dependent priming may also be stronger in medium proficient L2 learners than in high proficient L2-learners as the former rely more on the activation of item-specific representations. Indeed, the latter group was investigated by Bernolet et al. (2013), on which the predictions from the developmental account were based. The Chinese-English participants were considerably less proficient than the participants tested by Bernolet et al., suggesting that our observed proficiency effects (i.e., increasing lexically-dependent priming upon increasing proficiency) are associated with the difference between low proficient and medium proficient participants.

5.8 General discussion

The aim of the current study was twofold. First, we tested whether the proficiency effects on structural priming as observed in the choice data of Bernolet et al. (2013) in production can also be found in the onset latencies of the produced responses. Second, we investigated whether similar proficiency effects can be observed in comprehension, both in choice data and in reaction time data.

We tentatively replicated the proficiency effects as observed in Bernolet et al. (2013) in the choice data of the production experiments, which is also the modality and the measurement type in which the proficiency effects in Bernolet et al. were found. On the one hand, we found some evidence for stronger abstract structural priming of choices in more proficient L2 speakers than in less proficient L1 learners. More specifically, abstract structural priming was stronger in the more proficient Dutch-English participants than in the less proficient Chinese-English participants. On the other hand, we observed stronger lexically-dependent priming in less proficient L2 learners than in more proficient L2 speakers, as there was a tendency towards stronger lexically-dependent priming of the PO structure in less proficient Dutch-English participants than in the more proficient participants.

However, these paradoxical proficiency effects did not fully extend to the onset latencies in production, nor to comprehension. Although we did find stronger priming in the verb overlap condition for less proficient participants than for more proficient participants in the Chinese-English reaction times in comprehension, other observed proficiency effects did not follow the expected polarity of the correlation between priming and proficiency. In the onset latencies in production in the Dutch-English group, we found stronger priming in less proficient participants than in more proficient participants, and this proficiency effect was independent of verb overlap. In the comprehension experiment with the Dutch-English participants, priming of choices was stronger for more proficient participants than for less proficient participants in the condition with verb overlap between prime and target. Even though it seems that proficiency modulates priming in production as well as in comprehension both in the choice data and the reaction times, the exact mechanisms are therefore unclear.

The results suggest that the developmental account of L2 syntax (Hartsuiker & Bernolet, 2017) should be refined in terms of its description of the priming effects that occur at each stage of development as a consequence of the interaction between proficiency and the different mechanisms of priming, including residual activation and implicit learning. First, lexically-dependent priming may be stronger for less proficient participants than for more proficient participants only when comparing medium proficient learners to high proficient learners, but not if one takes into account very low proficient learners. Indeed, the developmental account assumes that the locus of lexically-dependent priming is different for very low proficient learners, relying on a copy-edit strategy (that is, explicit memory) or L1 transfer, and for medium proficient learners, showing priming of item-specific syntactic representations. Since residual activation only starts to play a role for medium proficient learners, while priming due to explicit memory may still play a role both in medium proficient learners, lexically-dependent priming may in fact increase upon increasing proficiency in the earlier stages of development. Once abstract structural representations are developed, learners may rely less on explicit memory, implying that lexically-dependent priming may only decrease at a higher level of proficiency.

Second, the developmental account is based on a residual activation account of priming, while a hybrid model of priming including both residual activation and implicit learning presumably is more accurate to explain priming effects (in the L1) (Reitter et al., 2011; Segaert et al., 2011). While residual activation of abstract structural representations and, consequently, residual activation-based priming may be stronger in more proficient participants than in less proficient participants due to stronger structural representations, the implicit learning account assumes larger prediction error and thus larger priming in less proficient participants than in more proficient participants also in the absence of lexical overlap. If one assumes a hybrid model of priming in which residual activation and implicit learning both play a role, one thus predicts contradictory effects of proficiency on abstract structural priming, and more research is needed to understand under which circumstances either mechanism gets the upper hand in a particular modality, context, or experimental setting.

While an interplay between different priming mechanisms (i.e., explicit memory, residual activation, and implicit learning) may account for the fact that the polarity of the correlation between proficiency and the magnitude of both lexically-dependent and abstract priming is not consistent, it is not fully understood why the interaction between priming and proficiency differs between modality and measurement type even within a particular group of L2 speakers. It is especially not clear why we did not find any significant priming effects in comprehension for the choice data in the Chinese-English participants. A language-specific effect is unlikely, given that X. Chen et al. (2022) observed abstract structural priming in L1 comprehension in Chinese, suggesting that the Chinese-English participants use predictive processing of syntactic structures at least in their L1. In addition, the reaction times in our comprehension data show that the Chinese-English participants have predictions on the upcoming structure based on the prime structure, implying that the absence of priming in the choice data cannot be attributed to an absence of structural representations for English ditransitives.

A potential reason may be that priming in production may be stronger than priming in comprehension. In production, priming does not only occur during the selection stage, but also during the planning stage, which may lead to cumulative effects. On top of that, the magnitude of priming is also affected by the relative frequency of structures, and the relative frequency of the DO and the PO differs between the production experiment and the comprehension experiment. In the production experiment, there is a clear PO preference (although the onset latencies in the Chinese-English group are overall lower for the DO than for the PO), whereas in the comprehension experiment the two structures are more evenly distributed. This is in line with what was found in L1 speakers of English (Chapter 4). As a result, the gap between the activation level of the DO and the activation level of the PO may have been larger in production than in comprehension. Consequently, competition between structures during the selection stage may have been easier to resolve in production than in comprehension.

As such, processing in comprehension may have put a higher load on the L2 speakers than structure selection during production, leaving less room for predictive processing. This might especially have been the case when participants needed to make a choice between the DO and the PO during reading. Not only did they have to select the noun that best matched their predictions, but they also needed to suppress a correct but unpredicted alternative. Given that the Chinese-English participants were considerably less proficient than the Dutch-English participants, this may explain why we were not able to observe priming effects in the choice data in comprehension in the former group, while the reaction time data imply that some predictive processing still took place.

5.9 Conclusion

In our study, we investigated whether the L2 proficiency effects on structural priming as observed in the choice data of Bernolet et al. (2013) in production are also found in the onset latencies of the produced responses, and whether similar proficiency effects can be observed in structural priming in comprehension. We replicated the proficiency effects as observed in Bernolet et al. in the choice data of the production experiment, showing a positive correlation between proficiency and abstract structural priming, and a negative correlation between proficiency and lexically-dependent priming. Although proficiency also modulated priming in the production onset latencies in comprehension, the observed polarity of the correlation between priming and proficiency was not consistent. The complex data pattern suggests that there may be an interplay between explicit memory, residual activation, implicit learning and the relative frequency of structures, which determines the magnitude of priming at a particular stage of L2 development. Given the complexity of this issue, a possible future direction might be computational modelling of L2 priming effects at different levels of proficiency in order to better understand the interplay between the different factors.

General discussion and conclusion

In this thesis, we studied the mental representation of syntactic structures at different stages of L2 learning. More specifically, in two parts, we looked into two different aspects of the developmental account of L2 syntax (Hartsuiker & Bernolet, 2017) that needed further elaboration. The aim of the first part was to investigate the L2 representations at the more extreme sides of the L2 proficiency continuum, namely in beginning learners and native-like L2 speakers. In the second part, we aimed to answer the question whether the developmental account, which was proposed on the basis of evidence from production, also applies to comprehension. In this general discussion, we will take together the results from both parts in order to consider the developmental account and to what extent it should be adapted on the basis of our findings.

The developmental account was based on the bilingual lexical-syntactic model (Hartsuiker et al., 2004), which in turn is rooted in the residual activation account of L1 priming (Pickering & Branigan, 1998). However, more recent studies acknowledge that structural priming in general (i.e., in the L1) is presumably best explained by a hybrid account, incorporating both residual activation and implicit learning (Reitter et al., 2011; Segaert et al., 2011). These developments in modelling L1 priming should be extended to L2 priming and to the developmental account.

As the way how L1 priming is modelled has consequences for the modeling of L2 priming and thus for the developmental account of L2 syntax, we will revisit the developmental account in a bottom-up fashion. First, we will address what our findings imply with regard to models of L1 priming. Then we will discuss the relevant results with regard to the model of L2 priming at the endpoint of learning, that is, at the final stage of the developmental account. Finally, we will consider the implications of our results in the light of the developmental account of L2 syntax.

A model of L1 priming: shared representations in production and in comprehension

In Chapter 3 and 4, we tested structural priming in production as well as in comprehension in the L1 to investigate to what extent priming is similar between modalities. It has been suggested that the abstract structural representations that are engaged during structural priming are the same for production and for comprehension (e.g., Giglio et al., 2022; Indefrey, 2018; Kempen et al., 2012; Pickering & Garrod, 2013). Therefore, a model of L1 priming should probably apply to both production and comprehension. This is especially supported by our results from Chapter 4, in which we studied structural priming of ditransitives in native speakers of English.There was significant abstract structural priming in production and comprehension, and in both modalities priming was sensitive to lexical overlap between prime and target.

Importantly, in comprehension, participants were slower to read a direct object dative structure after a prepositional object dative prime than after a baseline prime. This means that exposure to a prime structure may lead to inhibition while processing the alternative structure. As far as we know, such inhibitory effects are a novel finding and have not been observed in structural priming before. Nevertheless, inhibitory effects are predicted by the two-stage competition model (Segaert et al., 2011). Our results thus provide support for a model in which there is competition between alternatives during sentence processing.

The two-stage competition model, which was developed for structural priming in production, assumes two stages, namely a selection stage and a planning stage (see Figure GD.1). During the selection stage, there is competition between the two alternating structures. As soon as the activation level of one of the structures reaches the selection threshold, that structure is selected and passed on to the planning stage. Inhibition effects are the result of this competition during the selection stage. The nodes representing syntactic structures have a base level of activation. This base level is determined by implicit learning. As a consequence, more frequent structures have a higher base level of activation than less frequent structures. When the more frequent structure is primed and thus receives extra activation, the gap between the activation levels of the high frequent structure and the low frequent structure increases. Consequently, competition resolution proceeds faster and the selection time of the high frequent structure decreases. In contrast, when the less frequent structure is primed, the gap in activation level of the high frequent structure and that of the low frequent structure decreases. As a result, there is more competition between the two structures and the time it takes to select a structure increases. In other words, inhibition effects may occur if the less frequent structure of two syntactic alternants is primed. This explains why participants were slower to read a direct object dative structure after a prepositional object dative prime in the comprehension experiment reported in Chapter 4.

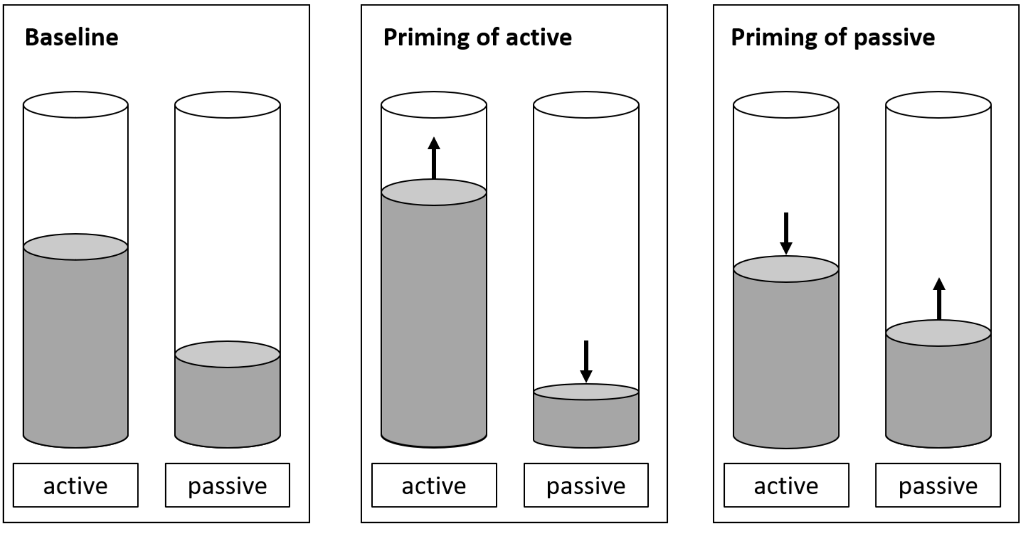


Figure GD.1: The two-stage competition model (Segaert et al., 2011)

The term “selection stage” may not be totally adequate for comprehension. Nevertheless, the process of competition between structures presumably works similarly in comprehension and in production. Competition between alternating structures is also part of constraint-based models of sentence processing as a way to solve syntactic ambiguity (e.g., MacDonald et al., 1994; cf. Humphreys & Gennari, 2014 for neurological evidence).

On a final note, the two-stage competition model is a hybrid model of priming, meaning that it attributes priming to both residual activation (cf. Pickering & Branigan, 1998) and implicit learning (cf. Chang et al., 2006). Chapter 1 provides additional evidence for such a hybrid account. In this chapter, we boosted the use of passive sentences over active sentences in an abstract structural priming paradigm by means of a lexical intervention halfway the experiment, which consisted of four items with verb overlap between prime and target. These items boosted the subsequent production of passives in the items without lexical overlap, not only in L2 speakers (which could indicate a learning effect), but also in native speakers. Normally, it is assumed that the lexical boost effect is due to residual activation (Hartsuiker & Pickering, 2008). Hence, a residual activation account does not predict a long-lasting boost effect. A hybrid model can explain lexically-boosted implicit learning. Because of the lexical boost effect caused by residual activation during the intervention, participants produce more passives than they would have done if there was no lexical overlap between prime and target. Different from what is assumed by Pickering and Branigan, the residual activation does not just disappear, but the base level activation of the passive is permanently updated by means of implicit learning. In other words, the lexical intervention leads to stronger priming, and consequently, a stronger adaptation of the base level activation of the boosted structure than if the experiment had only contained abstract structural priming.

Differences in priming between production and comprehension

The results from Chapter 4 are not exactly the same between production and comprehension. In addition, in Chapter 3, we did not find structural priming in comprehension, while we did observe priming in production. Our findings raise the question whether one would in fact expect identical results between the modalities, even if one assumes that the same abstract structural representations are involved and that the priming mechanisms leading to both abstract structural priming and lexically-boosted priming are shared.

First, according to the two-stage competition model (Segaert et al., 2011), relative frequency plays an important role in determining which of the syntactic alternants is subject to facilitation or inhibition respectively. Additionally, the magnitude of structural priming is affected by the relative frequency of structures due to the implicit learning mechanism that plays a role: one often finds an inverse preference effect, meaning that priming is stronger for less frequent structures than for more frequent structures (V. S. Ferreira & Bock, 2006). In Segaert et al., active and passive structures were tested, which have a very unbalanced distribution. Actives are much more frequent than passive structures. The gap between the base level activation of the two structure is therefore very large. Priming of the less frequent passive structure will decrease this gap, but the priming effect is too small to overrule the activation level of the active structure. In contrast, in our comprehension experiments (Chapter 3-5), we tested direct object datives (DO) and prepositional object datives (PO), for which the distribution is much more equal.

Therefore, priming of a structure may lead to a switch with regard to which structure has the highest level of activation

But also the experimental design may have an effect on the relative frequency of the structures. In Chapter 4, the production experiment revealed a preference for the PO structure. In the comprehension experiment, there was no clear preference for either structure, although the inhibition effect suggests a higher activation level of the DO structure than that of the PO structure. The picture description task in production may have boosted the production of PO datives, as the pictures display the object between the agent and the recipient, in accordance with the order in which entities are mentioned in PO datives. In the maze task paradigm used in comprehension, participants choose between an animate noun and an inanimate noun in order to determine whether the sentence will continue as a DO sentence or a PO sentence respectively. As animate nouns are conceptually more accessible than inanimate nouns (Bock & Warren, 1985), this may have promoted the choice for DO datives, in which the animate entity precedes the inanimate one (e.g., *The waitress passes the* **boxer** *the cake*) . Due to the different nature of the two modalities, differences in the experimental design and thus differences in the relative frequencies of the structural alternatives are inevitable. In the case of ditransitives, this means that it is not fixed which structure has the highest activation level (also prior to priming), which leads to differential predictions with regard to the polarity and the magnitude of the priming effect.

Second, structural priming effects are not only affected by lexical overlap or relative frequency, but may also be modified by other factors, which may be modality-specific. This became especially clear from Chapter 3, in which we compared priming from comprehension to production on the one hand to priming from comprehension to comprehension on the other hand. In other words, the primes were identical across experiments, but the targets differed in modality. Despite the identical primes, we found priming in production, but not in comprehension. This may at least partly be attributed to the experimental design we used, which seemed to be not sensitive enough to pick up on priming effects. Still, in comprehension, priming may be weaker due to message predictability (Ziegler & Snedeker, 2019). Listeners may not need to build a full representation in order to process the sentence, which may reduce priming effects as a result of residual activation. At the same time, we found evidence that priming in production was enhanced through self-priming. While in comprehension, the two target conditions (DO and PO) are evenly dispersed, in production participants often display a preference for one of the two structures. Due to their production choices, one of the two structures is processed more often than the other structure within the experiment, and this has an effect on the magnitude of priming.

In addition, according to the two-stage competition model of priming (Segaert et al., 2011), production does not only involve a sentence selection stage (which is arguably shared with comprehension), but also a planning stage. Segaert et al. argue that facilitatory priming also occurs during this stage (if the target structure is the same as the prime structure), which yields an additive effect to the priming effects from the selection stage. Due to this additional priming during the planning stage, the facilitatory priming effects for the most frequent structure from the selection stage may be stronger in production than in comprehension, while inhibitory effects for the less frequent structures may be weaker, since priming during the planning stage compensates for the inhibitory effects induced by priming of the less frequent structure during the selection stage. However, note that there may also be additional priming after the sentence selection stage during follow-up processing stages in comprehension.

With our results taken into account, future versions of the two-stage competition model (Segaert et al., 2011) should be more specific in its predictions in production and comprehension. Currently, the described priming mechanisms concern the situation in which the prime structure is the same as the target structure, such as the effect of an active prime on an active target. However, it is unclear what the predictions of the models are when the target structure is the alternative structure. When competition between the active and the passive decreases after an active prime is chosen, does that also mean that the model predicts faster reaction times for the passive structure after an active prime than after a baseline prime? This especially plays a role during comprehension, when the participant may not have a choice between structural alternatives, and is exposed to the alternative structure after a particular prime. But also during production, the priming effect is usually only between 5-10 percent point, meaning that the participant still often chooses to produce the alternative structure. Even when the target structure is different from the prime structure, the prime structure may still have an effect on the target structure, as the observed inhibition effects in Chapter 4 and 5 show. Furthermore, the model should be more explicit on the (modality-specific) follow-up processes after the selection stage, such as the planning stage in production. As priming during these follow-up processes may either accumulate or compensate for the priming effects from the selection stage, any result may currently be fit into the two-stage competition model. For instance, if one unexpectedly observes a facilitatory priming effect of the less frequent structure, one could attribute this due to strong facilitatory priming during the selection stage or another follow-up process, arguing that the effect is so strong that it does not only compensate for the predicted inhibitory effect, but also appears as a significant facilitatory priming effect.

In conclusion, our findings on L1 priming in comprehension from Chapter 3 and Chapter 4 confirm that structural priming effects follow from modality-specific processing on the one hand and a mechanism of priming of abstract structural representations which is shared between production and comprehension on the other hand. Given the observed inhibition effects, this shared mechanism is best explained by a hybrid model of structural priming which incorporates competition between structural alternatives, such as the two-stage competition model of priming of Segaert et al. (2011). Observed structural priming effects thus seem to be the result of a complex interplay between residual activation, implicit learning mechanisms, and relative frequency. The relative frequency of structures is partly context-dependent, and may for instance differ between experimental designs, even if the tested items are identical. As the relative frequency of structures may be different between a production and a comprehension experiment, the priming effects may differ between modalities, even though being the result of the same priming mechanisms. Since these priming mechanisms do not only play a role during L1 priming, but also during L2 priming (see Chapter 5), a refined model of L2 priming should also be based on the two-stage competition model of priming.

A model of L2 priming: The final stage of the developmental account

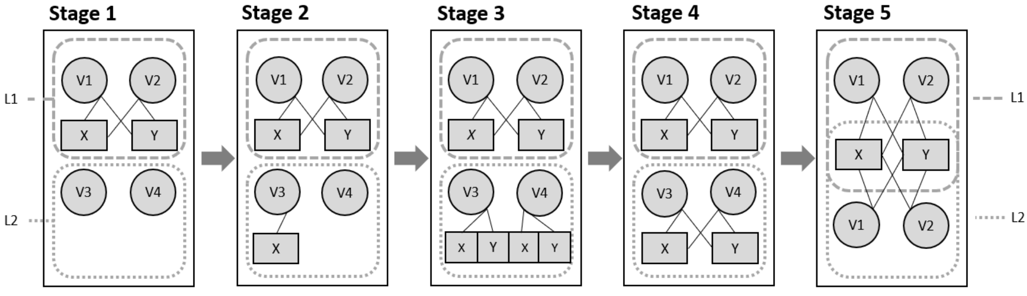


Figure GD.2: The five stages of the developmental account (Hartsuiker & Bernolet, 2017). X and Y are combinatorial nodes representing syntactic structures. V1, V2, etc. are lexical representations (for instance of verbs).

The bilingual lexical-syntactic model (Hartsuiker et al., 2004) aims to account for priming between languages, and also incorporates the idea that between-language priming may be equally strong between languages as within languages (cf. Kantola & van Gompel, 2011). The developmental account (Hartsuiker & Bernolet, 2017; see Figure GD.2) intends to explain how L2 learners who start to learn a language arrive at the point at which there is between-language priming. In other words, the model of L2 priming by Hartsuiker et al. equals the final stage of the developmental account.

There is no doubt that priming between languages occurs. However, it is not clear whether between-language priming is the result of shared combinatorial nodes (as it is represented in the developmental account), or whether similar syntactic structures may be connected rather than shared between languages (van Gompel & Arai, 2018). We addressed this issue in Chapter 2. Particularly, we investigated L2 representations in very proficient L2 speakers by comparing the production preferences and priming effects in Dutch between Arabic/Berber and Turkish heritage speakers with Dutch as their dominant L2 and native speakers of Dutch. Dutch has PP-final passives and PP-medial passives in addition to short, agentless passives, and speakers have a strong preference for the former alternative. Turkish has the same syntactic alternants available, but prefers the PP-medial passive over the PP-final passive. Arabic and Berber do not have full passives at all and exclusively use short passives. The results of our priming experiment showed that Arabic/Berber-Dutch participants were more likely to produce short passives in Dutch as well, which is an instance of facilitatory cross-linguistic influence. Turkish-Dutch participants, on the other hand, produced fewer PP-medial passives in Dutch than the Dutch participants did, which is suggestive of an inhibitory effect from Turkish on Dutch.

The developmental account states that structures become shared between languages whenever they are similar enough. In other words, similar structures in different languages share a single combinatorial node associated with that structure. However, inhibitory effects can presumably only be explained if one assumes competition between combinatorial nodes. As we found inhibition between similar syntactic alternatives, the Turkish PP-medial passive and the Dutch PP-medial passive must have separate combinatorial nodes rather than a joint combinatorial node. This is in line with van Gompel and Arai (2018), who argue that syntactic structures are connected instead of shared if structures are similar, but not identical in terms of hierarchical structure and constituent order. At least for languages with similar, non-identical syntactic structures, the final pane of the developmental account displaying the shared combinatorial nodes X and Y for the L1 and the L2 should therefore be replaced by a pane showing separate combinatorial nodes for the structures X_1 and X_2, and Y_1 and Y_2 (Figure GD.3). Different from stage 4, when there are language-specific representations, at stage 5 these combinatorial nodes are connected and thus interact between languages. In other words, X_1 and X_2 are on the same level as X_1 and Y_1. Again, we thus find support for the two-stage competition model of priming (Segaert et al., 2011), postulating that there is competition between syntactic alternants during sentence selection. The results from Chapter 2 show that competition does not only take place between alternants within a language, but also between structures in different languages, if they are represented in separate combinatorial nodes.

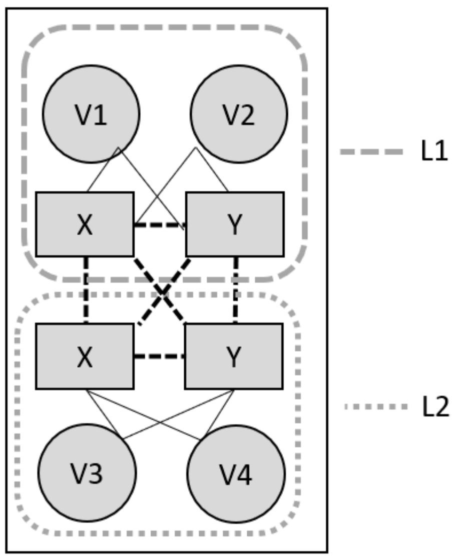


Figure GD.3: The adapted final stage of the developmental account, showing how combinatorial nodes are connected both within and between languages.

More research is needed to understand when, if ever, structures in multiple languages share a single combinatorial node. For this purpose, a better understanding of inhibition and facilitation of the production preferences of one language onto the other language is needed. The finding of facilitatory effects of the Arabic/Berber short passives on the use of short passives in Dutch should not be taken as evidence for shared syntax between Arabic/Berber and Dutch. Even in the case of connected structures, there may be facilitation under some conditions and inhibition in other cases. It is currently unclear what the conditions are that cause either facilitation or inhibition. As argued in Chapter 2, it is not the available alternatives in a language which determine the direction of the cross-linguistic influence. Instead, extralinguistic factors such as age of acquisition and experimental design may play a role.

Taken together, our findings suggest that - at least in the case of similar, non-identical structures - the final stage of L2 syntax is a stage in which structural alternatives are connected rather than shared between languages. As such, in bilingual speakers, there is not only competition between the different structures within each language as part of the two-stage competition model (Segaert et al., 2011), but also competition between the structural alternatives from different languages. Consequently, production preferences, reflecting the outcome of the competition between the alternatives, in the L1 are affected by those in the L2 and vice versa. A developmental account of L2 syntax should thus explain how L2 learners reach the point at which similar L1 and L2 structures compete with each other during sentence processing.

A model of L2 learning: Revisiting the developmental account

Before we will discuss the implications of our results for the developmental model, note that we focused on the transition between stage 3 and stage 4 of the developmental account. This was explicitly the case in our study with very beginning L2 learners of Dutch (Chapter 1), but also the L2 learners in Chapter 5 presumably were around stage 3 or 4, as we found lexically-dependent priming in the absence of abstract structural priming in some conditions. One may wonder to what extent the first stages of the developmental account are in fact testable. This is not only true for a natural setting, but also in an artificial language paradigm, as Muylle (2020) showed that participants may become proficient very fast in an artificial language. One of the problems is that structural priming in production can only be measured if participants produce both syntactic alternatives. To illustrate, if participants only use actives and no passives already in the baseline condition, there is no room for an increase in the proportion of actives after an active prime. This may not only be the case if participants have a structural representation for one of the alternatives but not for the other, but also if participants are hesitant to produce the more difficult structure. Moreover, it is assumed that L2 learners may show priming already from stage 1 if they use a copy-edit strategy. Participants may edit the prime sentence which is still in their explicit memory in order to produce the target sentence. There is no reason why participants would only be able to copy-edit the nouns and not the verb of the sentence. As such, some abstract structural priming may already occur at the first stage of language learning.

The crucial assumption of the developmental account of L2 syntax (Hartsuiker & Bernolet, 2017) is that L2 learners start with item-specific representations, which develop into abstract structural representations, and this happens between stage 3 and 4. As discussed above, these abstract structural representations eventually become connected or shared between languages, which constitutes the final stage of L2 syntax. The stages leading to the final stage of L2 syntax were mainly investigated in Chapter 1 and Chapter 5.

In Chapter 1, we investigated L2 representations in beginning learners. More specifically, we studied the development of active and passive structures in late learners of Dutch in both a cross-sectional and a longitudinal design. We found abstract structural priming for the less complex active structure before priming of the more complex passive structure occurred. Passive priming was only present in more proficient participants. In Chapter

5, we collected production data as well as comprehension data for Chinese-English and Dutch-English L2 speakers, testing ditransitives in English. The production choice data show that there is only priming in the presence of lexical overlap in the less proficient Chinese-English group, whereas there is abstract priming for the prepositional object dative in the more proficient Dutch-English group. In line with the developmental account (Hartsuiker & Bernolet, 2017), the aforementioned results of both Chapter 1 and Chapter 5 suggest that abstract structural representations are not present in the very beginning stages of language learners, and become stronger upon increasing proficiency.

However, not all our results are in accordance with the developmental account. Importantly, the developmental account predicts stronger lexically-dependent priming for less proficient participants than for more proficient participants, and stronger abstract structural priming upon increasing proficiency. Our findings include some instances where this is not the case. For example, in Chapter 1, we boosted the use of passive sentences over active sentences in an abstract structural priming paradigm by means of a lexical invention. During the intervention itself, which, in fact, consisted of prime-target pairs inducing lexically-dependent priming, priming was stronger for more proficient than for less proficient participants. Similarly, in Chapter 5, Dutch-English participants who were more proficient in English were primed stronger in their choice for either the DO or PO during comprehension than less proficient participants in the presence of lexical overlap. At the same time, the priming of the onset latencies during production was stronger for less proficient Chinese-English and Dutch-English participants than for more proficient participants, regardless of whether there was lexical overlap between prime and target.

The results suggest that proficiency may interact with the three different loci of structural priming, namely explicit memory, residual activation, and implicit learning. For low proficient L2 learners (stage 1), explicit memory is the only locus of priming (disregarding L1 transfer). Consequently, priming is the weakest and the least stable (but not absent) in beginning L2 learners, and this is true for lexically-dependent priming as well as abstract structural priming. This explains the proficiency effect in the lexical intervention in Chapter 1 and in the comprehension choice data in Chapter 5.

Once item-specific representations of structures are formed (stage 2 and 3) and are being developed into abstract structural representations (stage 4), residual activation and implicit learning contribute to the priming effect. As abstract representations develop over time, residual activation of the combinatorial nodes representing syntactic structures and hence, abstract structural priming increase as proficiency increases. These mechanisms may have contradictory effects. According to the developmental account, residual activation of lexical items and thus lexically-dependent priming decrease upon increasing proficiency, while residual activation of abstract combinatorial nodes representing syntactic structure and hence, abstract structural priming increase as proficiency increases. This leads to stronger lexically-dependent priming and weaker abstract structural priming in medium proficient L2 learners (stage 2 and 3) than in high proficient L2 learners (stage 4 and 5). This mechanism accounts for the proficiency effects reported in Bernolet et al. (2013), which we replicated in the production choice data from Chapter 5.

On the other hand, according to the implicit learning model, prediction error is stronger in less proficient participants than in more proficient participants, as the relative weights of structures are established less well in less proficient participants. This mechanism results in stronger priming in less proficient participants than in more proficient participants, independently of lexical overlap. This effect may be visible in the production onset latencies in Chapter 5, and may numerically also be present in the cross-sectional study of Chapter 1. Especially in low proficient participants such as those in Chapter 1, this implicit learning effect may be hard to detect, because low proficient participants may be hesitant to produce the structure they are less familiar with, in spite of stronger priming of the less frequent structure than of the more frequent structure.

All in all, we do not think that our results should be interpreted as evidence against the basic architecture of the developmental account, as there is evidence that participants start with item-specific representations before developing abstract structural representations even if their L1 has a parallel syntactic structure. Instead, if the developmental account aims to explain priming effects at different stages of L2 development, it should integrate the mechanisms that affect (the magnitude of) priming also in L1 priming, such as competition between structures and implicit learning. Our results show that proficiency interacts in different ways with the multiple mechanisms. More research is needed to understand the outcome of the interaction of the mechanisms at different stages of L2 development.

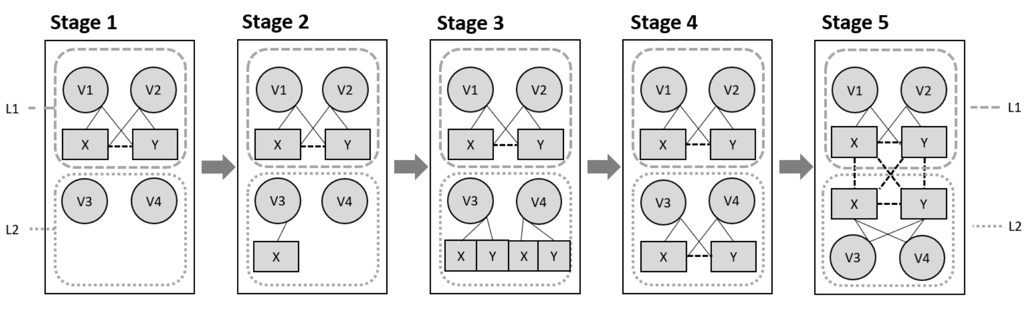


Figure GD.4: The five stages of the developmental account, displaying competing connected combinatorial nodes.

Future directions

This thesis is part of a series of three theses which all aim to test the predictions that follow from the developmental account. Muylle (2020) investigated the development of L2 representations in an artificial language paradigm. Sijyeniyo (2023) studied which factors affect the development of syntactic structures in a natural setting, such as explicit language instruction and guided implicit learning (i.e., using priming as a learning method rather than an experimental paradigm). This thesis as well as their theses show that residual activation of item-specific or abstract representations is not the only mechanism that causes structural priming. Additional mechanisms, including implicit learning, explicit memory, competition between structures, and modality, determine the magnitude of the observed priming effects as well, as has been shown separately for each mechanism in previous studies. All the aspects that play a role during L1 priming are present during L2 priming as well and, on top of that, they seem to interact with L2 proficiency. In addition, L1 interference is present at each stage of L2 development, ranging from a translation strategy at the first stage to competition between L1 and L2 combinatorial nodes at the final stage. The developmental model in its current form may therefore be too simplistic in its prediction of priming effects, and the next step must be to integrate the relevant factors in an updated version.

This is a very challenging task, as it will be impossible to predict a particular outcome of all the factors at play in a behavioral experiment. In addition, the power to detect structural priming effects decreases drastically if one adds interactions. For instance, while an abstract structural priming experiment with 48 subjects and 48 items has an estimated power of 0.87, the power to detect an interaction with the priming effect as large as the lag effect (i.e., intervening items between prime and target, which provides support for cumulative priming) with the same sample size is only 0.09 (Mahowald et al., 2016). A fruitful approach may be to conduct simulation experiments in order to arrive at a computational model for the development of L2 syntax.

An important first step is the bilingual version of the implicit learning account (Khoe et al., 2021). Note that Khoe et al. obtained better results for an error-driven implicit learning account rather than for the hybrid account of Reitter et al. (2011), but this is due to the assumption of Reitter et al. that identical word order is required in order to have between-language priming, an assumption that we do not share. The bilingual model of Khoe et al. in its current form assumes bilinguals who are equally proficient in the L1 and the L2. A future step may be to train the model for bilinguals at different levels of proficiency (also see Frank [2020] on computational modelling of multilingual sentence processing).

Of course, we still see a place for behavioral experiments as well. We would like to mention two concrete suggestions. First, we already commented on the testability of the early stages of the developmental model. One approach to remedy this problem may be to test structural priming in comprehension rather than in production in beginning L2 learners. In comprehension, priming effects from a copy-edit strategy will be ruled out. In addition, if learners do not produce a difficult structure because they are insecure, priming in comprehension may still reveal beginning structural representations. Finally, in comprehension, it is possible to measure structural priming with a single grammatical structure. Priming in comprehension does not require syntactic alternants, since the reaction times of the target structure after processing a prime sentence are compared to the reaction times of the target structure after an unrelated baseline prime. Thus, it is not necessary that participants are able to process both syntactic alternatives in order to investigate the representation of one of the alternatives.

The second suggestion concerns the final stage of L2 syntax. We only tested within-language priming in order to investigate production preferences, but from the assumption that structures are connected rather than shared, the prediction follows that between-language priming will be smaller than within-language priming (van Gompel & Arai, 2018). This is different from a shared syntax account, which assumes equally strong between-language and within-language priming. Note that this hypothesis is hard to test statistically, since it assumes a null effect, while statistical hypothesis testing normally aims to reject the null hypothesis (although equivalent testing or Bayesian modelling may provide a solution to this, cf. Harms & Lakens, 2018). A potential approach would be a comparison between two between-language priming experiments with different language pairs. In the first experiment, one might test structural priming between languages for which the alternatives in two languages are identical with regard to hierarchical structure and constituent order, such as the PP-final passive in English and the PP-final passive in French. In the second experiment, the tested alternatives may be similar but not identical between languages, which is the case for the PP-final passive in English and the PP-medial passive in German. One could then compare the magnitude of between-language priming between the different language pairs. In the case of shared syntax when structures are identical, but connected syntax for similar, non-identical structures, between-language priming may be stronger between the languages with identical structures (i.e., between English and French passives) than between the languages with similar, non-identical structures (i.e., between English and German passives).

In conclusion, this thesis provides support for the developmental account, assuming that L2 learners start with item-specific representations, which develop into abstract structural representations. Different from the developmental account, our findings suggest that these structural representations become connected rather than shared between languages. Structural priming effects may be different at each stage of L2 development. The magnitude of structural priming is the outcome of an interplay between different mechanisms, which include residual activation, implicit learning, explicit memory, competition between structures and modality-specific processing, and these mechanisms interact with L2 proficiency. The relevant mechanisms should therefore be integrated into a hybrid developmental account of syntactic representations in the L2 in order to better understand syntactic processing at different stages of L2 development.

Summary

When learning a language, one needs to acquire the grammatical structures of the new language. These syntactic structures are stored in mental representations. In this thesis, we investigated how mental representations of syntactic structures are formed during the process of learning a language, and how the syntactic representations of one language affect the syntactic representations of another language. Mental representations of syntactic structures can be investigated by means of structural priming (Bock, 1986). Structural priming is “the phenomenon by which processing one utterance facilitates processing of another utterance on the basis of a repeated syntactic structure” (Branigan, 2007, p. 1) and takes place both during sentence production and sentence comprehension. In a structural priming experiment in production (see Figure 1), participants who are primed by a passive sentence (*The elephant is treated by the veterinarian*), are more likely to describe a picture using a passive sentence than if the prime sentence had been an active sentence (*The veterinarian treats the elephant*). In comprehension, participants may be faster to process a passive sentence after a passive prime than in unprimed condition.

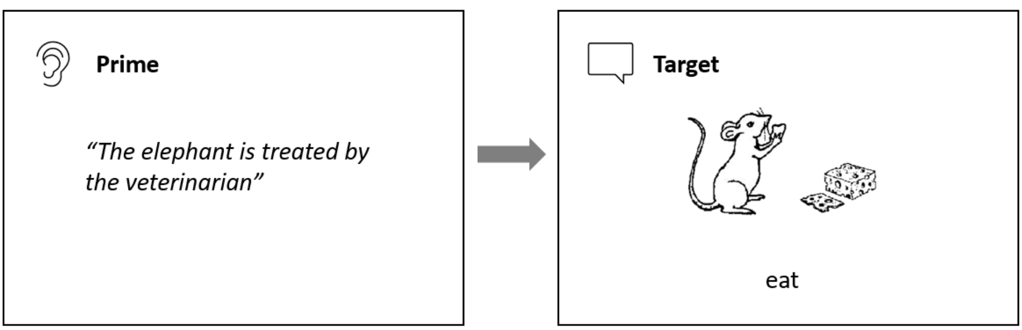


Figure 1: The structural priming paradigm.

Structural priming effects seem to be the result of two mechanisms, namely residual activation (Pickering & Branigan, 1998) and implicit learning (Chang et al., 2006). Structural priming leads to a temporarily higher level of activation of the memory node containing the abstract mental representation of the primed syntactic structure, which makes it easier to reactivate that syntactic structure. Processing a primed structure also strengthens the connections of its syntactic representation to conceptual and/or lexical memory nodes during error-based, implicit learning, which leads to permanent changes in the neural network. Hybrid models integrate both residual activation and implicit learning (see Reitter et al., 2011; Segaert et al., 2011) to account for structural priming effects.

Structural priming effects do not only occur within ones native language, but also in the second language (L2) and even between languages. For instance, Hartsuiker et al. (2004)

showed that participants produced more passive structures in English after a passive prime in Spanish than after a Spanish active prime. This finding suggests that mental representations may be shared between languages for similar structures. Cross-linguistic priming seems to be modulated by proficiency. Abstract structural priming (that is, the prime sentence [*The elephant is treated by the veterinarian*] and the target sentence [*The cheese is eaten by the mouse*] do not contain overlap in lexical items) becomes stronger upon increasing proficiency. By contrast, structural priming within the L2 is stronger for less proficient participants than for more proficient participants if there is lexical overlap between prime and target (Bernolet et al., 2013). This suggests that L2 learners might start with item-specific and language-specific syntactic representations. These representations become abstract and shared between the L1 and the L2 over time. The development from item-specific to abstract structural representations has been captured in the developmental account of L2 syntax (Bernolet & Hartsuiker, 2018; Hartsuiker & Bernolet, 2017), which aims to explain how syntactic representations develop during L2 acquisition in five stages (see Figure 2). At stage 1, learners have lexical representations of words. At stage 2, learners attach item-specific syntactic information to these lexical representations. During stage 3, the item-specific syntactic representations are expanded and lexical items may be used with more than one construction. At stage 4, syntactic constructions are generalized across lexical items and become abstract. At stage 5, structures become shared across languages.

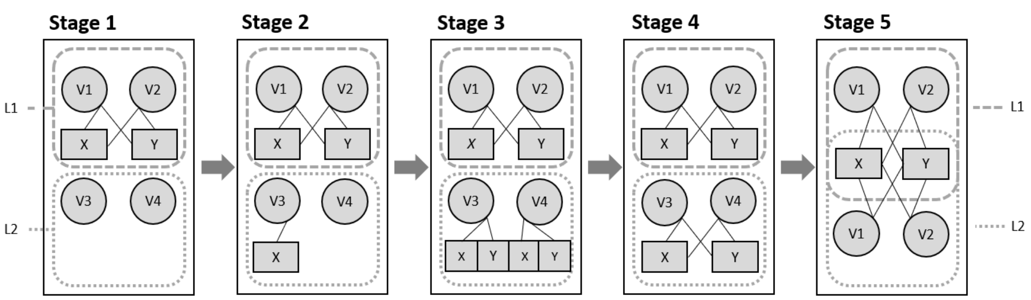


Figure 2: The five stages of the developmental account (Hartsuiker & Bernolet, 2017). X and Y are memory nodes representing syntactic structures. V1, V2, etc. are lexical representations of words (for instance of verbs).

The aim of this thesis was to elaborate upon the developmental account of L2 syntax (Hartsuiker & Bernolet, 2017) in two different ways. First, the developmental model was largely based on late L2 learners at an intermediate to advanced level. In the Production part, we therefore investigated the L2 syntax of the more extreme sides of the continuum of the language learners, namely beginning learners and native-like L2 speakers. Second, the developmental model is based on structural priming of L2 structures in production, and we studied to what extent the learning trajectory also applies to comprehension in the Comprehension part.

The Production part consists of two chapters. In Chapter 2, we tested beginning L2 learners of Dutch in a longitudinal and a cross-sectional study. We studied the transition from item-specific, lexically-dependent representations to abstract syntactic representations for Dutch active and passive sentences. The results showed that abstract priming of the more frequent and easier active structure takes place before priming of the more complex and less frequent passive structure, although abstract representations of the passive are formed quite rapidly after exposure. In Chapter 3, we investigated the production preferences and syntactic representations for Dutch passive sentences in highly proficient speakers of Dutch, who are proficient heritage speakers of another language. We found that in comparison to a Dutch control group, Arabic/Berber-Dutch speakers were more likely to produce agentless passives (*het broodje wordt gegeten* "de sandwich is being eaten") in Dutch, which is the common passive structure in Arabic and Berber. By contrast, Turkish-Dutch speakers produced more by-phrase-final passives (*het broodje wordt gegeten door de jongen* "the sandwich is being eaten by the boy") and fewer by-phrase-medial passives (*het broodje wordt door de jongen gegeten*) in Dutch than the control group did, although the by-phrase-medial passive is the most frequent passive structure in Turkish. This inhibition effect suggests that at the final stage of L2 syntax learners may have connected syntactic representations rather than shared representations if structures are similar but not identical between languages (see Figure 3).

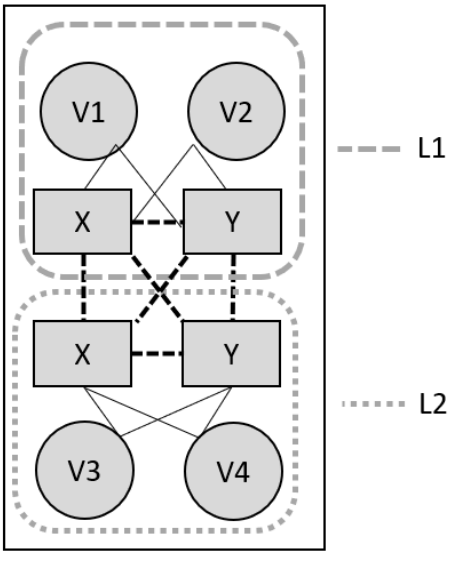


Figure 3: The adapted final stage of the developmental account, showing how combinatorial nodes are connected both within and between languages.

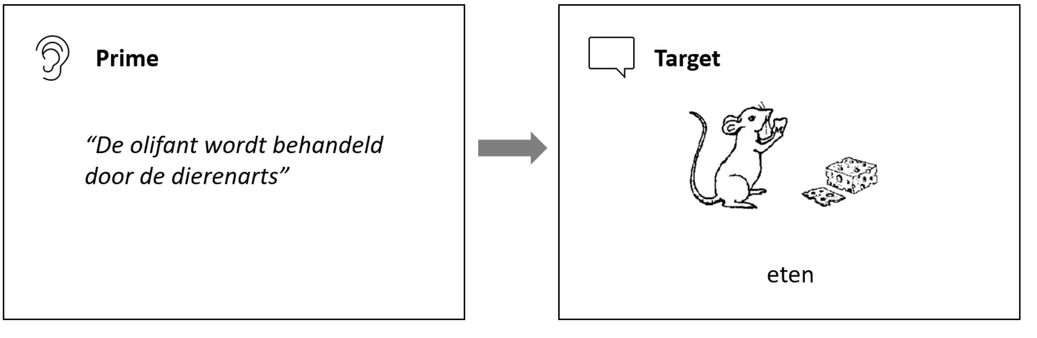
The Comprehension part contains three chapters. In Chapter 4, we studied to what extent modality-specific aspects affect the magnitude of structural priming by comparing priming in production to priming in comprehension. In our results, we could only detect priming effects in production and not in comprehension. In addition, we found evidence that self-priming played a role in the production experiment. Together, these results suggest that modality-specific aspects may lead to larger priming in production than in comprehension, despite a common mechanism causing priming. In Chapter 5, we compared structural priming of ditransitives between production and comprehension in L1 speakers of English. The results showed that structural priming does not only involve facilitation of repeated processing of structures, but priming with one structure may also lead to inhibition of the alternative structure. Double object datives (*The waitress passes the boxer the cake*) were read slower after a prepositional object dative prime sentence

(*The waitress passes the cake to the boxer*). This suggests that there is competition between alternatives during sentence selection, in line with the hybrid model of priming of Segaert et al. (2011). Finally, in Chapter 6, we examined structural priming of ditransitives in production and in comprehension in two different populations of intermediate to advanced L2 speakers of English, namely Chinese-English and Dutch-English bilinguals. We inquired whether the proficiency effects as predicted by the developmental account (decreasing lexically-dependent priming and increasing abstract structural priming upon increasing proficiency) could also be found in comprehension and in the onset latencies of production. We found that proficiency modulated priming in the onset latencies in production and in comprehension as well, but not always in the expected direction. Our results support a hybrid model of structural priming (Segaert et al.), assuming three different loci of structural priming, namely explicit memory, residual activation and implicit learning. Proficiency may interact with these three different loci of structural priming, meaning that the interplay between these different mechanisms may have different outcomes at different stages of language learning.

In short, this thesis provides support for the developmental account, assuming that L2 learners start with item-specific representations, which develop into abstract structural representations. Different from the developmental account, our results suggest that these structural representations become at least sometimes connected rather than shared between languages. Structural priming effects may be different at each stage of L2 development. The magnitude of structural priming is the outcome of an interplay between different mechanisms, which include residual activation, implicit learning, explicit memory, competition between structures and modality-specific processing, and these mechanisms interact with L2 proficiency. The relevant mechanisms should therefore be integrated into a developmental account of syntactic representations in the L2 in order to account for the priming effects observed at different stages of L2 development.

Samenvatting

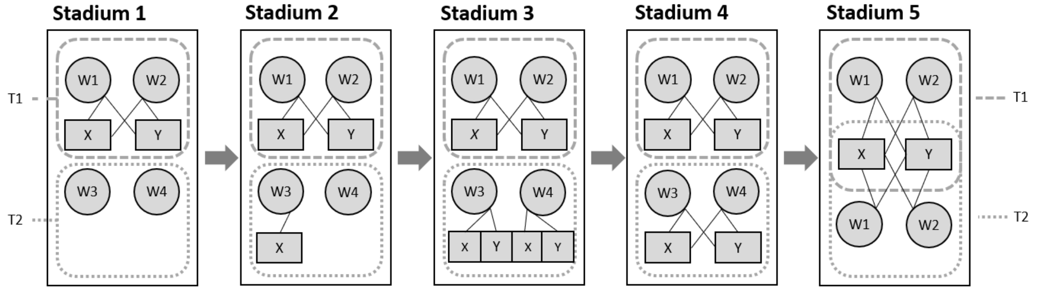
Bij het leren van een taal moeten leerders de grammaticale structuren van de nieuwe taal verwerven. Deze syntactische structuren worden opgeslagen in mentale representaties. In dit proefschrift onderzochten we hoe de mentale representaties van syntactische structuren gevormd worden tijdens het leerproces, en hoe de syntactische representaties van de ene taal de syntactische representaties van een andere taal kunnen beïnvloeden. De mentale representaties van syntactische structuren kunnen onderzocht worden door middel van structurele priming (Bock, 1986). Structurele priming is het fenomeen waarbij het verwerken van een zin met een bepaalde structuur de verwerking van een volgende zin met dezelfde structuur vergemakkelijkt (Branigan, 2007). Dit fenomeen vindt zowel tijdens zinsproductie als tijdens zinsbegrip plaats. In een priming-experiment in productie (zie Figuur 1) zijn participanten die geprimed worden met een passieve zin (*De olifant wordt behandeld door de dierenarts*), eerder geneigd om een afbeelding met een passieve zin te omschrijven dan wanneer de primezin een actieve zin (*De dierenarts behandelt de olifant*) was geweest. In begrip verwerken participanten een passieve zin sneller na een passieve prime dan in ongeprimede conditie.



Figuur 1: Een trial in een experiment met structurele priming.

Structureleprimingeffecten lijken het gevolg te zijn van twee mechanismes, namelijk residuele activatie (Pickering & Branigan, 1998) en impliciet leren (Chang et al., 2006). Structurele priming leidt tijdelijk tot een hoger activatieniveau van de geheugenknoop die de abstracte mentale representatie van de geprimede syntactische structuur bevat. Dit maakt het makkelijker om de betreffende syntactische structuur opnieuw te activeren. De verwerking van de primestructuur versterkt ook de verbindingen van de betreffende syntactische structuur met conceptuele en/of lexicale geheugenknopen tijdens impliciet, door fouten gestuurd leren. Dit leidt tot permanente veranderingen in het neuraal netwerk. Hybride modellen (zie Reitter et al., 2011; Segaert et al., 2011) integreren residuele activatie en impliciet leren om structureleprimingeffecten te verklaren.

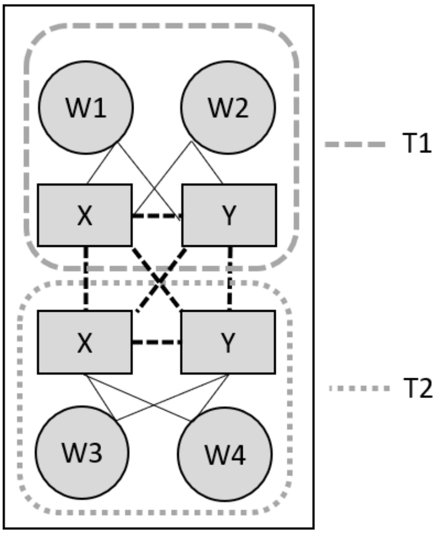
Structurele priming vindt niet alleen plaats in de moedertaal, maar ook in de tweede taal (T2) en zelfs tussen talen. Hartsuiker et al. (2004) toonden bijvoorbeeld aan dat deelnemers meer passieve zinnen produceerden in het Engels na een passieve primezin in het Spaans dan na een Spaanse actieve primezin. Dit suggereert dat mentale representaties van vergelijkbare structuren gedeeld kunnen worden tussen talen. Cross-linguïstische priming wordt gemoduleerd door taalvaardigheid. Abstracte structurele priming (wanneer de primezin [*De olifant wordt behandeld door de dierenarts*] en de targetzin [*De kaas wordt gegeten door de muis*] geen overlappende lexicale items bevatten) wordt sterker naarmate de taalvaardigheid toeneemt. Structurele priming binnen de T2 is echter sterker in minder taalvaardige participanten dan in meer taalvaardige participanten wanneer er lexicale overlap is tussen de prime en de target (Bernolet et al., 2013). Dit suggereert dat taalleerders in eerste instantie itemspecifieke en taalspecifieke syntactische representaties vormen, die later abstract worden en gedeeld tussen talen. De ontwikkeling van itemspecifieke naar abstracte structurele representaties is uitgewerkt in het ontwikkelingsmodel voor T2-syntaxis (Bernolet & Hartsuiker, 2018; Hartsuiker & Bernolet, 2017). Dit model omvat vijf stadia en beschrijft hoe syntactische representaties zich ontwikkelen tijdens tweedetaalverwerving (zie Figuur 2). In het eerste stadium hebben leerders lexicale representaties van woorden. In het tweede stadium voegen leerders itemspecifieke syntactische representaties toe aan deze lexicale representaties. Tijdens het derde stadium worden de itemspecifieke syntactische representaties uitgebreid en kunnen lexicale items met meer dan één constructie worden gebruikt. In het vierde stadium worden syntactische constructies gegeneraliseerd over lexicale items heen en ontstaan abstracte syntactische representaties. In het vijfde stadium gaan leerders de abstracte representaties delen tussen talen.



Figuur 2: De vijf stadia van het ontwikkelingsmodel (Hartsuiker & Bernolet, 2017). X en Y zijn geheugenknopen met abstracte representaties van syntactische structuren. W1, W2, etc. zijn lexicale representaties van woorden (bijvoorbeeld van werkwoorden).

Het doel van dit proefschrift was om het ontwikkelingsmodel voor T2-syntaxis (Hartsuiker & Bernolet, 2017) verder uit te werken op twee verschillende manieren. Ten eerste was het ontwikkelingsmodel grotendeels gebaseerd op late leerders met een redelijke tot gevorderde taalvaardigheid. In het Productiedeel onderzochten we daarom de T2-syntaxis aan de twee uiterste zijden van het continuum van taalleerders, namelijk beginnende leerders en vergevorderde leerders die de T2 beheersen op moedertaalniveau. Ten tweede is het ontwikkelingsmodel gebaseerd op structurele priming van T2-structuren in productie, en onderzochten wij in hoeverre het leerproces ook van toepassing is op taalbegrip in het Begripsdeel.

Het Productiedeel bestaat uit twee hoofdstukken. In hoofdstuk 2 testten wij beginnende leerders van het Nederlands in een longitudinale en een cross-sectionele studie. We onderzochten de overgang van itemspecifieke representaties naar abstracte syntactische representaties voor actieve en passieve zinnen in het Nederlands. De resultaten lieten zien dat abstracte priming eerder optreedt voor de frequentere en makkelijkere actieve structuur dan voor de minder frequente en complexere passieve structuur, hoewel abstracte representaties voor de passief eveneens vrij snel na blootstelling aan de structuur ontstaan. In hoofdstuk 3 onderzochten we de productievoorkeuren en syntactische representaties voor Nederlandse passieve zinnen in zeer taalvaardige sprekers van het Nederlands, die daarnaast zeer taalvaardig zijn in een erfgoedtaal. In vergelijking tot een Nederlandstalige controlegroep produceerden tweetalige sprekers van Arabisch/Berbers-Nederlands meer verkorte passieven (*het broodje wordt gegeten*), wat de gebruikelijke passieve structuur is in het Arabisch en het Berbers. Turks-Nederlandse sprekers produceerden echter meer passieven met een finale bijwoordelijke bepaling (*het broodje wordt gegeten door de jongen*) en minder passieven met een mediale bijwoordelijke bepaling (*het broodje wordt door de jongen gegeten*) dan de Nederlandstalige controlegroep, terwijl de passieven met mediale bijwoordelijke bepaling de meest frequente passieve structuur is in het Turks. Dit inhibitie-effect suggereert dat leerders in het laatste stadium van T2-syntax verbonden in plaats van gedeelde syntactische representaties hebben voor structuren die gelijkaardig maar niet identiek zijn tussen talen (zie Figuur 3).



Figuur 3: Het aangepaste laatste stadium van het ontwikkelingsmodel, waarbij combinatorische knopen verbonden zijn binnen en tussen talen.

Het begripsdeel bestaat uit drie hoofdstukken. In hoofdstuk 4 onderzochten we in hoeverre aspecten die specifiek zijn voor een bepaalde modaliteit de grootte van structurele priming beïnvloeden door priming in productie en priming in begrip met elkaar te vergelijken. In onze resultaten vonden we alleen primingeffecten in productie en niet in begrip. Daarnaast vonden we aanwijzingen dat zelf-priming een rol speelt in productie. Samen laten deze resultaten zien dat modaliteitsspecifieke effecten tot grotere primingeffecten in productie dan in begrip kunnen leiden, ondanks een gemeenschappelijk mechanisme dat verantwoordelijk is voor structurele priming. In hoofdstuk 5 vergeleken we structurele priming van ditransitieven tussen productie en begrip in moedertaalsprekers van het Engels. De resultaten lieten zien dat structurele priming niet alleen de verwerking van de geprimede structuur faciliteert, maar dat priming met een bepaalde structuur ook tot inhibitie van de alternatieve structuur kan leiden. Dubbelobject-datieven (*The waitress passes the boxer the cake* "The serveerster geeft de bokser de taart aan") werden langzamer gelezen na een primezin met een prepositionele datief (*The waitress passes the cake to the boxer* "The serveerster geeft de taart aan de bokser aan"). Dit suggereert dat er competitie is tussen alternatieven tijdens de zinsselectie, in lijn met het hybride model van Segaert et al. (2011). In hoofdstuk 6 onderzochten we tot slot de structurele priming van ditransitieven in productie en begrip in twee verschillende populaties taalleerders met een redelijk tot gevorderd taalniveau, namelijk Chinees-Engelse en Nederlands-Engelse tweetaligen. We onderzochten of de taalvaardigheidseffecten zoals voorspeld door het ontwikkelingsmodel (toenemende abstracte structurele priming en afnemende priming met lexicale overlap naarmate de taalvaardigheid toeneemt) ook gevonden kunnen worden in begrip en in de starttijden in productie. De resultaten toonden aan dat taalvaardigheid structurele priming in begrip en in de starttijden in productie ook beïnvloedde, maar niet altijd in de voorspelde richting. Onze resultaten sluiten aan bij het hybride model van structurele priming (Segaert et al., 2011), waarbij structurele priming uit drie loci voortkomt: expliciet geheugen, residuele activatie en impliciet leren. De taalvaardigheid interageert met deze drie loci. Hierdoor heeft de wisselwerking tussen de verschillende mechanismes een verschillende uitkomst in verschillende stadia van het leerproces.

Samengevat ondersteunen de bevindingen in dit proefschrift het ontwikkelingsmodel, waarbij tweedetaalleerders eerst itemspecifieke representaties vormen, die zich later ontwikkelen in abstracte structurele representaties. Anders dan het ontwikkelingsmodel stelt, laten onze resultaten zien dat deze structurele representaties in elk geval soms verbonden zijn tussen talen in plaats van gedeeld. Structureleprimingeffecten kunnen in elk stadium van het taalverwervingsproces anders zijn. De grootte van structurele priming is de uitkomst van een wisselwerking tussen verschillende mechanismes, waaronder residuele activatie, impliciet leren, expliciet geheugen, competitie tussen structuren en modaliteitsspecifieke verwerkingsprocessen, en er is sprake van een interactie tussen deze mechanismes en taalvaardigheid. De betreffende mechanismes moeten daarom geïntegreerd worden in het ontwikkelingsmodel voor T2-syntax om de geobserveerde primingeffecten in verschillende stadia van het leerproces te kunnen verklaren.

Bibliography

Anderssen, M., Lundquist, B., & Westergaard, M. (2018). Cross-linguistic similarities and differences in bilingual acquisition and attrition: Possessives and double definiteness in Norwegian heritage language. *Bilingualism: Language and Cognition*, *21*(4), 748–764. doi: 10.1017/s1366728918000330 62, 76

Anderssen, M., & Westergaard, M. (2020). Word order variation in heritage languages: Subject shift and object shift in Norwegian. In B. Brehmer & J. Treffers-Daller (Eds.), *Lost in transmission: The role of attrition and input in heritage language development* (pp. 100–124). John Benjamins Publishing Company. doi: 10.1075/sibil.59.04and 62, 63, 66, 75, 76

Arai, M., & Mazuka, R. (2014). The development of Japanese passive syntax as indexed by structural priming in comprehension. *Quarterly Journal of Experimental Psychology*, *67*(1), 60–78. doi: 10.1080/17470218.2013.790454 85, 103

Arai, M., van Gompel, R. P., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology*, *54*(3), 218–250. doi: 10.1016/j.cogpsych.2006.07.001 16, 22, 83, 84, 85, 86, 102, 103, 122, 124

Arman Ergin, S. (2019). *The effects of syntactic priming on Turkish-English bilinguals' production of passive sentences* (Unpublished master's thesis). Sosyal Bilimler Enstitüsü. 65

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*(1), 67–82. doi: 10.1016/j.jml.2009.09.005 86

Badawi, E., Carter, M. G., Gully, A., & Awad, M. (2004). *Modern written Arabic: A comprehensive grammar*. Routledge. doi: 10.4324/9781315856155 64

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. doi: 10.1016/j.jml.2012.11.001 42, 93, 94, 96, 97, 110, 111, 114, 116, 130, 132, 134, 135, 141, 142, 146, 147

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01 42, 93, 94, 96, 97, 98, 110, 111, 114, 116, 130, 132, 133, 135, 141, 142, 146, 147

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., . . . Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–380. doi: 10.3758/bf03196494 68

Benmamoun, E., Montrul, S., & Polinsky, M. (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, *39*(3-4), 129–181. doi: 10.1515/tl-2013-0009 66

Bernolet, S., Collina, S., & Hartsuiker, R. J. (2016). The persistence of syntactic priming revisited. *Journal of Memory and Language*, *91*, 99–116. doi: 10.1016/j.jml.2016.01.002 20, 36, 38, 84

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, *114*(3), 455–461. doi: 10.1016/j.cognition.2009.11.005 59, 60, 103, 127

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* (pp. 205–224). John Benjamins Publishing Company. doi: 10.1075/sibil.54.10ber 26, 36, 57, 60, 61, 63, 76, 122, 151, 168, 172

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2009). Persistence of emphasis in language production: A cross-linguistic approach. *Cognition*, *112*(2), 300–317. doi: 10.1016/j.cognition.2009.05.013 15, 40, 58, 64, 68, 77, 106, 117

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. doi: 10.1016/j.cognition.2013.02.005 25, 35, 37, 61, 65, 121, 122, 123, 124, 127, 137, 151, 153, 163, 168, 172

Bialystok, E., & Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Bilingualism: Language and Cognition*, *15*(2), 397–401. doi: 10.1017/s136672891100040x 67

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. doi: 10.1037/0278-7393.12.4.575 14, 15, 34, 58, 68, 83, 102, 122, 167, 171

Bock, J. K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*(3), 437–458. doi: 10.1016/j.cognition.2006.07.003 84, 87

Bock, J. K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*(2), 177–192. doi: 10.1037/0096-3445.129.2.177 20, 59, 83, 84

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, *21*(1), 47–67. doi: 10.1016/0010-0277(85)90023-x 99, 118, 119, 158

Boyce, V., Futrell, R., & Levy, R. P. (2020). Maze made easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language*, *111*, 104082. doi: 10.1016/j.jml.2019.104082 105, 106, 108, 109, 125, 138, 139, 140

Branigan, H. P. (2007). Syntactic priming. *Language and Linguistics Compass*, *1*(1-2), 1–16. doi: 10.1111/j.1749-818x.2006.00001.x 13, 83, 101, 122, 167, 171

Branigan, H. P., Pickering, M. J., & McLean, J. F. (2005). Priming prepositional-phrase attachment during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 468–481. doi: 10.1037/0278-7393.31.3.468 83, 84, 87, 103, 124

Brehmer, B., & Sopata, A. (2021). Word order in complex verb phrases in heritage Polish spoken in Germany. *Languages*, *6*(2), 70. doi: 10.3390/languages6020070 76, 77

Cai, Z. G., Pickering, M. J., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. *Journal of Memory and Language*, *66*(4), 833–849. doi: 10.1016/j.jml.2012.03.009 127

Cai, Z. G., Pickering, M. J., Yan, H., & Branigan, H. P. (2011). Lexical and syntactic representations in closely related languages: Evidence from Cantonese-Mandarin bilinguals. *Journal of Memory and Language*, *65*(4), 431–445. doi: 10.1016/j.jml.2011.05.003 25, 61

Chang, F., Dell, G. S., & Bock, J. K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. doi: 10.1037/0033-295x.113.2.234 19, 20, 34, 43, 55, 59, 84, 126, 157, 167, 171

Chang, F., Dell, G. S., Bock, J. K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*(2), 217–230. doi: 10.1023/a:1005101313330 19

Chen, Q., Xu, X., Tan, D., Zhang, J., & Zhong, Y. (2013). Syntactic priming in Chinese sentence comprehension: Evidence from event-related potentials. *Brain and Cognition*, *83*(1), 142–152. doi: 10.1016/j.bandc.2013.07.005 83, 84, 103, 124

Chen, X., Wang, S., & Hartsuiker, R. J. (2022). Error-based structure prediction in language comprehension: Evidence from verb bias effects in a visual-world structural priming paradigm for Mandarin Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(1), 60–71. doi: 10.1037/xlm0001048 17, 86, 103, 152

Corley, M., & Scheepers, C. (2002). Syntactic priming in English sentence production: Categorical and latency evidence from an internet-based study. *Psychonomic Bulletin & Review*, *9*(1), 126–131. doi: 10.3758/bf03196267 104, 107, 118

Coyle, J. M., & Kaschak, M. P. (2008). Patterns of experience with verbs affect long-term cumulative structural priming. *Psychonomic Bulletin & Review*, *15*(5), 967–970. doi: 10.3758/pbr.15.5.967 60

De Preter, J. (2011). Hier spreekt men illegaals. *De Morgen*. Retrieved from `https://www.demorgen.be/nieuws/hier-spreekt-men-illegaals~b171e2b3/` 57

De Rosario-Martinez, H. (2015). phia: Post-hoc interaction analysis [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=phia/` (R package version 0.2-1) 93, 94, 96, 97

Dell, G. S., & Chang, F. (2014). The p-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120394. doi: 10.1098/rstb.2012.0394 59

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816. doi: 10.1080/17470218.2017.1310261 90

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15. doi: 10.1111/1467-8721.00158 22, 86

Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, *1*(1-2), 71–83. doi: 10.1111/j.1749-818x.2007.00007.x 86

Ferreira, V. S., & Bock, J. K. (2006). The functions of structural priming. *Language and Cognitive Processes*, *21*(7-8), 1011-1029. doi: 10.1080/01690960600824609 15, 35, 60, 67, 86, 102, 126, 157

Fine, A. B., & Jaeger, T. F. (2016). The role of verb repetition in cumulative structural priming in comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(9), 1362–1376. doi: 10.1037/xlm0000236 16, 85, 103, 124

Flett, S., Branigan, H. P., & Pickering, M. J. (2013). Are non-native structural preferences affected by native language preferences? *Bilingualism: Language and Cognition*, *16*(4), 751–760. doi: 10.1017/s1366728912000594 61

Forster, K. I., Guerrera, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods*, *41*(1), 163–171. doi: 10.3758/brm.41.1.163 105, 125

Frank, S. L. (2020). Toward computational models of multilingual sentence processing. *Language Learning*, *71*(S1), 193–218. doi: 10.1111/lang.12406 165

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies* (Unpublished doctoral dissertation). University of Connecticut. 23, 86, 119

Giavazzi, M., Sambin, S., de Diego-Balaguer, R., Stanc, L. L., Bachoud-Lévi, A.-C., & Jacquemot, C. (2018). Structural priming in sentence comprehension: A single prime is enough. *PLOS ONE*, *13*(4), e0194959. doi: 10.1371/journal.pone.0194959 85, 103, 104, 105, 107

Giglio, L., Ostarek, M., Weber, K., & Hagoort, P. (2022). Commonalities and asymmetries in the neurobiological infrastructure for language production and comprehension. *Cerebral Cortex*, *32*(7), 1405–1418. doi: 10.1093/cercor/bhab287 82, 89, 123, 155

Göksel, A., & Kerslake, C. (2005). *Turkish: A comprehensive grammar*. Routledge. doi: 10.4324/9780203340769 64

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. doi: 10.1016/j.tics.2005.11.006 82

Gutova, E. (2013). Voice, lability, and causatives in Berber (I). *Études et Documents Berbères*, *32*(1), 105–121. doi: 10.3917/edb.032.0105 64, 65

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, *33*(2). doi: 10.18637/jss.v033.i02 72

Hadfield, J. D. (2019). MCMCglmm course notes. Retrieved from `https://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf` 72

Harms, C., & Lakens, D. (2018). Making 'null effects' informative: statistical techniques and inferential frameworks. *Journal of Clinical and Translational Research*, *3*(Suppl 2), 382–393. 165

Hartsuiker, R. J., Beerts, S., Loncke, M., Desmet, T., & Bernolet, S. (2016). Cross-linguistic structural priming in multilinguals: Further evidence for shared syntax. *Journal of Memory and Language*, *90*, 14–30. doi: 10.1016/j.jml.2016.03.003 65, 67

Hartsuiker, R. J., & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition*, *20*(2), 219–234. doi: 10.1017/s1366728915000164 13, 24, 26, 27, 33, 34, 35, 36, 37, 53, 54, 55, 61, 63, 65, 75, 76, 121, 122, 127, 137, 150, 151, 152, 155, 160, 162, 163, 168, 172

Hartsuiker, R. J., & Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychologica*, *128*(3), 479–489. doi: 10.1016/j.actpsy.2007.08.005 20, 34, 60, 83, 84, 113, 124, 128, 157

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. doi: 10.1111/j.0956-7976.2004.00693.x 24, 27, 34, 60, 65, 122, 155, 160, 167, 172

Hsieh, Y. (2017). Structural priming during sentence comprehension in Chinese–English bilinguals. *Applied Psycholinguistics*, *38*(3), 657–678. doi: 10.1017/s0142716416000382 123

Humphreys, G. F., & Gennari, S. P. (2014). Competitive mechanisms in sentence processing: Common and distinct production and reading comprehension networks linked to the prefrontal cortex. *NeuroImage*, *84*, 354–366. doi: 10.1016/j.neuroimage.2013.08.059 119, 157

Hwang, H., Shin, J.-A., & Hartsuiker, R. J. (2018). Late bilinguals share syntax unsparingly between L1 and L2: Evidence from crosslinguistically similar and different constructions. *Language Learning*, *68*(1), 177–205. doi: 10.1111/lang.12272 25, 123, 126

Indefrey, P. (2018). *The relationship between syntactic production and comprehension.* Oxford University Press. doi: 10.1093/oxfordhb/9780198786825.013.20 82, 89, 123, 155

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1-2), 101–144. doi: 10.1016/j.cognition.2002.06.001 18, 59

Jacobs, C. L., Cho, S.-J., & Watson, D. G. (2019). Self-priming in production: Evidence for a hybrid model of syntactic priming. *Cognitive Science*, *43*(7), e12749. doi:

10.1111/cogs.12749 86
Kantola, L., & van Gompel, R. P. G. (2011). Between- and within-language priming is the same: Evidence for shared bilingual syntactic representations. *Memory & Cognition*, *39*(2), 276–290. doi: 10.3758/s13421-010-0016-5 25, 60, 160
Kempen, G., Olsthoorn, N., & Sprenger, S. (2012). Grammatical workspace sharing during language production and language comprehension: Evidence from grammatical multitasking. *Language and Cognitive Processes*, *27*(3), 345–380. doi: 10.1080/01690965.2010.544583 82, 123, 155
Ketrez, F. N. (2012). *A student grammar of Turkish*. Cambridge University Press. doi: 10.1017/cbo9780511667077 64
Khoe, Y. H., Tsoukala, C., Kootstra, G. J., & Frank, S. L. (2021). Is structural priming between different languages a learning effect? Modelling priming as error-driven implicit learning. *Language, Cognition and Neuroscience*, 1–21. doi: 10.1080/23273798.2021.1998563 25, 35, 61, 165
Kidd, E., Tennant, E., & Nitschke, S. (2015). Shared abstract representation of linguistic structure in bilingual sentence comprehension. *Psychonomic Bulletin & Review*, *22*(4), 1062–1067. doi: 10.3758/s13423-014-0775-2 24, 123
Kim, C. S., Carbary, K. M., & Tanenhaus, M. K. (2014). Syntactic priming without lexical overlap in reading comprehension. *Language and Speech*, *57*(2), 181–195. doi: 10.1177/0023830913496052 85, 103, 104, 105, 107
Kootstra, G. J., & Şahin, H. (2018). Crosslinguistic structural priming as a mechanism of contact-induced language change: Evidence from Papiamento-Dutch bilinguals in Aruba and the Netherlands. *Language*, *94*(4), 902–930. doi: 10.1353/lan.2018.0050 61, 62, 77
Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, *109*(2-3), 68–74. doi: 10.1016/j.bandl.2008.06.002 122
Kroll, J. F., & Dussias, P. E. (2012). The comprehension of words and sentences in two languages. In *The handbook of bilingualism and multilingualism* (p. 216-243). John Wiley Sons, Ltd. doi: https://doi.org/10.1002/9781118332382.ch9 58
Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics*, *1*(1), 377–394. doi: 10.1146/annurev-linguist-030514-124937 58
Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In *The Oxford Handbook of Language Production* (p. 165-181). Oxford University Press. doi: 10.1093/oxfordhb/9780199735471.013.001 58
Kupisch, T. (2014). Adjective placement in simultaneous bilinguals (German–Italian) and the concept of cross-linguistic overcorrection. *Bilingualism: Language and Cognition*, *17*(1), 222–233. doi: 10.1017/s1366728913000382 62, 66, 76
Kupisch, T., & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, *22*(5), 564–582. doi: 10.1177/1367006916654355 66
Ledoux, K., Traxler, M. J., & Swaab, T. Y. (2007). Syntactic priming in comprehension. *Psychological Science*, *18*(2), 135–143. doi: 10.1111/j.1467-9280.2007.01863.x 16, 84
Lefcheck, J. S. (2016). scppiecewiseSEM/scp : Piecewise structural equation modelling in scpr/scp for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573–579. doi: 10.1111/2041-210x.12512 42
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. doi:

10.3758/s13428-011-0146-0 45, 67, 70, 109, 129, 140

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. The MIT Press. doi: 10.7551/mitpress/6393.001.0001 82

Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, *14*(1), 78–106. doi: 10.1016/0010-0285(82)90005-6 13

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1). doi: 10.1017/s0140525x99001776 18, 59, 82

Levshina, N. (2019). Bayesian logistic models with MCMCglmm: A brief tutorial. Retrieved from `http://www.natalialevshina.com/Documents/MCMCglmm_Tutorial.pdf` 72

Litcofsky, K. A., & van Hell, J. G. (2019). Bi-directional evidence linking sentence production and comprehension: A cross-modality structural priming study. *Frontiers in Psychology*, *10*. doi: 10.3389/fpsyg.2019.01095 84, 87, 88, 89

Long, J. (2022). jtools: Analysis and presentation of social scientific data [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=jtools/` (R package version 2.2.0) 98, 110

Loutfi, A. (2015, May). *The VP structure and the syntax-lexicon interface in Arabic.* Paper presented at the 3rd International Doctoral Seminar on Comparative Linguistics, Faculty of Letters and Human Sciences, Kenitra, May 14-15, 2015. Retrieved from `https://www.academia.edu/download/39059759/Conference.pdf` 64, 65

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703. doi: 10.1037/0033-295x.101.4.676 23, 86, 119, 157

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. doi: 10.1016/j.jml.2016.03.009 14, 34, 55, 85, 102, 103, 165

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. doi: 10.1044/1092-4388(2007/067) 45

Momma, S. (2022). Producing filler-gap dependencies: Structural priming evidence for two distinct combinatorial processes in production. *Journal of Memory and Language*, *126*, 104349. doi: 10.1016/j.jml.2022.104349 56

Montero-Melis, G., & Jaeger, T. F. (2019). Changing expectations mediate adaptation in L2 production. *Bilingualism: Language and Cognition*, *23*(3), 602–617. doi: 10.1017/s1366728919000506 39, 53

Muylle, M. (2020). *Development of syntactic representations in second-language learning* (Unpublished doctoral dissertation). Ghent University. 27, 162, 164

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. doi: 10.1016/j.jml.2021.104233 27, 37, 38

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*, *47*(9), 1471–1493. doi: 10.1037/xlm0000997 27, 39

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. doi: 10.1111/j.2041-210x.2012.00261.x 43, 48, 51

Öszoy, A. S. (2009). Argument structure, animacy, syntax and semantics of passivization in Turkish: A corpus-based approach. In *Corpus analysis and variation in linguistics* (pp. 259–279). John Benjamins Publishing Company. doi: 10.1075/tufs.1.16ozs 64

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., . . . Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. doi: 10.3758/s13428-018-01193-y 41, 108, 113, 140

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. doi: 10.1006/jmla.1998.2592 15, 18, 19, 20, 34, 55, 58, 59, 60, 61, 83, 84, 102, 118, 122, 126, 155, 157, 167, 171

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*(3), 427–459. doi: 10.1037/0033-2909.134.3.427 14, 34, 41, 83

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105–110. doi: 10.1016/j.tics.2006.12.002 82, 89, 123

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329–347. doi: 10.1017/s0140525x12001495 22, 82, 89, 155

Pickering, M. J., McLean, J. F., & Branigan, H. P. (2013). Persistent structural priming and frequency effects during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 890–897. doi: 10.1037/a0029181 82, 85, 102, 103, 124

Pickering, M. J., Traxler, M. J., & Crocker, M. W. (2000). Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. *Journal of Memory and Language*, *43*(3), 447–475. doi: 10.1006/jmla.2000.2708 23, 86, 119

Powell, M. J. D. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives.* Retrieved from `https://www.damtp.cam.ac.uk/user/na/NA_papers/NA2009_06.pdf` 42, 51, 93, 96, 98, 110, 114, 130, 133, 141, 146

Qualtrics. (2021, first release 2005). Qualtrics [Computer software manual]. Provo, Utah, USA. Retrieved from `https://www.qualtrics.com` 41, 50

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, *35*(4), 587–637. doi: 10.1111/j.1551-6709.2010.01165.x 20, 56, 59, 84, 152, 155, 165, 167, 171

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1-3), 107–142. doi: 10.1016/0010-0277(92)90041-f 18, 59

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*(3), 249–284. doi: 10.1016/s0010-0277(97)00027-9 18, 59

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, *56*(2), 153–171. doi: 10.1016/j.jml.2006.10.002 137

Segaert, K., Kempen, G., Petersson, K. M., & Hagoort, P. (2013). Syntactic priming and the lexical boost effect during sentence production and sentence comprehension: An fMRI study. *Brain and Language*, *124*(2), 174–183. doi: 10.1016/j.bandl.2012.12.003 82, 89, 102

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. doi: 10.1371/journal.pone.0024209 20, 21, 23, 56, 59, 63, 70, 84, 86, 104, 107, 117, 118, 119, 120, 125, 126, 127, 137, 152, 155, 156, 157, 158, 159, 161, 162, 167, 170, 171, 174

Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension–an fMRI study. *Cerebral Cortex*, *22*(7), 1662–1670. doi: 10.1093/cercor/bhr249 16, 82, 102

Segaert, K., Weber, K., Cladder-Micus, M., & Hagoort, P. (2014). The influence of verb-bound syntactic preferences on the processing of syntactic structures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1448–1460. doi: 10.1037/a0036796 20, 63, 125

Segaert, K., Wheeldon, L. R., & Hagoort, P. (2016). Unifying structural priming effects on syntactic choices and timing of sentence generation. *Journal of Memory and Language*, *91*, 59–80. doi: 10.1016/j.jml.2016.03.011 20, 42, 63, 125

Shaqråa, F. A. (2007). *Arabic: An essential grammar*. Routledge. doi: 10.4324/9780203088814 64

Sijyeniyo, E. (2023). *Structural representations in (late) learners of a second language: Effects of L2 proficiency and interventions on the priming of L2 syntactic structures* (Unpublished doctoral dissertation). University of Antwerp. 164

Sijyeniyo, E., Hartsuiker, R. J., & Bernolet, S. (2023). *L2 proficiency modulates structural priming effects of dissimilar L2 syntactic structures.* Manuscript in preparation. 126

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1521–1543. doi: 10.1037/0278-7393.24.6.1521 23, 86

Swain, M. (1985). *Communicative competence: Some roles of comprehensible input and comprehensible output in its development.* Newbury House. 122

Thothathiri, M., & Snedeker, J. (2008a). Give and take: Syntactic priming during spoken language comprehension. *Cognition*, *108*(1), 51–68. doi: 10.1016/j.cognition.2007.12.012 16, 22, 83, 85, 86, 103, 122

Thothathiri, M., & Snedeker, J. (2008b). Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language*, *58*(2), 188–213. doi: 10.1016/j.jml.2007.06.012 83, 85, 103, 122

Tooley, K. M., & Bock, J. K. (2014). On the parity of structural persistence in language production and comprehension. *Cognition*, *132*(2), 101–136. doi: 10.1016/j.cognition.2014.04.002 87, 88, 89, 102, 103, 140

Tooley, K. M., Swaab, T. Y., Boudewyn, M. A., Zirnstein, M., & Traxler, M. J. (2014). Evidence for priming across intervening sentences during on-line sentence comprehension. *Language, Cognition and Neuroscience*, *29*(3), 289–311. doi: 10.1080/01690965.2013.770892 103, 109, 124

Tooley, K. M., & Traxler, M. J. (2010). Syntactic priming effects in comprehension: A critical review. *Language and Linguistics Compass*, *4*(10), 925–937. doi: 10.1111/j.1749-818x.2010.00249.x 85, 102

Tooley, K. M., & Traxler, M. J. (2018). Implicit learning of structure occurs in parallel with lexically-mediated syntactic priming effects in sentence comprehension. *Journal of Memory and Language*, *98*, 59–76. doi: 10.1016/j.jml.2017.09.004 103, 124

Tooley, K. M., Traxler, M. J., & Swaab, T. Y. (2009). Electrophysiological and behavioral evidence of syntactic priming in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 19–45. doi: 10.1037/a0013984 16, 83, 84, 103, 124

Traxler, M. J. (2008). Lexically independent priming in online sentence comprehension. *Psychonomic Bulletin & Review*, *15*(1), 149–155. doi: 10.3758/pbr.15.1.149 85, 103, 122, 124

Traxler, M. J. (2015). Priming of early closure: evidence for the lexical boost during sentence comprehension. *Language, Cognition and Neuroscience*, *30*(4), 478–490. doi: 10.1080/23273798.2014.933243 84, 103, 124

Traxler, M. J., & Tooley, K. M. (2008). Priming in sentence comprehension: Strategic or syntactic? *Language and Cognitive Processes*, *23*(5), 609–645. doi: 10.1080/01690960701639898 16, 83

Traxler, M. J., Tooley, K. M., & Pickering, M. J. (2014). Syntactic priming during sentence comprehension: Evidence for the lexical boost. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 905–918. doi: 10.1037/a0036377 17

van Gompel, R. P. G., & Arai, M. (2018). Structural priming in bilinguals. *Bilingualism: Language and Cognition*, *21*(3), 448–455. doi: 10.1017/s1366728917000542 25, 60, 62, 65, 66, 76, 160, 161, 165

Weber, K., & Indefrey, P. (2009). Syntactic priming in German–English bilinguals during sentence comprehension. *NeuroImage*, *46*(4), 1164–1172. doi: 10.1016/j.neuroimage.2009.03.040 16

Wei, H., Boland, J. E., Brennan, J., Yuan, F., Wang, M., & Zhang, C. (2018). Lexicalized structural priming in second language online sentence comprehension. *Second Language Research*, *34*(3), 395–416. doi: 10.1177/0267658317723684 25, 123, 124

Wei, H., Boland, J. E., Cai, Z. G., Yuan, F., & Wang, M. (2019). Persistent structural priming during online second-language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(2), 349–359. doi: 10.1037/xlm0000584 124

Wei, H., Dong, Y., Boland, J. E., & Yuan, F. (2016). Structural priming and frequency effects interact in Chinese sentence comprehension. *Frontiers in Psychology*, *7*. doi: 10.3389/fpsyg.2016.00045 109, 140

Wheeldon, L. R., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *The Quarterly Journal of Experimental Psychology Section A*, *44*(4), 723–761. doi: 10.1080/14640749208401307 63

Wheeldon, L. R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language*, *33*(3), 332–356. doi: 10.1006/jmla.1994.1016 63

Witzel, J., & Forster, K. (2014). Lexical co-occurrence and ambiguity resolution. *Language, Cognition and Neuroscience*, *29*(2), 158–185. doi: 10.1080/01690965.2012.748925 105, 138

Witzel, N., Witzel, J., & Forster, K. (2012). Comparisons of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research*, *41*(2), 105–128. doi: 10.1007/s10936-011-9179-x 105, 138

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. doi: 10.1016/j.cogpsych.2007.04.002 37

Ziegler, J., & Snedeker, J. (2019). The use of syntax and information structure during language comprehension: Evidence from structural priming. *Language, Cognition and Neuroscience*, *34*(3), 365–384. doi: 10.1080/23273798.2018.1539757 22, 85, 86, 88, 90, 99, 103, 124, 158

Appendix A

Statistical models

A.1 Models for Chapter 1

Experiment 1: Control group

Table A.1.1: Model output for control group (N = 431, log-likelihood = -149.5).
glmer(TargetCondition ~ PrimeCondition + Intervention + (1 | Participant) + (1 | TargetItem))

	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**

Experiment 2: Longitudinal study

Table A.1.2: Model output for longitudinal study (N = 1,551, log-likelihood = -710.4). *glmer(TargetCondition ~ PrimeCondition + Session + Intervention + PrimeCondition:Session + Session:Intervention + (1 | Participant) + (1 | TargetItem))*

	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.

Experiment 3: Cross-sectional study

Table A.1.3: Model output for cross-sectional study (N = 664, log-likelihood = -218.9). *glmer(TargetCondition ~ PrimeCondition + Proficiency + Intervention + (1 | Participant) + (1 | TargetItem))*

	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***

A.2 Models for Chapter 2

Pre-experimental baseline

Table A.2.1: Model output for pre-experimental baseline (*N* = 1,488, sample size = 1,467). *MCMCglmm(TargetCondition ~ -1 + trait + trait:(PrimeCondition*Language), random = ~us(trait):Participant + us(trait):Item)*

	Posterior means	Lower confi-dence interval (95%)	Higher confi-dence interval (95%)	Effective sample size	*p*-value
SP(Target)	-4.2228	-6.0123	-2.6284	824.2	<.001***
PFP(Target)	-4.4675	-6.0931	-2.7586	318.6	<.001***
SP(Target) * Arabic	-0.1929	-1.2379	0.8284	1121.8	0.731
PFP(Target) * Arabic	0.2488	-0.7235	1.186	1088.5	0.608
SP(Target) * Turkish	0.3729	-0.7433	1.4646	1105.9	0.503
PFP(Target) * Turkish	-0.2778	-1.3602	0.7116	920.9	0.605

Priming experiment

Table A.2.2: Model output for priming experiment (*N* = 4,460, sample size = 1,467). *MCMCglmm(TargetCondition ~ -1 + trait + trait:(PrimeCondition*Language), random = ~us(trait):Participant + us(trait):Item)*

	Posterior means	Lower confi-dence interval (95%)	Higher confi-dence interval (95%)	Effective sample size	*p*-value
PFP(Target)	4.34957	3.46535	5.28872	1026.5	<.001***
PMP(Target)	-0.67464	-2.01484	0.76276	516.8	0.338
PFP(Target) * PFP(Condition)	0.29179	-0.37154	0.95492	763.7	0.374
PMP(Target) * PFP(Condition)	0.14085	-0.75489	1.01729	753.3	0.759
PFP(Target) *PMP(Condition)	0.09323	-0.53647	0.74589	756.8	0.787
PMP(Target) * PMP(Condition)	1.44412	0.54984	2.29962	719.7	<.001***
PFP(Target) * Arabic	-1.78194	-2.66735	-0.81326	1073.9	<.001***
PMP(Target) * Arabic	-2.40054	-4.19024	-0.71297	486.9	<.05*
PFP(Target) * Turkish	-0.5561	-1.55982	0.33301	875.5	0.249
PMP(Target) * Turkish	-2.41027	-4.27974	-0.46438	272.9	<.01**
PFP(Target) * PFP(Condition) * Arabic	0.12681	-0.64554	1.00989	927.7	0.790
PMP(Target) * PFP(Condition) * Arabic	-1.91803	-3.72541	-0.3705	246.5	<.05*
PFP(Target) * PMP(Condition) * Arabic	-0.01014	-0.85974	0.75419	848.9	0.991
PMP(Target) * PMP(Condition) * Arabic	-0.25224	-1.40039	1.03811	661.1	0.669
PFP(Target) * PFP(Condition) * Turkish	-0.41798	-1.31263	0.43647	869	0.352
PMP(Target) * PFP(Condition) * Turkish	-1.19233	-2.96143	0.22406	344.6	0.139
PFP(Target) * PMP(Condition) * Turkish	-0.46538	-1.31789	0.30697	802.3	0.263
PMP(Target) * PMP(Condition) * Turkish	0.39317	-0.97519	1.856	479.8	0.588

A.3 Models for Chapter 3

Experiment 1: Between participants

Table A.3.1: Model output for production data (between participants) (N = 2,195, log-likelihood = -1,214.2).
glmer(TargetCondition ~ PrimeCondition + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	Wald's *Z*	*p*-value
(Intercept)	-0.38	0.2717	-1.393	0.16
Prime DO baseline	-0.23	0.1493	-1.557	0.12
Prime DO	0.28	0.1479	1.903	<.1
Prime PO	-0.27	0.1489	-1.818	<.1

Table A.3.2: Model output for comprehension data (first click, between participants) (N = 5,797).
*lmer(RT ~ PrimeType * SentenceCondition + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	2.99	0.0625	128.10	47.755	<.001***
Type antiprimed	-0.01	0.0216	5634.00	-0.678	0.50
Type primed	-0.01	0.0215	5634.00	-0.343	0.73
Condition DO	0.01	0.0177	5634.00	0.592	0.55
Type(antiprimed) * Condition(DO)	0.02	0.0308	5635.00	0.561	0.58
Type(primed) * Condition(DO)	0.01	0.0307	5634.00	0.266	0.79

Table A.3.3: Model output for comprehension data (second click, between participants) (N = 5,334).
lmer(RT ~ PrimeType + SentenceCondition + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	0.73	0.0211	136.50	34.413	<.001***
Type antiprimed	-0.005	0.0077	5117.00	-.591	0.55
Type primed	-0.01	0.0077	5176.00	-.915	0.36
Condition DO	-0.02	0.0063	5178.00	-3.048	<.001**

Experiment 2: Within participants

Table A.3.4: Model output for production data (within participants) (N = 2,389, log-likelihood = -1,426.4).
glmer(TargetCondition ~ PrimeCondition + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	Wald's *Z*	*p*-value
(Intercept)	-0.48	0.2224	-2.169	.03*
Prime DO baseline	-0.002	0.1374	-0.016	.99
Prime DO	0.40	0.1362	2.911	<.01**
Prime PO	-0.09	0.1372	-0.666	.51

Table A.3.5: Model output for comprehension data (first click, within participants) (N = 5,837).
*lmer(RT ~ PrimeType * SentenceCondition + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	3.01	0.0419	264.50	73.449	<.001***
Type antiprimed	-0.001	0.0171	5578.00	-0.060	0.95
Type primed	-0.01	0.0171	5578.00	-0.411	0.68
Condition DO	0.01	0.0140	5579.00	0.984	0.33
Type(antiprimed) * Condition(DO)	0.02	0.0242	5578.00	0.810	0.42
Type(primed) * Condition(DO)	-0.003	0.0242	5578.00	-0.128	0.90

Table A.3.6: Model output for comprehension data (second click, within participants) (N = 5,586).
lmer(RT ~ PrimeType + SentenceCondition + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	0.74	0.0173	259.00	42.938	<.001***
Type antiprimed	-0.01	0.0071	5331.00	-1.524	0.13
Type primed	-0.003	0.0071	5330.00	-0.380	0.70
Condition DO	-0.01	0.0058	5331.00	-2.069	<.05*

Analysis of self-priming

Table A.3.7: Model output for analysis of self-priming (DO structure) (N = 4,584, log-likelihood = -2,959.1).
*glmer(TargetCondition ~ PrimeCondition + RunningCount(DO) * Experiment + (1 | Participant))*

	Coefficient	*SE*	Wald's *Z*	*p*-value
(Intercept)	-0.40	0.1257	-3.217	<.01**
PrimeCondition (PO baseline)	0.07	0.09098	0.814	0.42
PrimeCondition (DO)	0.34	0.09083	3.779	<.001***
PrimeCondition (PO)	-0.07	0.09132	-0.74	0.46
Running Count DO	0.02	0.01231	1.667	<.1.
Experiment (within)	-0.37	0.14416	-2.547	<.05*
Running Count DO * Experiment (within)	0.12	0.02374	4.897	<.001***

Table A.3.8: Model output for analysis of self-priming (PO structure) (N = 4,584, log-likelihood = -2,316.5).
*glmer(TargetCondition ~ PrimeCondition + RunningCount(PO) * Experiment + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	Wald's *Z*	*p*-value
(Intercept)	1.76	0.28413	6.204	<.001***
PrimeCondition (DO baseline)	-0.08	0.10347	-0.820	0.41
PrimeCondition (DO)	0.39	0.10331	3.789	<.001***
PrimeCondition (PO)	-0.20	0.10275	-1.959	<.1.
Running Count PO	-0.37	0.01776	-20.889	<.001***
Experiment (within)	0.10	0.12757	0.819	0.41
Running Count PO * Experiment (within)	-0.33	0.02932	-11.303	<.001***

A.4 Models for Chapter 4

Experiment 1: Comprehension

Table A.4.1: Model output for choice data in comprehension (N = 6,912, log-likelihood = -4,543.6).
*glmer(TargetCondition ~ PrimeCondition * VerbOverlap + (1 | Participant) + (1 | TargetItem))*

	Coefficient	exp (Estimate)	*SE*	Wald's *Z*	*p*-value
(Intercept)	0.03	1.03	0.12	0.25	0.80
PrimeCondition(DO)	-0.35	0.71	0.09	-3.96	<.001***
PrimeCondition(PO)	0.16	1.17	0.09	1.81	<.1.
VerbOverlap(yes)	-0.06	0.95	0.11	-0.51	0.61
PrimeCondition(DO) * VerbOverlap(yes)	0.14	1.15	0.12	1.12	0.26
PrimeCondition(PO) * VerbOverlap(yes)	0.31	1.36	0.12	2.48	<.05*

Table A.4.2: Model output for reaction times to DO targets in comprehension (disambiguating word, N = 6,774).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + MazeChoice + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.750	0.026	306.900	263.242	<.001***
PrimeCondition(DO)	0.0002	0.009	6367.000	0.019	0.985
PrimeCondition(PO)	0.018	0.009	6367.000	1.906	<.1.
VerbOverlap(yes)	0.001	0.025	373.700	0.034	0.973
Choice(yes)	0.003	0.025	374.800	0.109	0.913

Table A.4.3: Model output for reaction times to DO targets in comprehension (spillover word 1, N = 6,877).
*lmer(log(RT) ~ PrimeCondition * VerbOverlap + MazeChoice + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.526	0.025	221.400	256.730	<.001***
PrimeCondition(DO)	0.007	0.012	6470.000	0.594	0.553
PrimeCondition(PO)	0.026	0.012	6468.000	2.095	<.05*
VerbOverlap(yes)	0.020	0.025	536.200	0.820	0.413
Choice(yes)	0.035	0.023	378.600	1.539	0.125
PrimeCondition(DO) * VerbOverlap(yes)	-0.028	0.017	6473.000	-1.666	<.1.
PrimeCondition(PO) * VerbOverlap(yes)	0.001	0.018	6473.000	0.050	0.960

Table A.4.4: Model output for reaction times to DO targets in comprehension (spillover word 2, $N = 6{,}803$).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + MazeChoice + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.647	0.023	243.200	283.185	<.001***
PrimeCondition(DO)	0.006	0.008	6401.000	0.729	0.466
PrimeCondition(PO)	0.017	0.008	6400.000	2.092	<.05*
VerbOverlap(yes)	0.024	0.022	379.000	1.099	0.272
Choice(yes)	-0.062	0.022	380.100	-2.801	<.01**

Table A.4.5: Model output for reaction times to PO targets in comprehension (disambiguating word, $N = 6{,}824$).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + MazeChoice + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.712	0.025	343.700	272.787	<.001***
PrimeCondition(DO)	0.002	0.009	6423.000	0.217	0.828
PrimeCondition(PO)	-0.001	0.009	6422.000	-0.131	0.896
VerbOverlap(yes)	0.004	0.025	378.100	0.151	0.880
Choice(yes)	0.035	0.025	379.200	1.328	0.154

Table A.4.6: Model output for reaction times to PO targets in comprehension (spillover word 1, $N = 6{,}935$).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + MazeChoice + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.478	0.022	271.700	288.773	<.001***
PrimeCondition(DO)	-0.001	0.008	6532.000	-0.079	0.937
PrimeCondition(PO)	-0.013	0.008	6532.000	-1.732	<.1.
VerbOverlap(yes)	-0.005	0.022	380.000	-0.217	0.829
Choice(yes)	0.003	0.022	381.100	0.130	0.897

Table A.4.7: Model output for reaction times to PO targets in comprehension (spillover word 2, $N = 6{,}911$).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + MazeChoice + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.489	0.021	343.400	313.474	<.001***
PrimeCondition(DO)	0.002	0.007	6508.000	0.307	0.759
PrimeCondition(PO)	0.009	0.007	6508.000	1.348	0.178
VerbOverlap(yes)	0.0002	0.021	381.000	-0.010	0.992
Choice(yes)	-0.024	0.021	381.900	-1.147	0.252

Table A.4.8: Model output for reaction times to PO targets in comprehension (spillover word 3, $N = 6{,}841$).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + MazeChoice + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.645	0.025	211.500	265.021	<.001***
PrimeCondition(DO)	0.005	0.007	6435.000	0.632	0.527
PrimeCondition(PO)	0.019	0.007	6435.000	2.718	<.01**
VerbOverlap(yes)	0.025	0.023	380.000	1.102	0.271
Choice(yes)	-0.072	0.023	380.900	-3.125	<.01**

Experiment 2: Production

Table A.4.9: Model output for choice data in production ($N = 6{,}629$, log-likelihood = -2,665.0).
*glmer(TargetCondition ~ PrimeCondition * VerbOverlap + (1 | Participant) + (1 | TargetItem))*

	Coefficient	exp (Estimate)	*SE*	Wald's *Z*	*p*-value
(Intercept)	0.79	2.21	0.2875	2.754	<.01**
PrimeCondition(DO)	-0.17	0.84	0.1156	-1.505	0.13
PrimeCondition(PO)	0.28	1.32	0.1171	2.381	<.05*
VerbOverlap(yes)	1.33	3.79	0.3377	3.949	<.001***
PrimeCondition(DO) * VerbOverlap(yes)	-0.47	0.63	0.1718	-2.724	<.01**
PrimeCondition(PO) * VerbOverlap(yes)	1.13	3.10	0.1900	5.958	<.001***

Table A.4.10: Model output for onset latencies in production ($N = 6{,}590$).
*lmer(log(RT) ~ TargetCondition * PrimeCondition * VerbOverlap + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	8.155	0.050	206.700	164.266	<.001***
TargetCondition(PO)	-0.196	0.028	6498.000	-7.029	<.001***
PrimeCondition(DO)	0.002	0.027	6363.000	0.063	0.950
PrimeCondition(PO)	0.001	0.028	6360.000	0.031	0.975
VerbOverlap(yes)	-0.013	0.059	371.800	-0.211	0.833
TargetCondition(PO) * PrimeCondition(DO)	-0.011	0.035	6365.000	-0.309	0.757
TargetCondition(PO) * PrimeCondition(PO)	0.024	0.036	6363.000	0.684	0.494
TargetCondition(PO) * VerbOverlap(yes)	0.103	0.041	6474.000	2.487	<.05*
PrimeCondition(DO) * VerbOverlap(yes)	0.014	0.043	6365.000	0.331	0.740
PrimeCondition(PO) * VerbOverlap(yes)	0.075	0.051	6369.000	1.467	0.143
TargetCondition(PO) * PrimeCondition(DO) * VerbOverlap(yes)	0.037	0.052	6366.000	0.705	0.481
TargetCondition(PO) * PrimeCondition(PO) * VerbOverlap(yes)	-0.118	0.059	6370.000	-2.000	<.05*

A.5 Models for Chapter 5

Experiment 1A. Production (Chinese-English group)

Table A.5.1: Model output for choice data in Chinese-English production (N = 6,062, log-likelihood = -2,212.0).
*glmer(TargetCondition ~ PrimeCondition * VerbOverlap + Proficiency + (1+PrimeCondition | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	Wald's Z	*p*-value
(Intercept)	1.791	0.377	4.757	<.001***
PrimeCondition(DO)	-0.192	0.175	-1.098	0.272
PrimeCondition(PO)	0.363	0.212	1.711	<.1.
VerbOverlap(yes)	-0.772	0.511	-1.510	0.131
Proficiency	0.224	0.245	0.912	0.362
PrimeCondition(DO) * VerbOverlap(yes)	-1.287	0.227	-5.665	<.001***
PrimeCondition(PO) * VerbOverlap(yes)	1.011	0.275	3.670	<.001***

Table A.5.2: Model output for onset latencies in Chinese-English production (N = 6,017).
lmer(log(RT) ~ TargetCondition + PrimeCondition + VerbOverlap + Proficiency + TargetCondition:PrimeCondition + TargetCondition:VerbOverlap + PrimeCondition:VerbOverlap + TargetCondition:Proficiency + PrimeCondition:Proficiency + VerbOverlap:Proficiency + TargetCondition:PrimeCondition:VerbOverlap + TargetCondition:PrimeCondition:Proficiency + TargetCondition:VerbOverlap:Proficiency + PrimeCondition:VerbOverlap:Proficiency + (1 + TargetCondition | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	8.311	0.081	217.046	103.026	<.001***
Target(PO)	-0.149	0.063	230.636	-2.385	<.05*
Prime(DO)	0.047	0.049	5753.665	0.968	0.333
Prime(PO)	-0.016	0.050	5750.021	-0.313	0.754
VerbOverlap(yes)	0.086	0.094	214.590	0.921	0.358
Proficiency	0.036	0.063	187.645	0.577	0.565
Target(PO) * Prime(DO)	0.019	0.058	5730.822	0.330	0.741
Target(PO) * Prime(PO)	0.062	0.060	5731.007	1.039	0.299
Target(PO) * VerbOverlap(yes)	-0.078	0.082	215.896	-0.949	0.344
Prime(DO) * VerbOverlap(yes)	-0.150	0.063	5756.199	-2.358	<.05*
Prime(PO) * VerbOverlap(yes)	0.019	0.072	5757.062	0.267	0.789
Target(PO) * Proficiency	-0.097	0.055	160.110	-1.767	0.079
Prime(DO) * Proficiency	0.010	0.037	5749.371	0.282	0.778
Prime(PO) * Proficiency	-0.044	0.039	5737.142	-1.140	0.254
VerbOverlap(yes) * Proficiency	-0.093	0.086	168.210	-1.078	0.282
Target(PO) * Prime(DO) * VerbOverlap(yes)	0.221	0.081	5743.374	2.727	<.01**
Target(PO) * Prime(PO) * VerbOverlap(yes)	-0.110	0.086	5743.012	-1.283	0.200
Target(PO) * Prime(DO) * Proficiency	0.012	0.040	5746.091	0.304	0.762
Target(PO) * Prime(PO) * Proficiency	0.074	0.043	5729.271	1.732	<.1.
Target(PO) * VerbOverlap(yes) * Proficiency	0.065	0.065	94.140	0.988	0.326
Prime(DO) * VerbOverlap(yes) * Proficiency	-0.040	0.039	5772.897	-1.028	0.304
Prime(PO) * VerbOverlap(yes) * Proficiency	-0.063	0.039	5774.771	-1.626	0.104

Experiment 1B. Production (Dutch-English group)

Table A.5.3: Model output for choice data in Dutch-English production (N = 6,815, log-likelihood = -2,785.0).
*glmer(TargetCondition ~ PrimeCondition * VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	Wald's Z	p-value
(Intercept)	1.501	0.286	5.221	<.001***
PrimeCondition(DO)	-0.128	0.121	-1.062	0.288
PrimeCondition(PO)	0.329	0.123	2.668	<.01**
VerbOverlap(yes)	0.061	0.348	0.175	0.861
Proficiency	0.073	0.242	0.300	0.764
PrimeCondition(DO) * VerbOverlap(yes)	-1.140	0.169	-6.741	<.001***
PrimeCondition(PO) * VerbOverlap(yes)	1.262	0.186	6.771	<.001***
PrimeCondition(DO) * Proficiency	0.137	0.113	1.213	0.225
PrimeCondition(PO) * Proficiency	0.190	0.116	1.642	0.101
VerbOverlap(yes) * Proficiency	0.078	0.345	0.227	0.820
PrimeCondition(DO) * VerbOverlap(yes) * Proficiency	-0.249	0.160	-1.557	0.119
PrimeCondition(PO) * VerbOverlap(yes) * Proficiency	-0.322	0.173	-1.865	<.1.

Table A.5.4: Model output for onset latencies in Dutch-English production (N = 6,688).
lmer(log(RT) ~ TargetCondition + PrimeCondition + VerbOverlap + Proficiency + TargetCondition:PrimeCondition + TargetCondition:VerbOverlap + PrimeCondition:VerbOverlap + TargetCondition:Proficiency + PrimeCondition:Proficiency + VerbOverlap:Proficiency + TargetCondition:PrimeCondition:VerbOverlap + TargetCondition:PrimeCondition:Proficiency + TargetCondition:VerbOverlap:Proficiency + PrimeCondition:VerbOverlap:Proficiency + (1 + TargetCondition | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	t-value	p-value
(Intercept)	8.072	0.041	181.600	197.774	<.001***
Target(PO)	-0.238	0.028	355.200	-8.360	<.001***
Prime(DO)	-0.007	0.026	6423.000	-0.266	0.790
Prime(PO)	0.028	0.027	6443.000	1.048	0.295
VerbOverlap(yes)	-0.023	0.046	248.300	-0.493	0.623
Proficiency	-0.057	0.031	226.900	-1.802	<.1.
Target(PO) * Prime(DO)	0.045	0.031	6381.000	1.452	0.147
Target(PO) * Prime(PO)	-0.021	0.032	6411.000	-0.669	0.503
Target(PO) * VerbOverlap(yes)	0.075	0.040	355.500	1.888	<.1.
Prime(DO) * VerbOverlap(yes)	0.029	0.035	6445.000	0.829	0.407
Prime(PO) * VerbOverlap(yes)	-0.037	0.043	6444.000	-0.866	0.387
Target(PO) * Proficiency	0.032	0.025	243.300	1.287	0.199
Prime(DO) * Proficiency	0.050	0.021	6441.000	2.390	<.05*
Prime(PO) * Proficiency	-0.018	0.023	6425.000	-0.785	0.433
VerbOverlap(yes) * Proficiency	-0.044	0.042	180.300	-1.039	0.300
Target(PO) * Prime(DO) * VerbOverlap(yes)	-0.030	0.044	6404.000	-0.679	0.497
Target(PO) * Prime(PO) * VerbOverlap(yes)	0.026	0.049	6420.000	0.535	0.593
Target(PO) * Prime(DO) * Proficiency	-0.085	0.022	6414.000	-3.896	<.001***
Target(PO) * Prime(PO) * Proficiency	-0.018	0.025	6395.000	-0.741	0.459
Target(PO) * VerbOverlap(yes) * Proficiency	0.015	0.029	123.100	0.523	0.602
Prime(DO) * VerbOverlap(yes) * Proficiency	-0.020	0.021	6432.000	-0.978	0.328
Prime(PO) * VerbOverlap(yes) * Proficiency	0.026	0.021	6395.000	1.221	0.222

Experiment 2A. Comprehension (Chinese-English group)

Table A.5.5: Model output for choice data in Chinese-English comprehension (N = 6,912, log-likelihood = -4,466.9).
*glmer(TargetCondition ~ PrimeCondition * VerbOverlap + Proficiency + PrimeCondition:VerbOverlap + PrimeCondition:Proficiency + VerbOverlap:Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	Wald's *Z*	*p*-value
(Intercept)	0.465	0.116	4.025	<.001***
PrimeCondition(DO)	-0.099	0.090	-1.108	0.268
PrimeCondition(PO)	-0.092	0.090	-1.031	0.303
VerbOverlap(yes)	-0.057	0.118	-0.484	0.628
Proficiency	-0.188	0.073	-2.558	<.05*
PrimeCondition(DO) * VerbOverlap(yes)	0.060	0.126	0.475	0.634
PrimeCondition(PO) * VerbOverlap(yes)	0.060	0.126	0.476	0.634
PrimeCondition(DO) * Proficiency	-0.007	0.063	-0.103	0.918
PrimeCondition(PO) * Proficiency	0.076	0.063	1.198	0.231
VerbOverlap(yes) * Proficiency	0.275	0.092	2.978	<.01**

Table A.5.6: Model output for reaction times to DO targets in Chinese-English comprehension (disambiguating word, N = 6,261).
*lmer(log(RT) ~ PrimeCondition * VerbOverlap * Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.354	0.057	436.400	110.651	<.001***
PrimeCondition(DO)	0.010	0.020	5845.000	0.531	0.595
PrimeCondition(PO)	0.011	0.020	5846.000	0.541	0.589
VerbOverlap(yes)	0.022	0.079	411.500	0.277	0.782
Proficiency	-0.135	0.053	409.700	-2.533	<.05*
PrimeCondition(DO) * VerbOverlap(yes)	-0.021	0.028	5846.000	-0.771	0.441
PrimeCondition(PO) * VerbOverlap(yes)	-0.015	0.028	5847.000	-0.553	0.580
PrimeCondition(DO) * Proficiency	-0.028	0.019	5853.000	-1.487	0.137
PrimeCondition(PO) * Proficiency	0.005	0.018	5852.000	0.245	0.807
VerbOverlap(yes) * Proficiency	0.049	0.080	410.900	0.608	0.543
PrimeCondition(DO) * VerbOverlap(yes) * Proficiency	0.081	0.028	5849.000	2.901	<.01**
PrimeCondition(PO) * VerbOverlap(yes) * Proficiency	0.038	0.028	5850.000	1.356	0.175

Table A.5.7: Model output for reaction times to DO targets in Chinese-English comprehension (spillover word 1, N = 6,275).
*lmer(log(RT) ~ PrimeCondition * VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.167	0.045	438.705	136.111	<.001***
PrimeCondition(DO)	-0.029	0.017	5866.110	-1.719	<.1.
PrimeCondition(PO)	-0.024	0.017	5867.386	-1.427	0.154
VerbOverlap(yes)	-0.030	0.063	420.106	-0.471	0.638
Proficiency	-0.094	0.031	379.075	-3.045	<.01**
PrimeCondition(DO) * VerbOverlap(yes)	0.033	0.024	5866.916	1.354	0.176
PrimeCondition(PO) * VerbOverlap(yes)	0.050	0.024	5867.595	2.039	<.05*

Table A.5.8: Model output for reaction times to DO targets in Chinese-English comprehension (spillover word 2, $N = 6{,}273$).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.304	0.053	419.200	118.541	<.001***
PrimeCondition(DO)	-0.008	0.014	5866.000	-0.595	0.552
PrimeCondition(PO)	0.004	0.014	5868.000	0.277	0.782
VerbOverlap(yes)	0.013	0.073	379.400	0.180	0.858
Proficiency	-0.105	0.037	379.300	-2.862	<.01**

Table A.5.9: Model output for reaction times to PO targets in Chinese-English comprehension (disambiguating word, $N = 7{,}452$).
*lmer(log(RT) ~ PrimeCondition * VerbOverlap * Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.342	0.055	429.100	115.246	<.001***
PrimeCondition(DO)	0.003	0.017	7030.000	0.190	0.849
PrimeCondition(PO)	0.002	0.017	7031.000	0.105	0.917
VerbOverlap(yes)	-0.021	0.076	403.300	-0.276	0.783
Proficiency	-0.155	0.051	404.000	-3.042	<.01**
PrimeCondition(DO) * VerbOverlap(yes)	0.038	0.024	7031.000	1.610	0.108
PrimeCondition(PO) * VerbOverlap(yes)	0.020	0.024	7031.000	0.823	0.411
PrimeCondition(DO) * Proficiency	0.008	0.016	7035.000	0.499	0.618
PrimeCondition(PO) * Proficiency	0.019	0.016	7035.000	1.202	0.230
VerbOverlap(yes) * Proficiency	0.089	0.076	403.400	1.164	0.245
PrimeCondition(DO) * VerbOverlap(yes) * Proficiency	-0.044	0.024	7033.000	-1.846	<.1.
PrimeCondition(PO) * VerbOverlap(yes) * Proficiency	-0.036	0.024	7035.000	-1.490	0.136

Table A.5.10: Model output for reaction times to PO targets in Chinese-English comprehension (spillover word 1, $N = 7{,}471$).
*lmer(log(RT) ~ PrimeCondition * VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.106	0.043	436.900	142.715	<.001***
PrimeCondition(DO)	-0.004	0.015	7054.000	-0.250	0.803
PrimeCondition(PO)	-0.022	0.015	7056.000	-1.518	0.129
VerbOverlap(yes)	-0.020	0.059	411.700	-0.334	0.739
Proficiency	-0.091	0.029	378.400	-3.160	<.01**
PrimeCondition(DO) * VerbOverlap(yes)	0.040	0.021	7055.000	1.933	<.1.
PrimeCondition(PO) * VerbOverlap(yes)	0.021	0.021	7056.000	1.013	0.311

Table A.5.11: Model output for reaction times to PO targets in Chinese-English comprehension (spillover word 2, $N = 7{,}472$).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.113	0.044	419.100	140.480	<.001***
PrimeCondition(DO)	0.022	0.010	7057.000	2.239	<.05*
PrimeCondition(PO)	0.003	0.010	7058.000	0.285	0.776
VerbOverlap(yes)	0.0004	0.059	378.600	-0.007	0.994
Proficiency	-0.090	0.030	378.600	-3.042	<.01**

Table A.5.12: Model output for reaction times to PO targets in Chinese-English comprehension (spillover word 3, $N = 7{,}458$).
*lmer(log(RT) ~ PrimeCondition * VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.336	0.056	434.100	114.071	<.001***
PrimeCondition(DO)	-0.006	0.017	7041.000	-0.343	0.732
PrimeCondition(PO)	-0.019	0.017	7042.000	-1.101	0.271
VerbOverlap(yes)	-0.015	0.076	406.500	-0.199	0.843
Proficiency	-0.110	0.037	378.700	-2.945	<.01**
PrimeCondition(DO) * VerbOverlap(yes)	0.011	0.025	7042.000	0.428	0.669
PrimeCondition(PO) * VerbOverlap(yes)	0.055	0.025	7042.000	2.213	<.05*

Experiment 2B. Comprehension (Dutch-English group)

Table A.5.13: Model output for choice data in Dutch-English comprehension ($N = 6{,}912$, log-likelihood = -4,542.75).
*glmer(TargetCondition ~ PrimeCondition * VerbOverlap * Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	Wald's *Z*	*p*-value
(Intercept)	0.174	0.114	1.530	0.126
PrimeCondition(DO)	-0.348	0.097	-3.578	<.001***
PrimeCondition(PO)	-0.147	0.092	-1.596	0.111
VerbOverlap(yes)	-0.027	0.111	-0.240	0.811
Proficiency	-0.154	0.079	-1.952	<.1.
PrimeCondition(DO) * VerbOverlap(yes)	-0.300	0.138	-2.173	<.05*
PrimeCondition(PO) * VerbOverlap(yes)	0.384	0.131	2.945	<.01**
PrimeCondition(DO) * Proficiency	0.077	0.097	0.790	0.429
PrimeCondition(PO) * Proficiency	0.040	0.093	0.431	0.667
VerbOverlap(yes) * Proficiency	0.210	0.111	1.884	<.1.
PrimeCondition(DO) * VerbOverlap(yes) * Proficiency	-0.274	0.138	-1.979	<.05*
PrimeCondition(PO) * VerbOverlap(yes) * Proficiency	-0.136	0.131	-1.039	0.299

Table A.5.14: Model output for reaction times to DO targets in Dutch-English comprehension (disambiguating word, $N = 6{,}810$).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.758	0.017	102.600	395.356	<.001***
PrimeCondition(DO)	-0.025	0.009	6458.000	-2.618	<.01**
PrimeCondition(PO)	0.015	0.010	6449.000	1.536	0.12
VerbOverlap(yes)	0.025	0.015	374.400	1.702	<.1.
Proficiency	-0.016	0.007	371.900	-2.224	<.05*

Table A.5.15: Model output for reaction times to DO targets in Dutch-English comprehension (spillover word 1, $N = 6{,}740$).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.541	0.015	94.410	425.523	<.001***
PrimeCondition(DO)	-0.013	0.008	6382.000	-1.628	0.103
PrimeCondition(PO)	0.018	0.008	6377.000	2.238	<.05*
VerbOverlap(yes)	-0.013	0.013	374.700	-1.008	0.314
Proficiency	-0.012	0.006	373.2	-1.875	<.1.

Table A.5.16: Model output for reaction times to DO targets in Dutch-English comprehension (spillover word 2, $N = 6{,}710$).
*lmer(log(RT) ~ PrimeCondition * VerbOverlap * Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.639	0.018	79.520	371.379	<.001***
PrimeCondition(DO)	0.005	0.011	6341.000	0.432	0.666
PrimeCondition(PO)	0.008	0.011	6342.000	0.768	0.446
VerbOverlap(yes)	-0.019	0.016	877.000	-1.199	0.231
Proficiency	-0.026	0.012	913.600	-2.249	<.05*
PrimeCondition(DO) * VerbOverlap(yes)	0.020	0.015	6348.000	1.300	0.194
PrimeCondition(PO) * VerbOverlap(yes)	0.019	0.016	6350.000	1.230	0.219
PrimeCondition(DO) * Proficiency	0.015	0.018	6388.000	1.241	0.215
PrimeCondition(PO) * Proficiency	-0.009	0.018	6341.000	-0.764	0.445
VerbOverlap(yes) * Proficiency	0.015	0.016	881.300	0.937	0.349
PrimeCondition(DO) * VerbOverlap(yes) * Proficiency	-0.034	0.016	6372.000	-2.176	<.05*
PrimeCondition(PO) * VerbOverlap(yes) * Proficiency	-0.011	0.016	6347.000	-0.669	0.504

Table A.5.17: Model output for reaction times to PO targets in Dutch-English comprehension (disambiguating word, N = 6,834).
*lmer(log(RT) ~ PrimeCondition * VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.737	0.018	130.700	364.965	<.001***
PrimeCondition(DO)	0.002	0.013	6468.000	0.159	0.873
PrimeCondition(PO)	0.017	0.013	6460.000	1.318	0.188
VerbOverlap(yes)	0.017	0.019	763.800	0.860	0.390
Proficiency	-0.019	0.008	385.500	-2.315	<.05*
PrimeCondition(DO) * VerbOverlap(yes)	0.044	0.019	6476.000	2.344	<.05*
PrimeCondition(PO) * VerbOverlap(yes)	-0.020	0.018	6461.000	-1.121	0.262

Table A.5.18: Model output for reaction times to PO targets in Dutch-English comprehension (spillover word 1, N = 6,874).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.420	0.014	95.290	443.896	<.001***
PrimeCondition(DO)	0.021	0.007	6515.000	2.891	<.01**
PrimeCondition(PO)	-0.009	0.007	6499.000	-1.217	0.224
VerbOverlap(yes)	0.016	0.012	381.600	1.302	0.194
Proficiency	-0.015	0.006	382.100	-2.396	<.05*

Table A.5.19: Model output for reaction times to PO targets in Dutch-English comprehension (spillover word 2, N = 6,884).
lmer(log(RT) ~ PrimeCondition + VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.495	0.015	100.600	439.831	<.001***
PrimeCondition(DO)	0.0003	0.007	6515.000	0.044	0.965
PrimeCondition(PO)	-0.003	0.007	6500.000	-0.510	0.610
VerbOverlap(yes)	0.002	0.013	377.700	0.143	0.887
Proficiency	-0.017	0.006	378.000	-2.646	<.01**

Table A.5.20: Model output for reaction times to PO targets in Dutch-English comprehension (spillover word 3, N = 6,804).
*lmer(log(RT) ~ PrimeCondition * VerbOverlap + Proficiency + (1 | Participant) + (1 | TargetItem))*

	Coefficient	*SE*	df	*t*-value	*p*-value
(Intercept)	6.604	0.018	76.320	373.094	<.001***
PrimeCondition(DO)	0.003	0.010	6427.000	0.330	0.741
PrimeCondition(PO)	0.006	0.010	6425.000	0.605	0.545
VerbOverlap(yes)	0.003	0.015	722.800	0.184	0.854
Proficiency	-0.026	0.006	380.700	-4.121	<.001***
PrimeCondition(DO) * VerbOverlap(yes)	0.017	0.014	6439.000	1.202	0.229
PrimeCondition(PO) * VerbOverlap(yes)	0.027	0.014	6425.000	1.939	<.1.