Prepositions in Language Production: a Picture-Word Interference Experiment

Jorge Ortiz Amo

Máster en Lingüística Aplicada al Inglés

MÁSTERES DE LA UAM

2019 - 2020

Facultad de Filosofía y Letras

Universidad Autónoma de Madrid

Prepositions in Language Production:
a Picture-Word Interference Experiment

Jorge Ortiz Amo

Tutor: Dr Jesús Romero-Trillo

Master's Degree in English Applied Linguistics,
Facultad de Filosofía y Letras,
Universidad Autónoma de Madrid

Academic year: 2019/2020

Number of Words: 13,223

Table of Contents

Introduction

Since the decade of the 80s, the study of language processing in psycholinguistics has flourished thanks to the influential work of Levelt (1989), Roelofs (1992), Dell (1986), Bock (1982), Ferreira (2008), and many others. The reason for this can be found in their experimental methods. They approached language processing from three main sources: error analysis, priming experiments, and computational models. Firstly, error analysis studies have been able to find typical settings in which errors occur, and then, trace their stages of processing. Secondly, priming experiments have been used to test hypotheses on the mental lexicon, understanding it as a network. Finally, computational models have been used to replicate actual results from experimental data and find out if the reason for these results is justified in the stated hypothesis. Thanks to these studies, holistic models have been able to emerge. One of the most influential works was the book Speaking by Levelt, which was cited more than 11,000 times. He described a detailed model for language production and comprehension.

However, although these models are wide and can explain many phenomena, most of them have used content words as their primary source of investigation. If we think about the nature of content words, we find its explanation. They are easier to study, most of them can be understood in image format, allowing researchers to investigate the different layers in their meaning processing. Furthermore, the entities that these words refer to are similarly shared among all languages, allowing the results to be considered universal. On the other hand, function words, or closed class, are not so easy to examine. They cannot be described with standardized images, thus preventing researchers from applying traditional lines of experimentation. Also, closed words are very different among languages, making it difficult to extrapolate to universal truths. In this paper, I will follow Harley's (2005) distinction between functional and content words; the former is a class of words that is reluctant to accept new members, while the latter is unlimited and accepts new members constantly.

The aim of this paper is precisely to study one type of closed word using traditional experimentation. This type of closed word is prepositions, specifically the preposition "of". The reason for choosing this preposition is that it is often used to combine two different nouns, creating phrases of Noun-preposition-Noun (NpN). In this way, the preposition can be studied as a bridge between the two nouns, but with a specific weight in processing. In this way, incorporating modifications, it allows experiments to study prepositions in traditional priming experiment settings.

In addition, I will study the characteristics of those NpNs in the mind incorporating two variables in the experiment. The first will be their frequency of use, since some NpNs are used to denote individual entities in the world (e.g. house of dolls), this should have an impact in the way they are processed. The second variable will examine the linear processing of the two nouns to check if the processing is different from the first noun to the second or vice versa.

This experiment will only work with the preposition "of". This means that the results cannot be applied to other prepositions nor any other functional category; nonetheless, it will serve as a departure point for the future study of these.

The structure of this paper will consist of a theoretical background, a description of the experiment, results, discussion, and conclusion. In the theoretical background, I will firstly include a general description of language from the point of view of cognitive science; then, I will outline the most important models for language production; after that, since the nature of functional words is closely related to grammar, I will describe the major steps in grammatical encoding; and finally, I will provide with an overview of experimentation with prepositions in cognitive studies. In the experiment section, I will explain the methods that have been used, the research questions, and the apparatus used to carry out the experiment. Then, in the results section, I will present the obtained responses to the experiment, and the statistical analysis of these. After that, the discussion section will explain the results in a theoretical situation. Finally, the conclusion section will serve to comment on the implications of the results and the restrictions of the experiment.

Theoretical background

Much literature has been created in regards to language production in the last decades. Yet, this paper is going to focus on the original ideas by Levelt, Roelofs, Collins & Loftus, Laurence & Margolis, Dell, Bock, and Ferreira (among others) in the 90s and 2000s. The reason for not including more modern work is simply because the direction that language production has taken in recent years is moving towards specific processes of cognition (e.g. executive control and working memory), and neurological cognitive architecture (regions of the brain that account for language): separating from linguistic models in the mind. Some examples of this can be found hereunder. Shao et al (2012) linked the original language production models to executive control processes employing span tasks; Jongman et al. (2015) linked the same models to attentional models with an investigation of individual differences; Montero-Melis et at. (2019) studied language production from the perspective of working memory; Kemmerer (2015) studied meanings of verbs in the precentral motor cortices. Much more work can be found in these directions, for a review of modern contributions to language production, see Roelofs & Ferreira, 2019. These fields of study surely are interesting, but they must be separated from the linguistic intention of the present study.

To begin with, I will include an overview of general cognitive processes related to language production that will serve to better understand the abstraction of the more specific lexical selection (in the following sections).

Language in the mind

Throughout human history, philosophers and psychologists have used metaphors to interpret the mind. Some of the metaphors compared the mind to a blank sheet in which impressions were made, for example. In recent decades, with the emergence of computers, most cognitive researchers use computational metaphors to explain and transcribe the ways in which our mind works. In this paper, these metaphors will be used to denote the unconscious processes that occur before a fragment of language is uttered.

Since the best way to grasp the complexity of human thinking is to use multiple methods, especially combining psychological and neurological experiments with computational models, the dominant theory among cognitive researches is the representational-computational understanding of the mind. This theory consists of two parts: the mental representations accounting for data structures (i.e. information stored in our minds) and the computational procedures (i.e. the mechanical operations that connect the representations in the way algorithms work)(Thagard 1996). We could consider them to be the instructions that our mind uses to work with the mental representations.

In this way, in speaking, we should consider the meanings of words to be the mental representations, and grammar to be the computational procedures, but they are not completely separated. When we speak, we have to use both types of knowledge, even if we only produce one word, we have to encode its grammatical features and sounds/letters. Still, other types of knowledge are required to perform any type of conversation. Let us now revise the different stages of language production.

The first step in speech production of any formed speaker is the illocutionary intention: what information is to be transferred. Intention is understood as what aim(s) is the speaker trying to achieve by means of communicative expression. In this sense, we can talk about goals and subgoals (e.g. in giving instructions about how to get to a place, the main goal would be to lead the addressee to the destination, and the subgoals would be the specific information of how to reach the destination: "take the first street to your left, then right…"). Each subgoal needs to be planned in order to produce a speech act. However, we also have to take into account that these plannings are influenced by the fact that the speaker counts on the addressee to infer some information that is not explicitly expressed about the goal or subgoals of an utterance. The plannings are divided into two major steps: first plannings are called macroplanning, and they result in a sequence of Speech-Acts intentions, containing intended mood (declarative, interrogative, or imperative) and content. It involves communicative intention and planning of an utterance. The second stage, microplanning, is in charge of the informational perspective of an utterance: the topic, the focus, and the way it attracts the addressee's attention (Levelt 1989). Perspective, in this sense, represents the pragmatic directions from the speaker to the addressee: the choices that the speaker makes so that this shares the perspective. Clark (1997) refers to the obtaining of perspective by joint attention, both speaker and addressee share the same focus of attention, referring to the same entity or event. Joint attention can be achieved physically (i.e. both interlocutors are facing or looking at the same entities) or linguistically (i.e. reinforcements of shared referents).

Another element of any coherent conversation is the storage of what has been said. The speaker has to keep in mind what has been conveyed by himself and the interlocutor throughout the whole event of conversation. The information about the content of the conversation that is accessible to the speaker is denominated discourse record. It is the internal representation of the organized information of what happened in the conversation. This phenomenon happens in the storage of working memory (i.e. "brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning, and reasoning" (Baddeley 1992)), in the process of the conversation, but lost in long-term memory. The nature of discourse record includes information about the type of discourse (e.g. everyday conversation, narration, lecture…), the topic of discourse (i.e. necessary for topic shift), and the content of the discourse (i.e. predication about people, events, names, etc. and relations between them). From the discourse information, interlocutors create discourse models.

In order to create discourse models, interlocutors require several types of information. First, common ground is the knowledge that both (or all the) interlocutors share, e.g. the existence of chairs and that most of them are made of wood. This creates the possibility of a speaker uttering "I saw a metal chair, and I bought it for my kitchen" expressing, among other things, that a metal chair is uncommon. The interlocutors must be able to understand the implicit knowledge that allows this utterance to exist with no explanation about the strangeness of a metal chair. Common ground does not need to be actually shared, just believed by the speaker. Aside, own contribution is the representation of the information that the speaker believes he has conveyed through the conversation. At the same level, interlocutor contribution is the information that the speaker believes that the interlocutor has conveyed in the course of the conversation. Finally, the last type of structure is information to be conveyed, the information that the interlocutors intend to express in the following course of the discourse. Let us now consider the stages that a message has to go through before being uttered.

Once the intention, with all these variables, has been prepared, the message is sent to the next step in language production, which will is called Lexical Selection.

Lexical Selection

In order to explain the access to words, I will use the classification in stages of Levelt (1989) and Levelt et al (1999) (Fig. 1). However, I will not describe any further than the surface structure. The reason for this is that, at this stage, the word is already selected and grammatically coded.

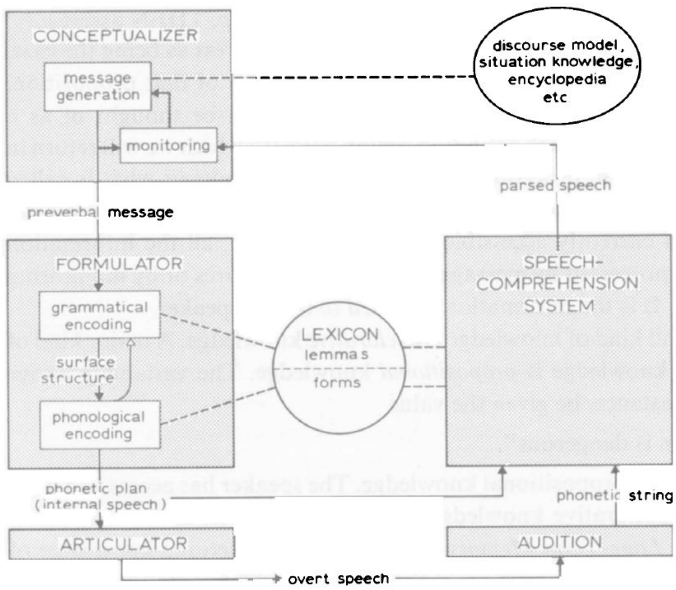


Fig 1. Levelt's (1989) model for speech production and comprehension based on three main components: conceptualizar, formulator/speech comprehension system, and articulator/audition; and also the specific encoding of each one.

Conceptualizer

Conceptualizing is the first process in language production, it requires all the information that intention includes (i.e. goal, form of utterance, language...). The processing system that serves to activate these processes is called Conceptualizer (Fig 1). It follows two kinds of knowledge: the first one is procedural knowledge (i.e. propositional expressions of the type IF X, THEN Y). The second type is declarative knowledge, which is subdivided into propositional knowledge (i.e. all the information stored through personal experience about the world and oneself, stored in long-term memory, also called encyclopedic knowledge), and situational knowledge (i.e. knowledge of conversation and context, e.g. acoustic and visual information). Based on these types of knowledge, the conceptualizer selects a concept.

Concepts are the mental representations of the real world, the information perceived by the senses. One of the characteristics of concepts is subjectivity: if concepts are created on the basis of experience, concepts must be subjective. For example, a chair is objective, but the mental representation of a chair must be subjective just because it is mental (Laurence and Margolis 1999) and individually acquired. The subjectivity is formed by all the experiential information of the individual related to the concept. Regarding learning of concepts, it is thought that some of them are innate, as the mental concept of object, and how an object should behave (i.e. when, in movement, they disappear behind another object and then reappear) (Thagard 1996). Other concepts are learnt by obtaining the general properties of different examples (e.g. discriminate dogs from other animals). Combination of different concepts into new ones is also a frequent way of creating new concepts. This can be made simple, e.g. pet fish, or more complex, e.g. blind lawyer.

Some models (named decompositional models) consider that the knowledge of a concept is molded in sets of slots learnt from happenings. In this way, slots do not need to hold for universal truths, but for typical truths. As an example, furniture has the slot to be typically made of wood as we most commonly encounter, but it is perfectly possible to find an exception to this slot. Keil (1979) indicated that concepts are hierarchically organized creating a taxonomy. In this hierarchy, the slots or properties of a concept pass down to the properties of another that are enclosed in the first one, a process that is known as inheritance, e.g. some of the typical characteristics of the concept chair are inferred from the typical characteristics of the concept furniture. However, other models (denominated non-decompositional models) state that this is not possible, if a concept activates its hierarchical concepts, and these subsequently activate other hierarchical concepts, when does the chain stop? Non-decompositional models find a solution in considering the concepts as individual, but related by links to other concepts. In this way, they are understood as a network in which concepts are related to other concepts.

Hence, concepts are understood as a network composed of all the concepts available to the speaker in which concepts are treated as nodes. Concept nodes are connected to others hierarchically, and also by means of the defining nodes, that are, at the same time, concepts as well (Dell 1986). Collins & Loftus (1975) examined these connections based on the work of Quilliam (1969), who applied a theory of semantic human processing in computer simulations of memory search and comprehension. They explained that connections or links are usually bidirectional: when a node points to another, the latter also leads to the former (Fig 2). Links of a node also present the characteristic of having critieralities (i. e. these can be understood as a number that expresses the importance from one node to another) indicating how important is a link from one concept to another concept. In this way, the double direction of the connections need not be equally relevant, e.g. an important link of

CHAIR leads to OBJECT, but it is not essential that a type of OBJECT is CHAIR for the understanding of OBJECT. Collins & Loftus (1975) also considered that there are links from concept nodes to the visual form of objects denoted by the concept. The selection of a concept node to be fully activated is always pragmatic and context-dependent (Clark 1997), "accompany" and "escort" have very similar meanings, but one or another will be used on the basis of the specific situation. However, both of them will have similar activation, but one will be more activated on the basis of the context. Once the concepts are selected, they pass to the next encoding process, and they receive the name preverbal message.

Formulator

The Formulator will, now, receive the preverbal message as input. So far, the intention and the conceptual information of the speaker are formulated, but it has yet to be coded lexically. To transfer that information into words, the message has to retrieve lexical items in the mental lexicon (i.e. the representation of stored words and their information). The first step of this transfer is the retrieval of the lemma. A lemma is the information of a word that includes meaning and also syntactical and morphological information about the words to be expressed (e.g. the lemma of "chair" bears the conceptual meaning of "CHAIR", and its use properties: it is a noun, it can be followed by adjectives, determiners…).

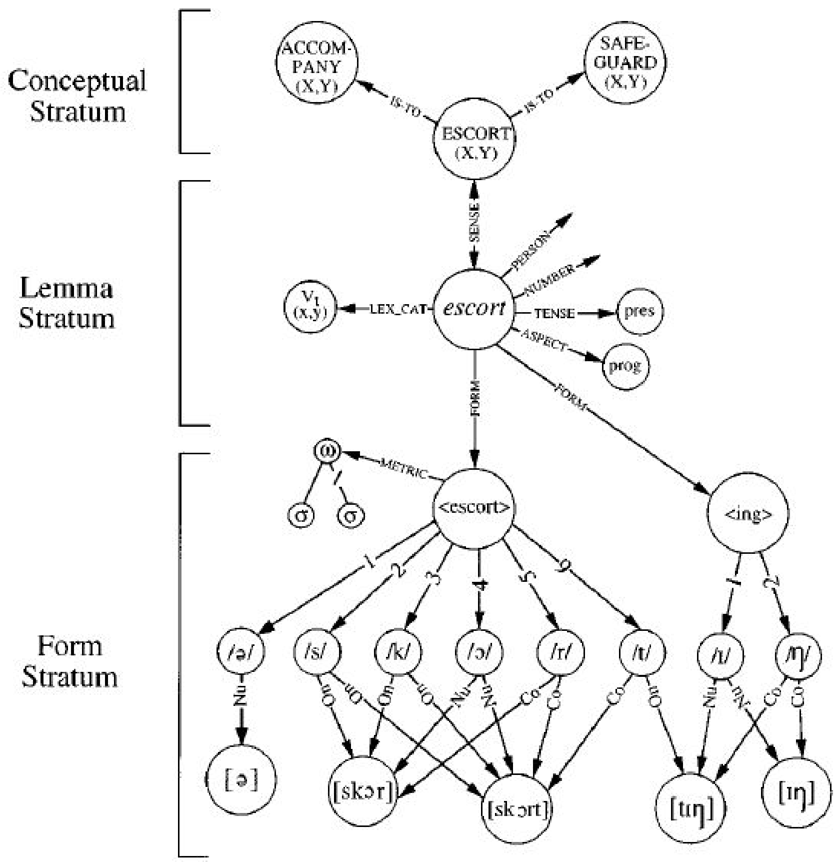


Fig 2. network lexical selection strata for the word "escort". ESCORT (X, Y) accounts for the concept, *escort* accounts for the lemma, and <escort> accounts for the form. Links are labeled by the relationship between nodes (e.g. SENSE, PERSON, LEXICAL CATEGORY…) (Levelt et al. 1999)

In this way, once the lexical concepts get activated, they will send a proportion of their activation to related lemma nodes (see Fig 2 for diagram of lexical selection). The highest activated lemma, that is, the lemma that has received more activation, will be retrieved. For example, in the verbalizing of the word "escort", the conceptualizer mainly selects ESCORT (X, Y) (i.e. X and Y refer to the arguments of the verb), but SAFEGUARD (X, Y) and ACCOMPANY (X, Y) (Figure 2) will also have activation because of their similarity in meaning. The three of these nodes will send a proportion of their activation to their lemma counterparts. However, the destination of the activation sent by SAFEGUARD (X, Y) and ACCOMPANY (X, Y) is only a proportion of a proportion of the activation. As a result, the most activated lemma will be the lemma "escort".

Then, following what Bobrow and Winograd (1977) denominated procedural attachment to nodes, each node has a procedure in which, when active, check whether the node links to the most activated node one level up (i.e. the lemma has to check if it corresponds to the most activated node in the concept). This is, the most activated lemma will be "escort", but it has to verify if it links to the most activated concept ESCORT (X, Y). This is denominated verification. Finally, the verified lemma node is sent to the form stratum to be codified phonologically and phonetically, however I will not describe it since it is not relevant for the present study.

This model in which the mental lexicon forms a network with concept nodes, lemma nodes and form nodes, and activation is spread was denominated spreading activation by Collins & Loftus (1975). As I have said, the activation of a node is transferred to same-level nodes, to one level up nodes, and to one level down nodes. To verify this model of lexical access, many experiments were tested. The method of analysis was mainly picture-word interference paradigm, many other authors also used this type of experimentation to describe the nature of the network (Dell, 1986; Stemberger, 1985; Dell & Reich,1981). In these experiments, the participants received a picture and a word on the same screen, and they only had to produce one of them, the other one served as a distractor. On the basis that pictures access directly the conceptual stratum, and words access directly the lemma stratum, the experiments were designed using different relations among words.

For example, in Glaser and Düngelhoff's (1984) experiment, the test included distractor words, and participants only had to produce the image. They created two groups, the control group had unrelated distractor words, and the experimental group had related distractor words. Besides, the experiment presented the stimuli (distractor word and target image) at different times. This was called Stimulus Onset Asynchrony (SOA) and was used to test the influence of time. Sometimes the word was presented before the image and vice versa. The SOAs that they used were, in milliseconds, -400, -300, -200, -100, 0, +100, +200, +300, +400 (minus symbol refers to the target image before the distractor word, and plus symbol refers to the distractor word before the target image). With these characteristics, the researchers recorded the answers of the participants and used their Response Time (RT) in milliseconds as the unit of study. Then the mean of each group was compared, and differences were found.

The reason for these results is that when the distractor word is presented, it activates the distractor lemma node; and when the target image was presented, it activated the target concept node. In order to produce the target word, the target lemma must be selected. In the case of unrelated distractor lemma activation, the target concept node would share its activation to the target lemma node, since the other lemma that is activated is unrelated, there is no competition, and thus, the target lemma node is selected easily and the response time of the participant is faster. On the other hand, if the distractor lemma was related, it produced competitiveness, because when the target concept node shares activation towards the target lemma node, there is another lemma node that is activated, and, since it is related, it would satisfy some of the characteristics of the target concept in the validation process, resulting in a naming latency (i.e. a delay in the production of the word).

Given that this process is complex to explain, I will illustrate it with an example. In an experiment that requires participants to produce the target image of a chair, the control group has the distractor word "dog" (unrelated), and the experimental group will have the distractor word "table" (related). The participant in the control group will activate the target concept node of CHAIR, and the distractor lemma node of "dog". Then, the target concept node of CHAIR will share activation to the target lemma node "chair". Now, the target lemma node "chair" will not have to compete with any other lemma node, because the lemma "dog" is not related and does not fulfill any characteristics of the target concept node CHAIR. On the other side, when the distractor word is "table", the target concept node shares its activation towards the target lemma node of "chair", however, in this case, the lemma of "chair" and "table" have to compete, because both of them activate the concept node of FURNITURE. Although the target lemma node is selected because it validates the target concept node CHAIR, the competitiveness between the two lemma nodes produces the naming latency.

The results that they found supported this hypothesis, related distractor words produced more naming latency than unrelated distractor words. They also found that different SOAs resulted in larger or shorter naming latencies. The SOAs that produced more naming latencies were -100, 0, and +100, especially 0 and +100. Schriefers, Meyer, and Levelt (1990) found a comparable semantic inhibition effect.

Although these results and many others seemed to support the spreading activation theory, there was a remote possibility that they could be justified by another cause. However, it was Roelofs (1992) who, using the data of Glaser and Düngelhoff (1984), created a computational simulation of the human lexical network. In this computational model, the parameters whose values were held constant across simulations are as follows: "(1) a real-time value in milliseconds for the smallest time interval (time step) in the model, (2) values for the general spreading rate at the conceptual stratum, and (3) values for the general spreading rate at the lemma stratum, (4) decay rate, (5) strength of the distractor input to the network, (6) time interval during which this input was provided, and (7) a selection threshold." Roelofs (1992). With this architecture, the input that participants received was applied to the computational model, then if the response time of the computational model is the same as the human participants, the hypothesis is verified. The program to carry out this validation is called Weaver ++, for a full description, see Roelofs (1997) As can be seen in Fig 3, the results of the computational model were surprisingly similar to those obtained by human participants, supporting the hypothesis of spreading activation.

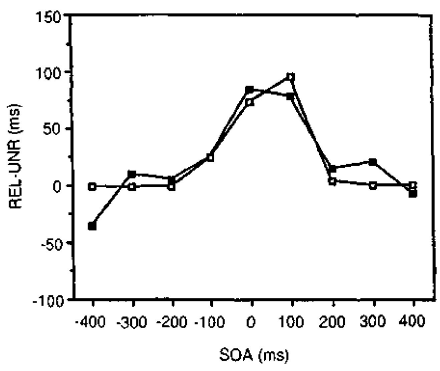


Fig 3. The real data (filled squares) are from Glaser and Düngelhoff (1984); the simulation data (open squares) are a simulation reported by Roelofs (1992), REL-UNR (ms) refers to the naming latencies and SOA (ms) refers to the presentation of the stimuli. Roelofs (1992)

Thus far, the mechanisms that govern the lexical access have been explained, yet, all these experiments and the computational models only account for content words, especially nouns and verbs. Function words, contrarily, have not found any place in this theory. The models do not prohibit them, because they could fit in the lemma stratum connected to nouns and other function words. However, they are not expressly included in the model, nor in the experimental studies. In a way, this makes sense, since functional words have more grammatical information than semantic information. What's more, the experimental method of picture-word interference paradigm cannot account for the function words, because of the characteristics of function words is that they do not represent real entities, but serve syntactical purposes; then, no universal image can describe them. In this way, to understand any type of functional word, we need to revise the model for grammatical encoding if we want to locate the functional words within language production.

Grammatical Encoding

For its part, Grammatical Encoding has been studied for a very long time and from many different perspectives. Since Chomsky proposed his generative grammar, many scholars have tried to build from his perspective or propose alternative solutions. Nonetheless, in this paper we are only going to focus on the aspects of grammar related to the production of language based on two types of evidence: speech error analysis and structural priming. Speech errors are amazingly useful because they allow us to see the systematic nature of grammar and what part of the encoding may be failing to produce those specific patterns, therefore, researchers can trace the origin of the errors to staged processes (Cutler 1988). Structural priming, on the other side, is fundamental to confirm the hypotheses drawn from speech errors and to create new ones (See Ferreira 2008 for a complete review of structural priming).

The first step in understanding the grammatical encoding is the question of whether lexical knowledge and grammatical knowledge are related or independent. And if they are related, is one dependent on the other, are they simultaneous? The first study that showed lexical influence in syntactic priming was Bock et al (1992). In their study, they primed participants with different sentences, they had the same grammatical properties (e.g. both primes were active subjects), but they changed the animacy or inanimacy of the entity of the subject, a semantic characteristic. They found that if the participants were primed with animate subjects, they would produce animate subjects in their utterances and vice versa. Furthermore, when they applied semantic priming to standard syntactic priming, combined, they observed that there was no difference to only syntactic priming, or to only semantic priming. This showed that both animacy priming and syntactic priming were independent. Both of them produced priming, but they did not influence each other. This finding also supports the hypothesis that constituents in the sentence (e.g. agents, themes..) are influenced by the characteristics of the semantics of the entity rather than a thematic specification at the message level. This means that the message is bound to the primitive features of words, not to the thematic roles.

I am now going to describe the stages involved in grammatical encoding. I have included Bock & Levelt's (1994) overview (Fig 4) because it accounts for all the processes that I am going to describe.

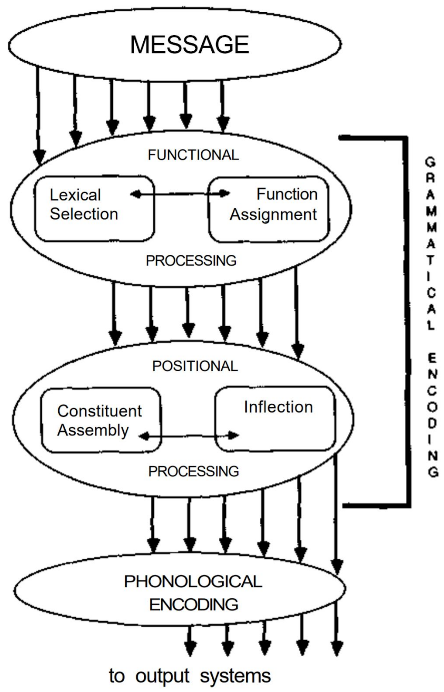


Fig 4. Bock & Levelt's (1994) overview of language production processes in regards to grammatical encoding. The first step is the message, then the functional encoding (consisting of lexical selection and function assignment), afterwards the positional processing (including constituent assembly and inflection) and finally phonological encoding.

Bock and Levelt's model follows four main steps of language processing. These levels are Message, Functional Processing, Positional Processing, and Phonological Encoding. The message accounts for all the information that we stated in previous sections under the same name. As a consequence, it will not be explained again.

Functional processing

Once the message has sent its activation, the functional processing starts. This comprises the lexical selection, explained above, and the function assignment. The former is in charge of selecting the lemmas and the thematic role of each element (e.g agent, theme, recipient....). The thematic roles are used to establish logical relationships between the elements in the future utterance. This is, in a sentence like "The bird was singing a song", the lemmas corresponding to "bird", "sing", and "song" get activated and assigned the roles of agent, verb, and theme based on the properties of the lemmas (e.g. bird is a countable, common, concrete noun) and the intention of the utterance.

Once the lemmas with the thematic roles are assigned, they require function assignment. This is in charge of transforming those into syntactic relations/grammatical relations. In this way, in an active sentence, the agent is assigned the nominative function/subject, and the theme is assigned the dative function/direct object, and thus, the order of the elements in the sentence can be arranged. However, the order is set by another mechanism called positional processing.

The functional encoder has been shown in error analysis studies, especially in phrase exchanges. In these errors, the problem was that a whole phrase, not only a word, was moved (e.g. "I went to the mechanical mouse for an economy five and dime" instead of "I went to the economy five and dime for a mechanical mouse"; Garrett, 1980). Since the whole phrase is moved unchanged, there must be a step in processing in which the phrases are assigned with a function. Also, Stemberg (1982) showed examples of misplacements of pronouns (e.g. "you must be too tight for them" instead of "they must be too tight for you"). In this type of sentences, we see how the whole phrase is moved, but the grammatical encoding of the word is appropriate (them in the error sentence is the right grammatically encoded word and not for they). Examples like these show that the function must be assigned in a different moment than the grammatical encoding of words. Similar examples can be found in Berg (1987) and Bock et al. (1992).

Regarding the order in which these functions get assigned, Bock and Cutting (1992) carried out an experiment of sentence completion. In their experiment, they tested a type of error called attraction error. In these types of errors, speakers produce wrong verb agreements when the subject was far from the verb in the sentence. For example, "*The claim that he had committed the crimes were rejected". Here, "the claim" is far from the verb "were", and therefore, the speaker is more likely to produce an error. Bock and Cutting used this error to design an experiment that could show whether the order of the function assignment was linear (i.e. with the order of final production, e.g. the nominative first, then the verb, then the theme) or was verb-driven (i.e. the verb is selected first and then its arguments). They presented participants with two primes that were complex subjects; later, the participants had to complete the sentences. The prompts had a noun (The claim) and either a post modifier phrase (about the newborn babies) or a postmodifier clause (that wolves had raised the babies). With this, they intended to examine whether the postmodifier phrases or the postmodifiers clauses produced more errors. They predicted that the postmodifier phrases would produce more errors, and the order would be verb-driven because the clause phrases present a new verb with new arguments, thus the word "babies" is bound to another verb (e.g. had raised). In this way, "babies" is not bound to "was/were", and it is less likely to modify it. Their results showed this. When there is a clause in the subject, errors were less likely to appear while long phrases in the subject produced more errors. These results were used to confirm that the order of production was verb-driven. In this manner, the verb gets assigned the function and its arguments are assigned a function afterwards.

Positional processing

The second main part of grammatical encoding is the positional processing, which includes two types of processing. The first is constituent assembly. In this step, the sentence is organized in a control hierarchy for the grammatical dependencies. It manages to produce the order of production that would be finally uttered. This hierarchy follows a frame, which, depending on the language, varies. This frame can be depicted in a traditional tree diagram that establishes the relations of the elements (see Fig 5 for an English mapping of the sentence "I think that the boy gave the book to the girl")

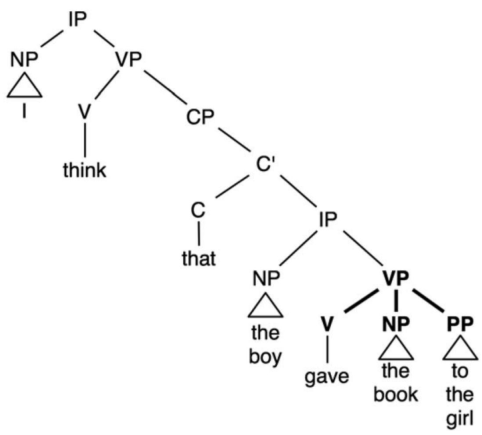


Fig 5. Phrase structure tree for a prepositional-object structure in an embedded clause. IP inflectional phrase; NP noun phrase; VP verb phrase; V verb; PP prepositional phrase; Det determiner; N noun; CP complementizer phrase; C complementizer. (Pickering & Ferreira 2008)

At the same time, another codification takes place, it is denominated the inflection process, in this stage, the most detailed parts of the utterance are codified, including the grammatical properties of verbs (e.g. progressive, past,…), or grammatical properties of words (number and gender of nouns, affixes for case languages...) and also, all the functional words that are required for the production of the sentence. Again, this was shown by error analysis, in this case, the error is called stranding. It refers to utterances where the lexical items are wrongly assigned, but the inflections are correctly assigned. Garrett (1975) found several cases of utterances like "You ordered up ending some fish dish" instead of "You ended up ordering some fish dish". In these examples, the inflections of the verbs are well placed, but the verbs are mislocated, indicating that a new stage must be created to encode the inflection. Sternberg (1985) also found errors in the placement of affixation, these are called shifts. In these types of error, the intended affixation of a word is placed in another word as in the imaginary example "he was give humming some chocolate", supporting the hypothesis of an individual stage for inflection assignment.

During the inflection assignment, functional words are also supposedly encoded, and I say supposedly because, as Bock and Levelt (1994) suggest, there is controversy about where this class of words is assigned. The reason for this knowledge vacuum is that functional words do not take place in speech errors in the same way that content words do. Even in sound errors, which are more likely to indiscriminate about lexical errors and structural errors, it is not common to find errors with functional words. In order to solve this gap of knowledge, two main hypotheses have been proposed.

The first, in the hands of Garret (1982), proposed that functional words were an intrinsic part of the grammatical frame. This is that the frame of a phrase or sentence includes in itself the functional words which that specific phrase or sentence requires. In this way, during functional assignment, each function is tagged with the specific characteristics including functional words. In this way, if the function that is assigned is a subject that is specific and plural, the frame that would be used is the one in Fig 6.

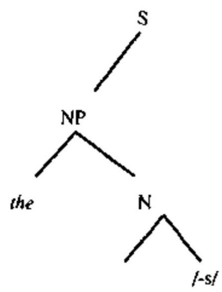


Fig 6. Functional words as part of the frame in the way that Garret (1982) explains. (Bock and Levelt 1994)

However, Lapointe (1985) noted a difference in the production of affixes and functional words, rejecting the proposal above. He carried out an analysis of speech errors in English and Italian aphasics. He realized that when there was an error with functional words, they tended to be omitted, while if the error occurred in the affixes, it tended to be a substitution of the affix by another affix, Therefore, the affixes and the functional words should be processed by different operators. Along with Dell, he proposed a mixed model (Lapointe and Dell, 1989). In this new model, the frame would not include the functional word, but a specification of it and also the affix, as seen in Fig 7.

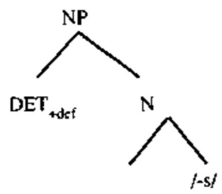


Fig 7. A mixed model for functional words described in a frame, but not included as a word in the way Lapointe and Dell (1989) describe. (Bock and Levelt 1994)

This paper will assume the mixed model to be true. In this way, the aim is to study what is that operator that Lapointe and Dell (1989) refer to, what processes are required to produce functional words.

Prepositions in the mind

As we have seen, if we try to locate functional words in the mental access to words, they are not explicitly included in Levelt's model, at least any experiment has shown otherwise. In a wide view, this makes sense since prepositions are not present in all the languages of the world (Saint-Dizier 2006). In some languages, the prepositions are included in the sentence as morphological marks (e.g. cases). Other languages do not use prepositions but postpositions (e.g. Hindi, Telugu, Tamil). It would be difficult to prove the concept representation of a preposition. However, they could perfectly exist in the lemma stratum, where the words are codified with their grammatical properties.

In regards to the grammatical encoding, the function words are supposed to be encoded in a different stage to content words. This is, content words get activated in the lexical selection, and function words get encoded or activated by other operator in the inflection assignment. Nonetheless, even if a different operator is in charge of activating content words and function words, we could expect that both of them are stored similarly.

It is plausible to hypothesize that functional words find a place in Roelofs' network. Even if they get activated at different moments in the processing stages, ultimately the content words and the functional words must be connected before the utterance is phonetically encoded, and the lemma stratum in the spreading activation network. Finding whether this is true or not would be interesting for the new questions that it would bring to the table. If functional words and content words are stored in the same network, does grammatical encoding access this network twice in different moments? and how are the content and functional words connected in the same network? If on the other side, functional words do not belong to Roelofs' network, where are closed words stored? Do they form a network similar to that for content words?

For this reason, the experiment that follows will investigate whether it is possible to find the presence of a functional word in the spreading activation network. In order to do so, the experiment will only use one type of functional word, a preposition. The results of this experiment will not be able to be applied to all functional words, but they will serve as a starting point where we can begin hypothesizing about the questions that I have mentioned above. The selection of prepositions is not arbitrary. Prepositions, in most cases, are used to connect two different content words, and thus they open a window for several ways of spreading activation mechanisms. Let us not take a look at the experimental literature that has been produced in regards to prepositions

The experimentation with prepositions has been reduced and is mostly related to syntactic priming. Levelt and Kelter (1982) carried out an experiment in which they asked shopkeepers different questions. In some cases, they asked in Dutch "Om hoe laat gaat uw winkel dicht?" (At what time does your shop close?) and in others they asked "Hoe laat gaat uw winkel dicht?" (What time does your shop close?). They found that questions with prepositions had answers with prepositions and questions without them had answers without them. However, these results have been explained because of structural priming, the prepositional structure is primed and therefore the participants answer with that construction; however, this does not have any implication with the questions I asked before.

Another example of an experiment with prepositions was Bock (1982) that used prepositions with structural priming. In their experiment, speakers heard and had to repeat prepositional-object or double-object prime sentences. These primes included either the preposition "to" (the secretary was taking a cake to her boss) or the preposition "for" (the secretary was baking a cake for her boss). The participants, then, had to describe target pictures that could be described with prepositional-object or double-object structures, but where the prepositional-object form always had to include the preposition "to". If function-word priming caused the appearance of structural priming, then the prepositional datives (with the preposition "to") should have only caused prepositional dative target descriptions. Instead, Bock found that prepositional datives with to or with for caused equivalent production of prepositional dative targets with to. This finding was replicated by Tree & Meijer (1999). These results suggest that what is primed is the production of an abstract prepositional phrase structure rather than the production of particular function words.

Again, this shows another example of prepositions priming the production of abstract structures, but not specifying its processing or place of storage. In the light of these findings, although syntactic priming has been successful in much grammatical encoding investigation (for a review, see Pickering & Ferreira 2008, and for a meta-analysis, see Mahowald et al 2016), it will not be useful in the present study as it would only produce results of grammatical processing, not prepositional processing.

In order to find how the prepositions are processed, the aim of the following experiment is to test if one example of a preposition can find a place in Roelofs' (1992) spreading activation network. Traditional experiments in that line of research have used the picture-word interference paradigm. As we have said before, in these experiments, the investigators present an image to be named and a word to be read or listened to. With these, the researchers have been able to trace the processes of producing and comprehending nouns and actions. Nevertheless, functional words cannot have any universal representation because of their semantics and grammatical use. Yet, a way of using picture-word interference method to test prepositions is proposed, but with the difference that it includes a system of double priming in which the first prime is a phrase consisting of Noun-preposition-Noun (NpN).

This type of phrases is not very frequent in English and other languages because they are usually substituted for a phrase where a noun premodifies the other noun (e.g. "dolls house" instead of "house of dolls"). However, this premodification is not possible in the Spanish language, and thus the preposition is compulsory to create those NpNs. Since the experiment and participants are going to be Spanish, we can take advantage of these constructions, which also allow to include some variables in the design of the experiment.

Experiment

The main objective of the experiment is to study prepositions in language production, however, by the nature of the experiment, other two analyses of the NpNs can be carried out. The first one is the study of these NpNs in terms of their frequency of use. This is due to the fact that some NpNs are used to name a single entity that cannot be named in any other way in Spanish (e.g. "chair of wheels", "house of dolls", "glasses of sun",...), hence, they could have different processings in the conceptual and lemma strata. These results could elucidate more information about the Roelofs' (1992) spreading activation network and the prepositions in it. The second one is the study of linear order of processing in constructions of NpN. Is there any difference in processing the first noun then preposition and then second noun than second noun, then preposition, and then first noun?

Research questions

(1) Do prepositions take part in spreading activation in Roelofs' (1992) network in the same way as content words?

(2) Is there any difference between high-frequency NpN phrases and low-frequency NpN phrases processing?

(3) Is presenting the N1 as prime different from presenting the N2?

Method

Procedure

In picture-word interference experiments, pictures directly access the conceptual nodes and words directly access the lemma nodes. In this way, since prepositions cannot be described with a picture, they can only be examined in their lemma nodes. Whether they have nodes in the conceptual stratum or not will not be answered in this paper. With these restrictions, the design for this experiment is based on the potential effect of preposition nodes when the preposition relates to two different nouns. This is, the way to observe the existence of preposition nodes will be done by testing if the preposition can further activation between two noun nodes. For example, "house" and "dolls" are not related, but if we include "of" and create "house of dolls", then those words become related by an intermediary node.

To find the effects of prepositions in the network, I will prime the participants with a construction like "house of dolls" with a noun, a preposition, and another noun. This construction will be denominated NpN (Noun preposition Noun). This prime will activate the lemmas of the first noun, the preposition, and the second noun; and, also, the connection between them. In this way, when the participants are primed with an NpN, then, if they receive the first noun (N1) of that construction, they will automatically activate the preposition and noun that were previously related to them because they have recently activated in that way. To exemplify this, if the participant is primed with "house of dolls", and then is presented with "house", he/she will unconsciously activate "of" and "dolls". On the other hand, if, instead of NpN, I prime the participant with the construction without the preposition "house dolls" only NN (Noun Noun), and then present the participant with the N1, it won't further activation to the N2 dolls, because the preposition node was not activated and it cannot serve as nexus.

With this distinction, I can create two groups, a control group that will be primed with a NN construction, and an experimental group that will be primed with NpN construction. After this priming, both groups will have to complete the same task. The task will be a standard picture-word interference test in which either the N1 or the N2 of the NpN will be distractors, and the participants will only produce the target image on the screen. The image will be semantically related to the other noun in the NpN that is not the prime. This is, the participant will be primed with either an NpN or a NN construction ("house of dolls" or "house dolls"), then both groups will have on the screen a picture related to one noun (castle, being related to house) and the other noun (dolls) written.

The logic behind this experiment is that if prepositions have nodes in the lemma stratum, in the experimental group, after the first priming (house of dolls), the second priming (dolls) will further activation to the "of" node and this one to the node of "house" and then, this hidden activation of "house" will produce competition with the production of the image castle and result in a naming latency (i.e. a delayed Response Time). Meanwhile, in the control group, after presenting the first prime "house dolls", the second prime (dolls) will not further activation to the node of "house" because they have not been previously connected with the node of the preposition. This will not produce the activation of "house", and then, no competition with the image castle, resulting in a shorter Response Time.

Materials

The participants had to complete 12 trials in total. Originally, the experiment was designed using 24 trials; however, after four pilot tests, the pilot participants stated that they became tired and, at the end of the experiment, they were not able to pay as much attention as they did in the beginning. For that reason, I decided to reduce it to 12 trials.

In order to answer the second research question, the NpN phrases had to be split into frequent and infrequent ones. In that way, 6 NpN were frequent and 6 were infrequent. The selection of the frequent NpNs was done by looking at the images in the set and selecting frequent NpN in which one of their nouns were related to some image. A total of 63 NpN were chosen as valid for the experiment. Since I only had to use 6, all the NpN were looked for in the CREA, the reference corpus for the current Spanish language to check for their frequency. The 6 most frequent NpNs were selected. On the other side, in order to choose their counterparts, the infrequent NpNs, the NpNs with 0 instances in the corpus were chosen. However, since they had been selected as possible for the frequent NpNs, either the N1 or the N2 were substituted by another noun to make sure that they were novel or completely infrequent in the minds of the participants. With the substitution, they were All the NpNs, frequency, and images can be found in Appendix A.

Regarding the third research question, the 6 frequent and 6 infrequent NpNs were subdivided depending on the noun that was used as distractor word, out of 6 frequent NpNs, 3 had N1 as the distractor, and 3 had the N2 as the distractor. The same was applied to the 6 infrequent NpNs.

Since the participants were native speakers of Spanish, the NpNs, and the distractor words that appeared in the experiment were written in Spanish. The instructions were also in Spanish, and the responses of all participants are in Spanish

The participants received the instructions written on the screen before they started. The instructions did not have duration, the experimenter told the participants to press the spacebar when they had read and understood the instructions. Then, the trials started. In each trial, the participants had to read aloud the first prime (NpN or NN). This screen lasted for 5 seconds. Then, on the next screen, the second distractor appeared at the same time as a short beep that lasted for 50 ms., 100 ms after the appearance of the word and the beep, the target image appeared, this is an SOA of +100. The reason for choosing 100 ms is that Collins and Loftus (1975) and Roelofs (1992) found more priming effects in related word-picture in -100 msec, 0 msec and +100 msec, as can be seen in image 3. Any of them would produce similar results, the specific selection of +100 was arbitrary.

The images used in the experiment come from the standardized set of 360 images by Moreno-Martínez & Montoro (2012), which are an alternative to the original 260 set of images by Snodgrass & Vanderwart M (1980). The reason for choosing these images is that they are in color, which makes a more natural environment, and also because these images were tested in Spanish participants, which would produce less naming problems with the Spanish participants in the present experiment.

Analysis

The unit of analysis in this experiment is the Response Time (RT) of participants to target images. The answers were measured in milliseconds. In each trial, a starting sound (musical note A) was used to know the exact moment in which the image appeared. This sound lasted for 50 msec., and the time was counted from the moment the beep sounded and the distractor word appeared to the moment that the participant made the first sound. Since the target image appeared 100 milliseconds after the distractor word and the beep, all the response times include 100 milliseconds in which the participant had not seen the image yet. It was designed in this way to avoid that the beep sounded when the target image appeared in case it would produce any effect in the results.

Once all the responses were collected, the average of each participant was calculated. This was done because the responses of the same participant to the twelve trials were dependent on the participant. To solve this problem, the mean of each participant in each group was taken into account.

In this way, 4 analysis were carried out: the response time of the control group was compared with the response time of the experimental group, the response of the frequent trial was compared to the response of infrequent trials, the frequent trials that had the N1 as distractors were compared with the frequent trials that had the N2 as distractors, and the infrequent trials that had the N1 as distractors were compared with the infrequent trials that had the N2 as distractors.

Participants

Since this experiment searches for general cognitive processes, any native speaker of Spanish would be appropriate to perform the task. However, due to the fact that some cognitive abilities decrease over time, there will be an attempt to conform the groups with similar age ranges. No other distinction will be made in the participants of each group on the grounds that gender, social group, ethnicity, and level of studies make a difference in these general processes.

The sampling technique used to find participants was snowball sampling, in which existing participants recruited new study subjects from among their acquaintances. Given that the intention for the groups was to be homogeneous in terms of age, some participants were rejected when they belonged to an age group that already had several participants.

32 Spanish native speakers participated in the experiment. All of them were healthy in regards to cognitive abilities and were assigned to the control and experimental groups randomly. The age of the participants ranges from 18 to 72. The total average age is 43,5 with a standard deviation of 17,8. Regarding each group, the average age of the control group is 41,5 with a standard deviation of 19,9; the average age of the experimental group is 45,5 with a standard deviation of 15,8. To see all the ages of participants, see Appendix C.

Apparatus

The stimuli were presented to participants using the software Psychopy, with which stimulus can be presented with any type of SOA. The creator of this program published a tutorial video on Youtube in which basic procedures in the program are explained (Jon Peirce 2010). A personal computer was used to run the experiment. The responses were recorded with a regular smartphone with an integrated microphone.

The recordings were then transferred to the personal computer and analyzed with Audacity, (i.e. an editing program which allows analyzing time in fractions of 1 msec). The extraction of the msec of the RT was done manually in Audacity looking at the spectrogram (this allowed to recognize the musical note A before the actual response of participants because it shows the different symmetrical frequencies of the note A, see Image 1 for the musical note and the response, see Appendix B for a view of a complete experimental in Audacity). Once the data were collected, they were transferred to a google spreadsheet to develop the means and standard deviations. Finally, R studio with the extension R Commander was the program used to develop the statistical results.

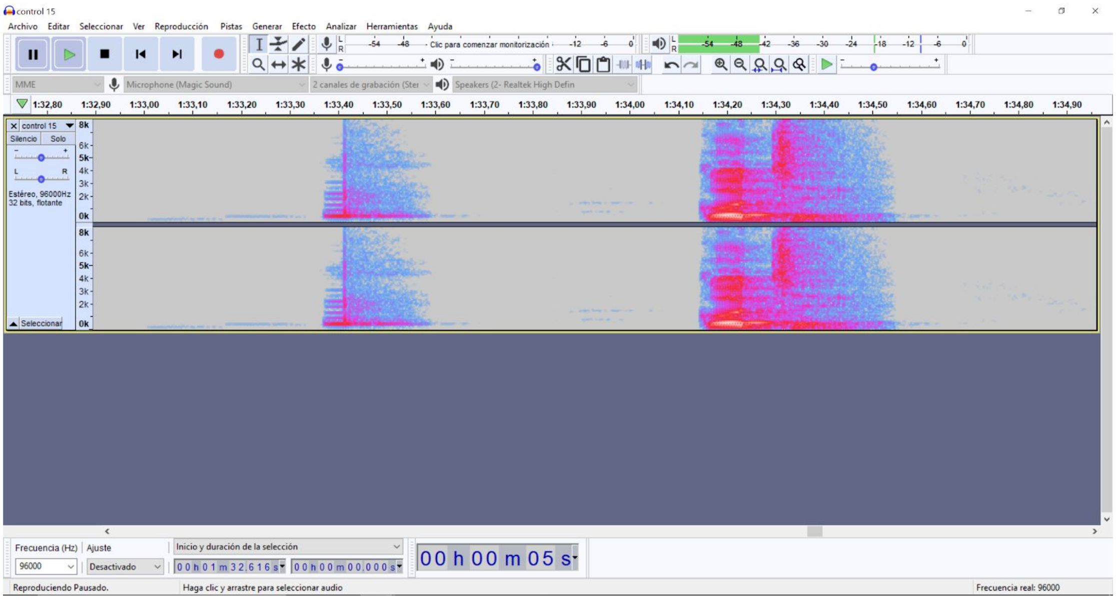


Image 1. An example of the musical note (on the left) and answer (on the right) in Audacity. The musical note has the straight parallel horizontal lines that express that it is a classical musical note, then the human response is clearly distinguishable for its strength. The spectrogram is doubled by default of the application, both spectrogram are identical.

Results

Some participants stated that the ninth target image "moon" was difficult to understand. This can be due to the fact that the moon was the only image in the set that had background color (the ski surrounding the moon). Yet, this image was standardized and was the same in both the control and experimental groups. This means that if it had produced any effect, it would have affected in the same way in both groups. For this reason, the results of these trials were included. Apart from this, the results were obtained with normality.

1st research question:

The RT of each participant in each trial can be found in Table 2 in Appendix 1. The total number of responses to trials was 384. However, each participant produced 12 responses and thus they should be considered dependent on its participant and placed in the sequence of trials; in this way, the mean of the 12 trials of each participant was taken into account, instead of the 12 responses for each participant. In this way, the number of results is equivalent to the number of participants. With a mean RT for each participant, the final number of responses is 32, of which 16 are in the control group and 16 in the experimental group (table 1).

Con Participant	*Con means*	*Exp Participant*	*Exp means*
con 1	*1345,5*	*exp 1*	*1068,833333*
con 2	*899,6666667*	*exp 2*	*941,3333333*
con 3	*1308,5*	*exp 3*	*1238,666667*
con 4	*1121,416667*	*exp 4*	*1039,416667*
con 5	*885,9166667*	*exp 5*	*1138,416667*
con 6	*971,4166667*	*exp 6*	*1723,583333*
con7	*986,4166667*	*exp 7*	*1030*
con 8	*1971,333333*	*exp 8*	*1095*
con 9	*1303,5*	*exp 9*	*942,8333333*
con 10	*1254*	*exp 10*	*1540*
con 11	*960,9166667*	*exp 11*	*1437,25*
con 12	*758,4166667*	*exp 12*	*846*
con 13	*760,3333333*	*exp 13*	*777,6666667*
con 14	*1112,25*	*exp 14*	*1224,333333*
con 15	*893,75*	*exp 15*	*900,1666667*
con 16	*990,25*	*exp 16*	*1009,333333*
	$\bar{x}$=1095		*$\bar{x}$=1122*
	sd=299,9		*sd=265,5*

Table 1. Averages of the 12 responses of each participant in the control and experimental group.

The control group produced a mean response time of $\bar{x}$=1095 msec and a standard deviation of SD=299.9; and the experimental group a mean of $\bar{x}$=1122 msec and a standard deviation of SD=265.5. These results apparently show a tendency towards the hypothesis of increased time in the experimental group. Nevertheless, the difference between the groups has to be tested statistically.

A Shapiro-Wilk test was used to check the normality of control and experimental groups. Neither group showed a normal distribution. With the significance level set at $\alpha=0.05$, the control group showed the p-value: 0.844; and the experimental group showed a p-value: 0.480.

Since both groups include non-parametric data and are independent, the test that will be used is a Mann Whitney U test (also known as Wilcoxon rank-sum test). In this test, medians are used for the comparison between groups. With a null hypothesis H0: control group RT=experimental group RT and an alternative hypothesis H1: Control group RT≠Experimental group RT. With a significance level set at $\alpha=0.05$, the alternative hypothesis is rejected at a p-value: 0.6106. With this result, the hypothesis is rejected.

2nd research question:

Regarding the second research question, a new analysis was carried out. This time the two groups are the 6 frequent trials in comparison with the 6 infrequent trials. Only the experimental group was taken into account because it showed the complete NpN as a prime, while the control group only showed the NN, without the preposition.

The data in each group was gathered by selecting the frequent and infrequent trials responses of each participant and calculating their mean independently, resulting in two different means for each participant, the mean of their frequent trials, and the mean of their infrequent trials (Table 2).

As shown in the table, the frequent trials obtained a mean of $\bar{x}$=1216,2 with a standard deviation of SD=314,486, while the infrequent obtained a mean of $\bar{x}$=1027,8, with a standard deviation of SD=214,065.

Participant	***Frequent trials mean***	***Infrequent trials mean***
exp 1	*1159*	*978,6666667*
exp 2	*1002,166667*	*880,5*
exp 3	*1429,5*	*1047,833333*
exp 4	*1163,5*	*915,3333333*
exp 5	*1182*	*1094,833333*
exp 6	*2010,666667*	*1436,5*
exp 7	*1161*	*899*
exp 8	*1205,666667*	*984,3333333*
exp 9	*970,5*	*915,1666667*
exp 10	*1612,333333*	*1467,666667*
exp 11	*1586,333333*	*1288,166667*
exp 12	*920*	*772*
exp 13	*844*	*711,3333333*
exp 14	*1328,833333*	*1119,833333*

exp 15	*882,3333333*	*918*
exp 16	*1001,833333*	*1016,833333*
	x̄=1216,2291	*x̄=1027,875*
	SD=*314,486*	SD=*214,065*

Table 2.Averages of each participant in their frequent and infrequent trials.

A Shapiro-Wilk test was used to check the normality of control and experimental groups. Neither of both groups showed normal distribution. With the significance level set at α=0.05, the frequent trials group showed the p-value: 0.083; and the experimental group showed a p-value: 0.126.

Opposite to the previous analysis, these results are dependent on the same participant, and since they have asymmetrical distribution, the test that I will apply is the Wilcoxon Signed-Rank Test. With a null hypothesis H0: frequent group RT=infrequent group RT and an alternative hypothesis H1: frequent group RT≠infrequent group RT. With a significance level set at α=0.05, the alternative hypothesis is validated at a p-value: 0.00078. With this result, the alternative hypothesis is approved.

3rd research question:

To answer the third research question, each group of frequent and infrequent trials was subdivided into two. The 6 frequent trials were divided into 3 frequent trials with the distractor word being the N1 (Frequent 1), and 3 frequent trials with the distractor word being the N2 (Frequent 2). The same happened in the infrequent group. The reason for dividing the trials into 4 groups instead of 2 groups was to avoid any influence of frequency and infrequency. The means of each participant in each group were considered, as shown in table 3. The total means and standard deviation of each group are also included in the table.

Participant	Frequent 1	Frequent 2	Infrequent 1	Infrequent 2
exp 1	*1129,333333*	*1188,666667*	*874*	*1083,333333*
exp 2	*1181*	*823,3333333*	*858,6666667*	*902,3333333*
exp 3	*1623*	*1236*	*1023*	*1072,666667*
exp 4	*1414,333333*	*912,6666667*	*837,6666667*	*993*
exp 5	*1142*	*1222*	*825*	*1364,666667*
exp 6	*1872*	*2149,333333*	*1580,666667*	*1292,333333*
exp 7	*1240,333333*	*1081,666667*	*924,6666667*	*873,3333333*
exp 8	*1269,333333*	*1142*	*917,3333333*	*1051,333333*
exp 9	*972,3333333*	*968,6666667*	*932,3333333*	*898*
exp 10	*1583,333333*	*1641,333333*	*1707*	*1228,333333*
exp 11	*1622,666667*	*1550*	*1285,333333*	*1291*
exp 12	*766,3333333*	*1073,666667*	*783*	*761*

exp 13	*915,6666667*	*772,3333333*	*684*	*738,6666667*
exp 14	*1372,666667*	*1285*	*1095*	*1144,666667*
exp 15	*853*	*911,6666667*	*998,6666667*	*837,3333333*
exp 16	*926,3333333*	*1077,333333*	*1107,333333*	*926,3333333*
	x̄=*1242,729167*	x̄=*1189,729167*	x̄=*1027,104167*	x̄=*1028,645833*
	SD=*319,775342*	SD=*348,168227*	SD=*281,024414*	SD=*195,150595*

Table 3. Averages of the responses of each participant to the frequent 1 group, the Frequent 2 group, the Infrequent group 1, and the Infrequent group 2.

Again, a Shapiro-Wilk test was used to check the normality of the four groups. With the significance level set at α=0.05, the Frequent 1 group did not show normal distribution with a p-value: 0.7092; the Frequent 2 group showed normal distribution with a p-value: 0.0301; the Infrequent group 1 showed normal distribution with a p-value: 0.0126; the Infrequent group 2 showed normal distribution with a p-value: 0.0422

Concerning the comparison of frequent groups, since the Frequent 1 group does not have a normal distribution, the comparison between Frequent 1 and Frequent 2 had to be non-parametric. Since the results are paired because the same participant produced them, the statistical test that was used is the Wilcoxon Signed-Rank Test. The null hypothesis H0: frequent 1 group =Frequent 2 group and an alternative hypothesis H1: Frequent 1 group ≠ Frequent 2 group. With a significance level set at α=0.05, the alternative hypothesis is rejected at a p-value: 0.35238.

In regards to the comparison of the infrequent groups, both groups show a normal distribution and are also paired by the participants. In this way, a Paired T Student test was carried out. With a null hypothesis H0: Infrequent 1 group = Infrequent 2 group and an alternative hypothesis H1: Infrequent 1 group ≠ Infrequent 2 group. With a significance level set at α=0.05, the alternative hypothesis is rejected at a p-value=0.9785.

Discussion

With these results, we can start describing what is their contribution to the above-mentioned models. The results do not validate the initial suppositions that I presented about prepositions in the mind. Nonetheless, we are going to decipher these results one by one answering the research questions.

Regarding the first research question, the data did not show any statistical significance in the comparison of the control and experimental group. However, although the p-value is high (0.6106), I believe that we should not directly reject the hypothesis. The reason for this is that the difference between the means of each group is large (control: x̄=1095, experimental: x̄=1122). This can be due to the fact that the number of participants was small. The statistical test had to be run with the mean of all the responses of each participant, not with the raw responses of each participant. Thus having small groups of participants produced that the results could be more dependant on chance. An indicative test with all the raw results of all participants was carried out. This cannot be considered statistically valid because the responses are considered independent, but as we said before, they have to be considered dependent on the same participants. However, this test has been used as an indication to check if the p-value moves towards signification when the data has the same mean but with a larger number of responses. Again, a Mann Whitney U test was used to compare both control and experimental groups with 192 responses in each group. In this case, the p-value=0.186, still not significant, but much closer to significance than the other test. As I have already said, the test with 384 responses is not to be considered seriously, but it shows how the p-value reduces having more responses, even if the mean is the same. In this way, the experiment should be carried out by many more participants to ultimately confirm or refute my results.

Nevertheless, if this result is replicated and again shows that the difference is not significant, we would have to conclude that prepositions are not stored in the same place as content words, or, at least, that the prepositions do not take part in Roelofs' network for spreading activation. This would mean that prepositions and possibly other types of functional words are stored and selected by other operators as Lapointe & Dell (1989) selected. In that way, a new line of research should be implemented to decipher the nature of the processing of these words, but separating from the picture-word interference paradigm because it would only explain their nature in Roelofs and Levelt's model.

With reference to the second research question, the high-frequency responses and low-frequency responses did show a great significant difference (p-value=0.00078). This result can have two different interpretations, depending if we reject the previous hypothesis or not.

If we do not reject the previous hypothesis and consider that prepositions are included in the Roelofs' network, the difference between the frequent and infrequent groups could be explained by the strengths of the links. As Collins & Loftus (1975) showed, the more frequent a connection between nodes is, the stronger it becomes, and then it shares more activation than weaker links. This is, in the case of "house of dolls", the connections between "house" - "of" - "dolls" is much stronger than the connection of "house" - "of" - "people", for example, because the former has been activated in that way many more times than the infrequent "house of people".

If, on the other hand, we reject the previous hypothesis, as the results show, and we consider that these prepositions are not included in the network, we would have to find the explanation somewhere else. As I stated before, the frequent words that were included in the experiment are NpNs that not only are frequent, but that also are used to denominate entities that could not be denominated in any other way. This is, the construction NpN was required to refer to those entities (e.g. in Spanish "casa de muñecas, silla de ruedas, gafas de sol,..."). In this way, this type of NpNs could be understood as individual entities and have individual nodes in the conceptual stratum and lemma stratum, even if they consist of three words in overt speech.

Then, the longer response time in the frequent group would be explained because the single node of the NpN serves as mediation of the nodes of the N1 and the N2. To exemplify this, I will use "house of dolls". When presented with the first prime "house of dolls", the participants' network activates the lemma node of "house of dolls", this shares its activation to the lemma nodes of "house" and "dolls" because they are related; then, when presented with the second prime "dolls", it activates the lemma node of "dolls" and shares its activation towards the lemma node "house of dolls" because of the previous prime, and "house of dolls" itself shares a portion of its activation to the lemma node “house” because they are related, producing a naming latency with the related word “castle”. If this happens to be true, it would be possible to expect other types of multiple words constructions to exist as individual concept nodes and individual lemma nodes.

In terms of the third research question, neither of the comparisons (order in the frequent group and order in the infrequent group) presented statistically different results. In the case of the infrequent NpNs, the comparison showed a p-value=0.9785, stating that there is very little possibility that any difference exists. This was expected by the external comparison of the means in both groups (Infrequent 1: $\bar{x}$=1027,104; Infrequent 2: $\bar{x}$=1028,645). This means that no matter if the N1 or the N2 were presented as primes, the results were unchanged. If the prepositions do not take part in the network, this result would be explained by the fact that the preposition is not present in the network and does not share activation, and then, it does not matter what noun is used as prime because any of them will share the activation. If, on the other hand, the prepositions take place in the network, these results would prove that there is no linear activation of the infrequent NpNs, this is, the strength of links is the same in N1→p→N2 than in N2→p→N1. This would make sense because this type of NpNs was novel to participants and consequently they have not previously activated N1→p→N2 nor N2→p→N1.

In relation to the order in the Frequent group, the difference is also not statistically significant. This was not expected because, since they are frequent NpNs, the strength of the links in the original direction (N1→p→N2) should be stronger than the opposite direction (N2→p→N1) because it has been more frequently activated in that way. This result, however, correlates with the idea of single nodes for frequent NpNs, because there wouldn’t be any linear order, only three related nodes: the node for the N1, the node for the NpN, and the node for the N2. With this view, the relation of the node of the N1 to the node of the NpN would be the same as the relation between the node of the N2 and the node of the NpN. Let us exemplify this, when the lemma node for the NpN “house of dolls” is activated, it shares activation to the nodes of “house” and “dolls”; then, when the word distractor is presented, it is not important which one is presented because the relation of the N1 and N2 with the NpN would be the same (“house”→“house of dolls”→“dolls” would be the same as “dolls”→“house of dolls”→“dolls”).

In light of all these findings, the results tend to indicate that the preposition “of” did not take the expected place in Roelofs’ network. Although there is a small difference in the response time means of the experimental and control group, it could be explained by chance. What has been most interesting is the fact that the frequent NpNs produced significantly more naming latencies than the infrequent NpNs. If prepositions do not take place in Roelofs’ network, this result should not be so statistically different. Nonetheless, a possible solution has been found in the hypothesis that these types of NpNs exist as a single unit in its lemma form. In this way, the preposition “of” would only be included semantically in the network when it forms a necessary part of an NpN that stands for a single entity. Moreover, the third analysis in this paper, the analysis of the linear order of processing, supports this hypothesis. If no difference is found between the order of N1→p→N2 and N2→p→N1 in the frequent group, where NpNs are frequently processed with the former order, the only explanation would be understanding those elements as a single node in the network.

Conclusion

This paper has aimed at exploring the workings of prepositions in language production using a novel experimental method. With this experiment, I intended to locate the activation of prepositions in the lexical selection as a part of Roelofs's model for spreading activation. Unfortunately, the main research question has not found an answer. The results have proved not to be significant, and thus refuted the first hypothesis that stated that prepositions are included in the network. Still, the absence of these should now be verified with more participants taking part in this experiment, a replication of the experiment with other units of analysis, or with a similar experiment of picture-word interference paradigm. With similar results or different results, a computational model that accounts for these results should be carried out as well. This would resolve the question of whether my results are valid or they need more participants. In this way, the doubts that I have explained above could be answered.

Meanwhile, the characteristics of NpNs have also been studied in the experiment and they have proved to be more interesting than expected. The significant difference between frequent and infrequent groups showed that the NpNs that are used to define single entities have a high probability to present an individual node in the network, at least in the lemma stratum. I say that the single node is at least present in the lemma stratum because it is required to understand the results; however, the existence of a single node for NpNs in the conceptual node would be logical, but not necessary. Thus, we cannot assume its presence as a single concept node because there could be other ways in which the single lemma node of the NpN is activated without its single concept node. For example, a similar activation of the concept nodes for the N1 and N2 could send activation to the NpN lemma node.

Yet, it would be more plausible to find a single concept node for the NpN because they denote individual entities. Although they are clearly influenced by two different concepts, the N1 and N2, they form a new meaning that is more than just the sum of the N1 and N2. This was already considered by Thagard (1996). He explained that a way of acquiring new concepts could be the combination of two different concepts. This explanation would satisfy the existence of NpNs as single nodes, but with a preposition combining the two nouns with a logical relationship.

Again, these results should be verified by obtaining similar results in the same type of experiment or with a simplified version of this experiment. Since these NpNs represent single entities, a normal picture-word interference experiment could be applied. In these experiments, the entities that the NpN represent could be included as pictures with a single distractor word. Also, these results should be confirmed with a computational model that accounts for the real response time by the participants. This could be done easily by including the single nodes of the NpNs in the network with the characteristics that I have explained in the discussion and run the model with the same stimuli that the participants received. In this manner, if the results of that computational model are similar to the ones I have obtained, we could conclude that the frequent NpN form a single entity in our mind in terms of processing.

When it comes to the third research question, it helped elucidate the nature of the previous results. The presentation of the N1 or the N2 as distractor words to participants was indifferent. If prepositions take place in the model as expected, there should not be a difference in the infrequent group, but there should be a difference in the frequent one. Still, none of the groups showed any significant difference, contrary to what was expected. Thus, they support the idea that frequent NpNs are independent nodes in the network, and the order of processing is not linear, the relation between the N1 node and the NpN node is the same as the relation between the N2 node and the NpN node.

Besides, apart from the results, the paper has been used to explore some of the most important models in language production. The theoretical background has served to give a general understanding and an overview of several models of language production such as general cognitive processes of intention behind speaking, lexical selection, and grammatical encoding. These models have not been explained in detail but have included all the necessary elements to understand them.

In summary, the results were not what was expected, but they have shown interesting results in another direction. Overall, they have contributed to a better understanding of language production.

References

Baddeley, A. (1992). Working memory. Science, 255(5044), 556-559.

Berg, T. (1987). The case against accommodation: Evidence from German speech error data. Journal of Memory and Language, 26, 277-299.

Bobrow, D. G., & Winograd, T. (1977). An overview of KRL, a knowledge representation language. Cognitive science, 1(1), 3-46.

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. Journal of Memory and Language, 31, 99-127.

Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. Psychological Review, 99, 150-171.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. Psychological Review, 89, 1-47.

Bock, J. K. & Levelt, W. J. M. (1994) Language production: Grammatical encoding. In: Handbook of psycholinguistics, ed. M. A. Gernsbacher, pp. 945–84. Elsevier/Academic Press.

Clark, E. (1997) Conceptual perspective and lexical choice in acquisition. Cognition 64:1–37.

Collins, A.M., & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82, 407-428.

Cutler, A. (1988). The perfect speech error. In L. M. Hyman & C. S. Li (Eds.), Language, speech and mind: Studies in honour of Victoria A. Fromkin (pp. 209-223). London: Routledge

Dell, G.S., & Reich, P.A. (1981). Stages in sentence production: An analysis of speech error data.

Dell, G. S., Juliano, C, & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. Cognitive Science, 17, 149-195.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. Psychological review, 93(3), 283.

Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. Learning and Motivation, 49, 209–246.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), The psychology of learning and motivation (pp. 133-177). New York: Academic Press.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), Language production (Vol. 1, pp. 177-220). London: Academic Press

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), Normality and pathology in cognitive functions (pp. 19-76). London: Academic Press

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture-word interference. Journal of Experimental Psychology: Human Perception and Performance, 10(5), 640-654.

Harley, T. A. (2005). The Psychology of Language from Data to Theory (2nd Ed.). Scotland: Tylor and Francis Group.

Jongman, S. R., Roelofs, A., & Meyer, A. S. (2015). Sustained attention in language production: An individual differences investigation. Quarterly Journal of Experimental Psychology, 68(4), 710–730.

Journal of Verbal Learning and Verbal Behavior, 20, 611-629.

Keil, F. C. (1979). Semantic and Conceptual Development: An Ontological Perspective. Cambridge, Mass: Harvard Univ. Press

Kemmerer, D. (2015). Are the motor features of verb meanings represented in the precentral motor cortices? Yes, but within the context of a flexible, multilevel architecture for conceptual knowledge. Psychonomic Bulletin & Review, 22(4), 1068-1075.

Lapointe, S. G., & Dell, G. S. (1989). A synthesis of some recent work in sentence production. In G. N. Carlson & M. K. Tanenhaus (Eds.), Linguistic structure in language processing (pp. 107-156). Dordrecht: Kluwer.

Lapointe, S. G. (1985). A theory of verb form use in the speech of agrammatic aphasics. Brain and Language, 24, 100-155.

Laurence, S. & Margolis, E. (1999). Concepts and Cognitive Science. In Eric Margolis & Stephen Laurence (eds.),Concepts: Core Readings. MIT Press. pp. 3-81.

Levelt, W. J. M., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. Behavioral and Brain Sciences, 22(1), 1-38.

Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. Cognitive Psychology, 14, 78-106.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T. & Havinga, J. (1991) Normal and deviant lexical processing: Reply to Dell and O'Seaghdha. Psychological Review 98:615–18.

Levelt, W. J. M. (1989) Speaking: From Intention to Articulation, MIT Press, Cambridge, Massachusetts

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.

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture–word interference experiments. Journal of memory and Language, 35(4), 477-496.

Montero-Melis, G., van Paridon, J., Ostarek, M., & Bylund, E. (2019). Does the motor system functionally contribute to keeping words in working memory? A pre-registered replication of Shebani and Pulvermüller (2013, Cortex).

Moreno-Martínez, F. J., & Montoro, P. R. (2012). An Ecological Alternative to Snodgrass & Vanderwart: 360 High Quality Colour Images with Norms for Seven Psycholinguistic Variables. PLoS ONE, 7(5), e37527. doi:10.1371/journal.pone.0037527

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. Psychological Bulletin, 134(3), 427.

Quillian, M. R. (1969). The Teachable Language Comprehender: A simulation program and theory of language. Communications of the ACM, ,12, 459-476.

Roelofs, A., & Ferreira, V. S. (2019). The architecture of speaking. Human language: From genes and brains to behavior, 35-50.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. Cognition 42:107–42

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. Cognition, 64(3), 249-284.

Saint-Dizier, P. (ed.) (2006). Syntax and semantics of prepositions. Dordrecht: Springer

Shao Z., Roelofs A., & Meyer A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control, The Quarterly Journal of Experimental Psychology, 65:10, 1927-1944

Snodgrass, L., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. Journal of Experimental Psychology: Human Learning and Memory, 6, 174–215.

Stemberger, J. P. (1982). Syntactic errors in speech. Journal of Psycholinguistic Research, 11, 313-345.

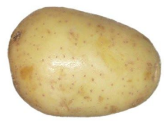
Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), Progress in the psychology of language (pp. 143-186). London: Erlbaum.

Thagard, P. (1996). Mind: Introduction to cognitive science (Vol. 4, pp. 3-22). Cambridge, MA: MIT press.

Tree, J. E. F., & Meijer, P. J. (1999). Building syntactic structure in speaking. Journal of Psycholinguistic Research, 28(1), 71-90.

Appendix A

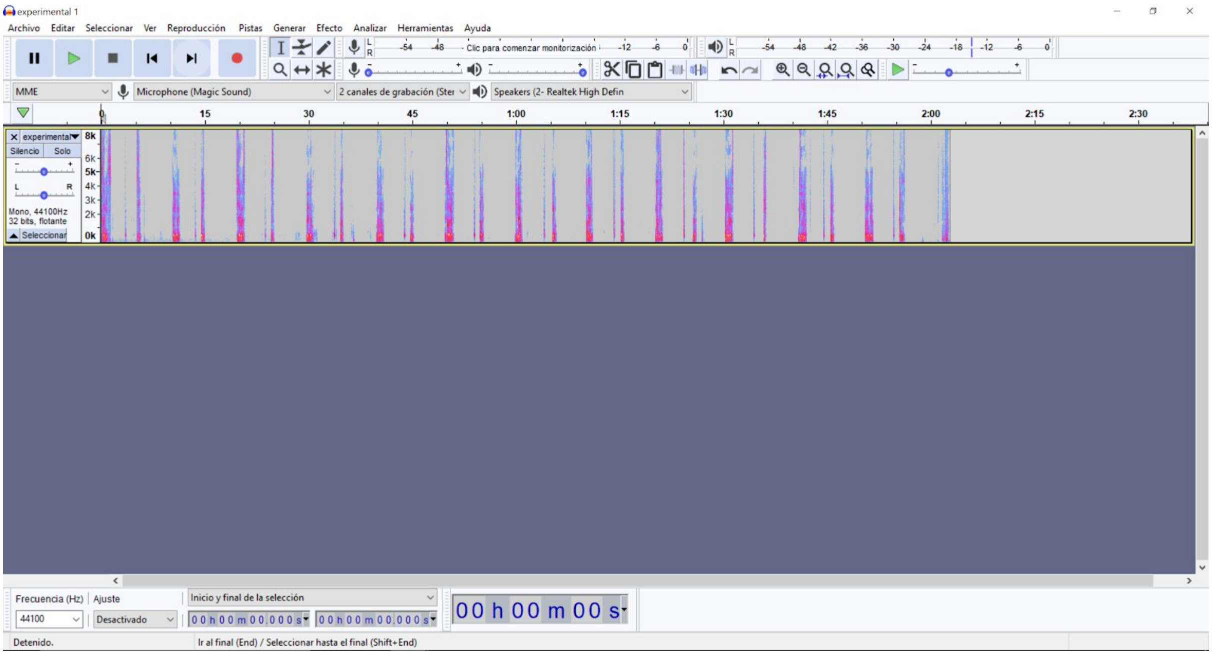
Table with the order of trials, the Noun-preposition-Noun phrase that was used to prime participants (the word in bold is the word related to the picture), the frequency of the NpN in the CREA corpus, the distractor word that appeared in the screen, and the target image that participants had to produce.

Order	***Prime NpN***	***Frequency NpN***	***Distractor word***	***Image***
1st	***Casa** de muñecas*	*41/26*	*muñecas*	
2nd	***Cable** de internet*	-	*internet*	
3rd	***silla** de ruedas*	*674/326*	*ruedas*	
4th	*puré de **puerro***	-	*puré*	
5th	*Vapor de **agua***	*415/125*	*vapor*	

6th	*ensalada de* ***maíz***	-	*ensalada*	
7th	*Diente de* ***ajo***	*407/43*	*diente*	
8th	*Pata de* ***camello***	-	*pata*	
9th	*Gafas de* ***sol***	*221/141*	*gafas*	
10th	***camión*** *de ruedas*	-	*ruedas*	
11th	***Vestido*** *de novia*	*94/61*	*novia*	
12th	***Pantalón*** *de tela*	-	*tela*	

Appendix B

Example of recording for the 12 trials of a participant, the view is spectrogram, and each peak is either the NpN produced by the participant, a musical note A, or the target response of the participant



Appendix C

Table with the number of participant, age of the participants, response time of each participant in each trial, and mean of each participant in their 12 trials.

participant	Age	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10	trial 11	trial 12	means
con 1	29	1226	1820	1666	1644	951	1122	1386	1651	991	1040	1558	1091	1345,5
con 2	57	960	770	824	797	895	849	764	1467	1080	886	728	776	899,66666
con 3	66	1406	1095	855	1251	1191	1341	1334	1328	1982	1386	1374	1159	1308,5
con 4	61	998	1586	713	1081	836	876	1849	773	2068	892	942	843	1121,4166
con 5	18	1844	778	767	693	856	1046	882	662	725	703	1024	651	885,91666
con 6	46	1291	962	1215	681	877	1036	837	759	1288	985	1013	713	971,41666
con7	18	795	871	715	748	809	1605	1220	729	1136	674	1753	782	986,41666
con 8	72	2934	2324	1746	1528	1330	2656	2152	1971	1657	1792	2310	1256	1971,333
con 9	52	1644	1255	1094	1073	960	1338	1152	952	1897	1659	1136	1482	1303,5
con 10	68	1541	1410	921	1043	1084	1072	1536	1055	2418	886	1058	1024	1254
con 11	24	1151	831	1122	712	749	715	849	927	1005	790	2002	678	960,91666
con 12	25	772	645	672	578	622	722	687	610	1016	722	1289	766	758,41666
con 13	20	789	676	595	722	645	770	739	742	814	673	1162	797	760,33333
con 14	56	1576	1533	1305	815	881	1249	965	1044	1213	976	968	822	1112,25
con 15	27	1190	847	760	891	761	1007	793	668	1057	772	1030	949	893,75
con 16	26	1626	1174	940	826	870	822	882	882	1010	863	1066	922	990,25
exp 1	62	1674	905	884	817	870	781	1204	1024	1314	1463	1008	882	1068,8333
exp 2	26	795	745	715	710	1395	1022	1015	844	1133	1024	960	938	941,3333333
exp 3	68	1782	992	698	1119	998	1154	1777	796	2094	1309	1228	917	1238,666667
exp 4	59	907	807	807	760	984	856	870	897	2389	1476	1024	696	1039,416667
exp 5	48	1790	1322	766	726	785	837	1384	912	1257	840	1110	1932	1138,416667
exp 6	46	2528	1459	2378	1277	1136	1185	2502	2280	1978	1320	1542	1098	1723,583333
exp 7	38	959	834	1140	805	1041	837	936	1132	1744	859	1146	927	1030
exp 8	67	1685	1114	812	917	991	859	1067	976	1750	979	929	1061	1095
exp 9	49	1103	1246	907	856	861	1091	913	850	1143	724	896	724	942,8333333
exp 10	69	1709	1367	1286	1431	1321	1251	1865	2439	1564	1043	1929	1275	1540
exp 11	31	1489	1518	1333	1596	1382	1194	1191	1066	2295	1072	1828	1283	1437,25
exp 12	24	1123	847	849	776	621	890	805	683	873	689	1249	747	846
exp 13	48	968	774	573	604	864	782	776	666	1107	728	776	714	777,6666667
exp 14	26	1297	1210	1109	806	960	1168	1513	1311	1645	1028	1449	1196	1224,333333
exp 15	32	1124	945	708	928	686	1322	934	746	939	746	903	821	900,1666667
exp 16	35	1207	786	827	969	889	1355	821	998	1069	985	1198	1008	1009,333333