Similarity-based competition in relative clause production and comprehension

Humphreys, G. F.

University of Manchester

Mirković, J.

York St John University and University of York

Gennari, S. P.

University of York

Corresponding author

Silvia P Gennari

Silvia.gennari@york.ac.uk

Department of Psychology

University of York

YO10 5DD

York, UK

Abstract

This work investigates the role of semantic similarity in sentence production and comprehension. Previous research suggests that animacy and conceptual similarity of the noun concepts within complex descriptive phrases modulate structural preferences in production, and processing cost in comprehension. For example, animate-head phrases such as *the girl that the boy is pulling* are rare in production and more difficult to understand in comprehension. In contrast, phrases with passive clauses such as *the girl being pulled by the boy* are commonly produced and more easily understood, as are inanimate-head structures such as *the truck the boy is pulling*. In three picture-based studies, we examined the mechanisms underlying semantic similarity effects in producing and comprehending these phrases. Study 1 investigated structural preferences in production, whereas Study 2 investigated processing cost in comprehension. Study 3 used eye-tracking to examine the time-course of production processes. The results showed that semantic similarity elicited competition during phrase planning, influenced the choice of syntactic structure in production, and engendered comprehension difficulty in animate-head active configurations. Structural preferences, fixation probabilities reflecting production planning processes and comprehension cost significantly correlated with measures of conceptual similarity across the three studies. We argue that similarity-based competition modulates sentence production and comprehension processes when verbs are planned or interpreted, i.e., when event-based semantic or syntactic roles are determined. In addition to task-specific processes, we suggest that a similar and shared semantic competition mechanism underlies both production and comprehension, a view consistent with existing evidence for common brain regions recruited in both tasks.

## 1. Introduction

Traditionally, psycholinguistic research has studied production and comprehension as separate processing systems (Bock & Levelt, 1994; MacDonald, Pearlmutter, & Seidenberg, 1994). This appeared necessary because production and comprehension do involve different processes such as word retrieval vs. recognition, or motor planning vs. sensory analysis. More recently, however, researchers have started to investigate relationships between these two processing systems and what they have in common (Bock, Dell, Chang, & Onishi, 2007; Chang, Dell, & Bock, 2006; Gennari & MacDonald, 2009; Humphreys & Gennari, 2014; MacDonald, 2013; Menenti, Gierhan, Segaert, & Hagoort, 2011; Pickering & Garrod, 2007). Chang et al. (2006), for example, suggest that the learner's experience shapes production processes, and that learning and production share the same computational architecture and the abstract structural knowledge underlying these processes. Gennari and MacDonald (2009) further suggest that the mapping of semantic roles into syntactic arguments may be shared across production and comprehension, in part because both abilities emerge from experience-based learning, which extracts common patterns of mapping words to meanings. More generally, MacDonald (2013) argues that production mechanisms play a significant role in shaping comprehension processes because they determine the input that learners and comprehenders are exposed to. Taken together, these accounts converge in suggesting that production and comprehension share a common knowledge base, which contains lexical knowledge and abstract mapping patterns extracted from linguistic experience.

Beyond shared knowledge, however, it remains an open issue whether sentence production and comprehension also share specific processing mechanisms. Previous behavioral evidence has shown strikingly similar processes in both tasks, and thus suggests potential candidate mechanisms. For example, in sentence comprehension research, multiple studies have shown that relative clauses such as those in (2) are difficult to comprehend and that this comprehension difficulty is modulated by animacy and conceptual similarity (Gennari & MacDonald, 2008, 2009; Mak, Vonk, & Schriefers, 2006; Traxler, Morris, & Seely, 2002). Object relative clauses with two animate nouns as in (2) are harder to understand than those with nouns of different animacy as in (4), and increased similarity between the two animate nouns leads to increased difficulty (Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, Johnson, & Lee, 2006). This comprehension difficulty has been attributed to different mechanisms. Some authors have argued for similarity-based interference in working memory as two similar nouns (e.g., *man* and *woman* in (2)) must be retrieved and integrated with the upcoming verb (Gordon et al, 2001). Others have argued for competition between incoming and expected alternative structures and interpretations as the relative clause unfolds (Gennari & MacDonald, 2008, 2009). These two views are not mutually exclusive, because different types of linguistic information may contribute to competition between interpretations and/or retrieval interference in working memory (Staub, 2010).

(1) Animate Head, Passive: *The man who's being punched by the woman*

(2) Animate Head, Active*: The man that the woman is punching*

(3) Inanimate Head, Passive: *The sand bag that's being punched by the woman*

(4) Inanimate Head, Active: *The sand bag that the woman is punching*

Similarly, in production research, studies have shown that speakers and writers select relative clause structures according to animacy and conceptual similarity. Producers very rarely use structures such as (2), whereas they are more likely to use the ones in (1) and (4) (Gennari & MacDonald, 2009; Gennari, Mirkovic, & MacDonald, 2012; Roland, Dick, & Elman, 2007). For example, a sentence elicitation task using pictures such as that in the left panel of Figure 1 indicated that in response to a question, e.g. *Who is bald?,* speakers overwhelmingly prefer structure (1) rather than (2). In contrast, in response to *What is orange?*, speakers are equally likely to use (3) or (4) (Gennari et al, 2012). These findings indicate that animacy influences production choices. Moreover, in animate-head cases such as (1) the conceptual similarity between the nouns of the relative clause positively correlated with the rate of agent omissions: for the case of man and woman in (1), the preference for agent omission is relatively large (e.g., *the man being punched*), whereas for cases where the similarity is lower as in *the baby being held by the woman* (Figure 1, right panel), this preference is reduced (Gennari et al, 2012). This graded similarity effect clearly indicated that specific semantic features of the nouns played a role, rather than only categorical animacy. The authors argued that conceptual similarity may cause interference between alternative concepts or words to be produced. In planning relative clauses with two animate nouns, where the second noun could potentially follow the first noun, sentence planning may involve some temporal overlap in the preparation of the two nouns, and hence there is potential for interference or competition between their meanings, as in comprehension of object relative clauses. This interference or competition may result in inhibition or reduced activation of the agent concept following lexical selection of the head noun *man*, leading to the use of a passive construction in which the agent noun is mentioned last or omitted altogether.

Figure 1: Example images used in the elicitation of relative clauses in Gennari et al, 2012 and in Studies 1 and 3. The left panel displays a case of high similarity between the animate entities participating in an event, and the right panel a case of low similarity.

More generally, outside of the domain of relative clauses, much research has also demonstrated the role of similarity-based interference in many language behaviors, including production, comprehension and verbal working memory (Baddeley & Hitch, 1974; Baddeley, 1986; Freedman, Martin, & Biegler, 2004). For example, production studies investigating factors that modulate choices of referential expressions (e.g., pronouns vs. definite descriptions) suggest that the conceptual similarity between the characters to be described influences the choice of referential expressions (Arnold & Griffin, 2007; Fukumura, van Gompel, Harley, & Pickering, 2011; Slevc, 2011; Smith & Wheeldon, 2004). These findings are consistent with Smith and Wheeldon's proposal that some temporal overlap exists, if only at a conceptual level, in the planning of two nouns to be produced in sequence, thus leading to semantic interference (see also picture-word interference effects between distractors and targets, e.g. Costa, Alario, & Caramazza, 2005; Glaser & Düngelhoff, 1984; Glaser & Glaser, 1989; Schriefers, Meyer, & Levelt, 1990). Further, models of sentence comprehension inspired by working memory research suggest that similarity-based interference pervades many aspects of sentence comprehension and leads to a greater processing cost when the target to be retrieved from memory syntactically or semantically overlaps with distractors available in working memory (Gordon et al., 2001; Gordon, Hendrick, & Johnson, 2004; Lewis & Vasishth, 2005; Van Dyke, Johns, & Kukona, 2014; Van Dyke & McElree, 2011; Van Dyke, 2007).

The evidence reviewed above strongly suggests that similarity-based processes take place in both production and comprehension in a variety of circumstances. In all cases, these processes appear to involve a competitive phase in which alternative candidates are entertained to varying degrees, leading to a cost in selecting or maintaining the targeted information, and reduced accessibility or inhibition of the non-selected alternative. We will use the term *interference* or *competition* interchangeably to refer to these processes, as both terms are used to refer to empirically equivalent mechanisms (Badre & Wagner, 2007; Conway & Engle, 1994; Friedman & Miyake, 2004; Levy & Anderson, 2002). Relative clauses provide a useful ground to investigate these mechanisms further, because both production and comprehension of relative clauses appear to involve similarity-based interference due to maintenance and retrieval of highly similar noun concepts. Moreover, relative clauses allow using identical stimuli across both tasks in contexts where relative clauses are felicitous or expected. Note that in the absence of context, as in many reading comprehension studies, an isolated sentence such *The sandbag that the woman punched was orange* is pragmatically odd because relative clauses typically require a context in which there is more than one sandbag. In picture contexts such as that of Figure 1, the presence of an alternative or contrasting sandbag makes it more relevant for the comprehender and the speaker to expect or to produce a relative clause specifying which sandbag is being referred to. Thus in what follows, we investigate the production and comprehension of relative clauses in contexts where they are felicitous for both the speaker and the comprehender.

Our goal in this work is twofold. First, we investigate the relationship between production and comprehension by examining in more detail the role of similarity-based competition in producing and comprehending different types of relative clause structures. To this end, we report a production and a comprehension study using the same picture materials, which attempt to establish correspondences across the two tasks. Second, we aim to investigate in more detail the time-course of production processes operating in relative clauses, which unlike those in comprehension have not been extensively investigated, and we examine the role of conceptual similarity in relative clause production. Previous production studies have only investigated structure preferences and initiation times (Gennari et al., 2012; Montag & MacDonald, 2014), but it is currently unclear how and when similarity-based mechanisms take place. Together, these studies have the potential to highlight common as well as distinct mechanisms at play in production and comprehension.

Specifically, we test the hypothesis that there may be points in the structure at which semantic similarity plays a role both in production and comprehension, suggesting similar mechanisms operating on the same kinds of representations. We reasoned that both producing and understanding sentences require representing the event being communicated, in particular, the relationship between the nouns (who is doing what to whom). Event-based information is typically carried by verbs, which also act as the syntactic and semantic link between the nouns of the sentence and the entities of the events. Indeed, in comprehension of object relative clauses, such link is typically assumed to occur at verbs, because the two relative clause nouns preceding the verb receive a semantic and syntactic role interpretation at this point (Gibson, 1998). Competition may thus arise when the two nouns are good candidates for the same syntactic and/or semantic role. In production, however, it is unclear whether some sort of semantic integration or structural linkage occurs when a verb is planned. This is in part because subject-verb-object structures have been more thoroughly investigated, the simple linearity of which, as in the comprehension of simple sentences, does not provide the opportunity to examine the consequences of entertaining more than one entity at a time. The investigation of complex structures in production therefore may be revealing in this respect. Indeed, Gennari et al. (2012) suggested that in Spanish competition between similar entities appeared to determine the choice of syntactic function for the head noun (object or subject) in relative clause production —less similar entities are mapped into active impersonal structures whereas more similar entities are mapped into passives as in English. This preliminary evidence was interpreted to suggest a role for similarity-based competition at the point in which syntactic functions are established in the formulation process, which have direct repercussions on the planning of verb morphology (e.g., *is being punched* vs. *is punching*) (Gennari et al, 2012). It is therefore possible that competitive processes may occur in both production and comprehension at similar points in the structure, providing supporting evidence for common competitive mechanisms.

## 2. Study 1: The role of similarity on structure choice in production

Production involves mapping conceptual representations onto a linear structure. The form of this structure is known to be largely dependent on *accessibility*, the ease with which words or concepts can be retrieved. Uttering easier-to-retrieve words early provides additional time to plan less accessible components, thus maximizing fluency (De Smedt & Kempen, 1987; Kempen & Hoenkamp, 1987; Levelt, 1989). In particular, animate nouns, which are inherently more salient and accessible than inanimate nouns, tend to precede other words in a structure and thus constrain the form of the resulting sentence, e.g., whether it is an active vs. passive sentence, or a double object vs. prepositional dative verb phrase (Bock & Warren, 1985; Bock, 1987; McDonald, Bock, & Kelly, 1993). For example, the passive *The boy was hit by the truck* is preferred over the active *The truck hit the boy*, because the animate noun is easier to retrieve from memory and thus pronounced first. This is consistent with the general tendency in English to locate animate concepts in early sentence subject positions, even when they are not agents (Bock, Loebell, & Morey, 1992; Clark, 1965).

The accessibility of referential expressions is known to be modulated by context. For example, many picture-word interference studies have shown that picture targets that are semantically similar to distractor words are named more slowly than when unrelated distractors are present (Costa et al., 2005; Vigliocco, Vinson, Damian, & Levelt, 2002). Other studies investigating pronoun use—an index of the high accessibility of an entity in the speaker's mind—have shown that when an animate target is being talked about, the mere presence of an animate—but not inanimate—competitor in the context reduces the likelihood that speakers would use pronouns (Arnold & Griffin, 2007; Arnold, 2010; Fukumura et al., 2011; Fukumura & van Gompel, 2010). This type of result suggests that the accessibility of the entity to be named is reduced in the presence of other potential candidates, i.e., competitor entities interfere with the lexical retrieval of the target reference due to their semantic similarity. Interestingly, in a picture description task, the actions that characters perform also contribute semantic features leading to interference: if the target and the competitor are performing similar actions (e.g., both are sitting on a horse), more descriptors and fewer pronouns are used, compared to when they are performing different actions. This effect is larger if the action to be described in the utterance is related to the action being performed (e.g., getting off the horse) (Fukumura et al., 2011).

In the context of these findings, relative clauses are interesting to investigate because they identify a specific entity being talked about relative to the other entities present in the context, and therefore all entities are susceptible to the sort of interference effects described above. Gennari et al. 2012 (henceforth GMM) have indeed suggested that animate heads are typically followed by a passive structure as in (1) rather than active structures like (2), because the retrieval and selection of the head noun involves competition with other alternatives (other people in the picture), and in particular those who are taking part in the same event to be described (e.g., the man and woman involved in the punching event). GMM suggest that the selection of *man* leads to the inhibition of *woman,* making it less accessible to be named right after the head. This results in a passive structure, rather than an object relative clause. Moreover, the conceptual similarity between the two animate nouns of a relative clause (the agent and the patient of the punching event or the holding event in Fig. 1) was found to correlate with the tendency to omit the agent of the event (*by*-phrase omission), with more similar entities resulting in more by-phrase omissions, suggesting stronger inhibition for more similar entities.

Interestingly, however, similarity did not correlate with the choice of active vs. passive relative structures (GMM). Indeed, passive structures are overwhelmingly preferred for animate-heads in English and this strong language-wide structural bias leaves little room for exploring the influence of specific semantic features on structure choices besides animacy. Thus in order to examine the extent to which the similarity-based account applies to structural choices, we aimed to decrease the overwhelming passive bias by manipulating the experimental context and thus increase the chances to detect the influence of factors other than animacy in structure choices. Specifically, we increased the availability of both active phrase structures and agent mentions within the experimental context. We reasoned that increasing the availability of these relative clause properties in the experimental context should increase the likelihood of using an active structure, as in syntactic priming studies (Bock, 1986; Bock et al., 2007) and in particular, it should increase the likelihood of producing an animate agent right after the animate head in an object relative structure. Critically, if noun similarity in relative clause production shapes phrase structure choices, the contextual facilitation should be modulated by conceptual similarity. Specifically, highly similar animate nouns are expected to engender similarity-based interference, and thus structural facilitation from previous trials should be more likely to occur with less similar animate nouns. In other words, previous contextual experience with a structure only facilitates a given utterance if it maximizes production efficiency, rather than taxing it.

The goal of the present study was therefore to evaluate these predictions by examining relative clause production in contexts in which agent mentions and alternative structures are made relatively more accessible compared to previous studies. In doing so, we aim to provide evidence that similarity-based competition can influence the choice of passive vs. active clauses, i.e., syntactic function assignment—whether the head noun is the subject or the object of the verb as in (1) and (2)—and consequently it can also (directly or indirectly) influence the corresponding choice of passive or active verb morphology. Because the verb morphology must be planned near or together with the verb, this would provide initial support for an influence of similarity-based competition on verb planning or encoding.

In the present study, we used a paradigm in which blocks of production trials were intermixed with blocks of comprehension trials. Both production and comprehension tasks were picture-based. The production task instructed participants to describe an entity highlighted on a picture like that of Figure 2. The comprehension task instructed participants to indicate whether a phrase appearing below the picture correctly described a character highlighted in it (see Figure 2). Unlike other priming paradigms, we did not seek to prime the production of active structures with a specific preceding trial using the same structure. Instead, we aimed to increase the overall availability of all alternative structures in the comprehension blocks. Thus, the comprehension trials exemplified all possible relative clause structures—animate- and inanimate-head actives and passives—and always mentioned the agents (there were no agent omissions). We assumed that across many production and comprehension trials, participants would have available representations of the alternative structures to use and be encouraged to name agents in both passive and active structures. By the distribution of trials across and within blocks (see below), it would often be the case that a production trial was preceded by a similar structure either in production or comprehension, at least more so than in a neutrally designed experiment with unrelated fillers. We therefore predicted that the availability of agent mentions and active structures across the experiment will increase the likelihood of active structures with two animate entities, resulting in a decreased overall preference for passives compared to that found in previous research. Crucially, we predicted that similarity-based interference between animate nouns would nevertheless influence structure choice, leading to an increased use of passives for more similar nouns.

## 2.1 Methods

*2.1.1. Participants:* 16 native English speakers from the undergraduate population at the University of York completed the task for course credit.

*2.1.2. Materials:* 42 grey-scale scenes were constructed using images from Clipart.com. 21 of the scenes were taken from GMM and an additional 21 items were constructed in order to increase the similarity range. The full list of pictures is provided in the supplemental materials. Because action verbs that can have both animate and inanimate patients (e.g., *push a person* or *push an object*) are relatively few, a number of pictures represented the same verb/action (e.g., *push*). For example, there were two and sometimes three different pictures in which pushing or hugging events were depicted with different objects and characters. The scenes had the general structure of Figure 1, with an animate and an inanimate entity being acted upon, plus additional entities triggering the felicity of a relative clause. Within the animate items, some cases were highly semantically and functionally similar (e.g. a boy and a girl), and other cases were less similar (e.g. a man and a crab, a girl and a dog).

A pre-test study determined that characters and actions in the pictures were equally codable across conditions. In this study (N=13), participants provided relative clauses to describe each of the animate and inanimate recipients of the action (using the same instructions as described for the main study). The participants provided descriptions for all 84 experimental events (two queries per picture). Across participants, the verbs used to describe the actions had a high name agreement, which was matched across conditions (mean animate = 78.83%, SD = 22.24; mean inanimate = 82.36%, SD = 35.40; $t(82) = -.55$, ns). There was also high name agreement both for the head noun and for the agent noun in the relative clause, and these were also matched across conditions (N1: mean animate = 85.35%, SD = 22.72; mean inanimate = 80.95%, SD = 21.30; $t(82) = .91$, ns; N2: mean animate = 76.61%, SD = 22.28; mean inanimate = 83.74%, SD = 23.43; $t(82) = -1.43$, ns).

For the comprehension trials of this experiment, a picture and a description of a character in it was provided. Each of the 42 pictures was paired with four relative clauses as in (1)-(4) that described either the animate or the inanimate entity being acted upon with either a passive or an active structure. This resulted in 168 trials that were split into two lists (see below). The words used in these descriptions were taken from the pilot study above and therefore they contained nouns and verbs that had been used by producers. All relative clauses mentioned the agent of the event, including all passives and thus *by*-phrases were always used to encourage attention to the agents of the events.

*2.1.3. Similarity rating study:* An online similarity rating study was conducted in order to obtain a measure of similarity of the two relevant characters in the event. Here, the participants were presented with the pictures in which the to-be-rated entities were each highlighted with a red box. The ratings were given on a 7 point scale with 1 being "not at all similar" and 7 being "highly similar". When giving their ratings the participants were encouraged to focus not only on any physical similarity between entities but also their semantic similarity (e.g. similarity in function) as in GMM. Ratings were collected from a total of 25 participants. There were two lists of items with the animate and inanimate prompts of each picture presented on separate lists. We used inanimate entities in this study to provide the full range of similarity that would be seen in the production task. The item analysis showed that, as expected, the difference in similarity between the animate-animate pairs and the inanimate-animate pairs was statistically significant ($M_{animate}$ = 4.70, SD = 1.13; $M_{inanimate}$ = 1.79, SD = .42; $t(41,1) = 15.77$, $p < .001$). Also, as intended, within the animate condition there was a good range of similarity ratings (from 2.57 to 6.18) with steady increases across the 42 items.

*2.1.4. Design and procedure:* The 42 pictures were presented in a web-based questionnaire with production and comprehension trials. The pictures had two versions each, one in which the animate entity was highlighted by a red square and another in which the inanimate entity was highlighted (see Figure 2). In the production trials, each picture was presented in the center of the screen, with an empty text box below the picture. Participants were instructed to look at the picture and answer the question *Who or what is the highlighted entity?* by typing their response in the text box provided. Before beginning the study participants were presented with instructions that gave several examples and encouraged the use of relative clause structures (the instructions were identical to those used in GMM). In the comprehension trials, the pictures were presented with a relative clause (either an active or a passive) underneath. The task for comprehension trials was to indicate whether the sentence accurately described the highlighted entity (Figure 2). 35 filler-trials in which the sentence and the picture did not match were intermixed with the comprehension trials in the order indicated below. 35 filler production trials in which the response was not one of the targeted responses were intermixed with the production trials. The filler trials contained a variety of definite descriptions (e.g., *the man on the phone, the boy swimming, the man wearing a hat, the man sitting at the piano*).


Figure 2: Examples of production and comprehension trials in Study 1

The manipulation of animacy and structure on comprehension trials in this study did not intend to elicit priming from individual items. In other priming paradigms, and in particular in the priming from comprehension to production (e.g., Bock, et al. 2007) the order of items is carefully manipulated such that a priming item is located some number of trials before a to-be-primed target. In the current study we aimed to generally increase the accessibility of all passive and active relatives without favoring a particular one, or expecting a specific priming effect at some point in the list. Generally, production trials were preceded by a variety of comprehension trials containing examples of all relative clause types with equal probability, and across the comprehension trials there was an equal number of cases instantiating each type. To achieve this, an algorithm was developed to compute transitional probabilities between trials: The order of comprehension trials in a list was such that the probability of one relative clause type following all four types or fillers was equal. For example, if an animate-head passive was shown in one trial, it was followed by an animate-head active, animate-head passive, inanimate-head active, inanimate-head passive or a filler trial with equal probability. A similar transitional probability constraint was used for production trials where a given query was followed by an animate or inanimate query or a "filler" query with equal probability.

In addition to this general constraint, the animate and inanimate versions of each picture to be described were assigned to two different lists, but each participant experienced an equal number of cases per animacy condition. There was a total of 160 trials in each list. In each of the two lists, 7 blocks of comprehension trials (varying between 9 and 10 trials) were intermixed with 7 blocks of production trials (with either 5 or 7 trials), each including filler items. Because comprehension trials were more numerous than production trials, half of them were allocated to list 1 and the other half to list 2 in such a way that a given picture would appear with active or passive structures of the opposite animacy to those queried in the production trials. For example, if the list contained a picture querying the animate entity in the production trial, it would only appear in two comprehension trials paired with inanimate-head structures. This eliminated noun (character) repetitions across blocks, as different events of the same picture would be described in the list, although some verbs or actions may repeat. A given picture appeared within a list on average 41 trials apart from the previous trial using the same picture (either in production or comprehension). A given verb used in a comprehension trial or potentially to be used in a production trials also appeared in a list with 28 trials apart on average. This precludes priming by repeated pictures or repeated verbs from recently presented trials in most cases.

*2.1.5 Data coding and analyses*: The production data were coded for syntactic structure (active or passive), and accuracy. Production responses were coded as errors and removed from the analysis if they did not include a relative clause or if they provided an inaccurate description of the highlighted entity. Production and comprehension accuracy overall was very high (production: M = 96.7%; SD = 5.46; comprehension: M = 95.8%; SD = 8.32) so most of the trials were included in the analysis. In the comprehension trials, there was slightly better performance in relative clauses with inanimate heads than animate heads ($M_{inanimate}$ = 99%; $M_{animate}$ = 97%). The analyses reported below used non-parametric statistics (Wilcoxon Signed Ranks and Mann-Whitney tests) because proportional data are not normally distributed, although these statistics do not provide straightforward ways to test for interactions.

## 2.2 Results

### *2.2.1 Relative clause structure choice and agent omissions*

As in previous studies, there was a significantly greater proportion of passive structures produced for animate-head than inanimate-head relatives (Wilcoxon Signed Ranks Test: by-participants: -2.99, p=.003, by-items: z= -3.13, p=.002) (see Table 1). However, as expected, there were many more animate-head actives and thus fewer passives produced compared to GMM's written production study 1a (Table 1). Comparing the present results with those in GMM indicated that the animacy effect was modulated by the experimental context: Using as the dependent variable the difference between the proportion of passives produced for animate- and inanimate-head conditions for each participant, there was a main effect of experiment (Mann-Whitney test: -3.33, p=.001), suggesting that the animacy effect was smaller in the present experiment compared to that of GMM. Specific contrasts comparing the proportion of passives in each condition separately across the experiments indicated a main effect of experiment only for animate-head structures (Mann-Whitney test: -5.91, p=.009). These results indicate that the contextual manipulation succeeded in increasing the likelihood of producing active object relatives for animate-heads compared to the previous study, despite more passives being produced for animate-head than inanimate-head clauses in each experiment. Interestingly, as shown in Table 1, structural preferences did not change across experiments for inanimate-heads clauses, indicating that inanimate-head relative structures were not influenced by the context, perhaps because production is already easy in these cases and cannot be facilitated further.

| | Proportion of passives | | Proportion of *by*-phrase omission | |
|---|---|---|---|---|
| | Animate head | Inanimate head | Animate head | Inanimate head |
| Study 1 | .60 | .46 | .22 | .15 |
| GMM | .97 | .50 | .51 | .25 |

Table 1: Results of Study 1 and a previous study from Gennari et al. 2012 (GMM)

The overall rate of agent omissions was also smaller in the current study than in GMM, suggesting contextual effects: 22% of all animate-head passives had agent omissions compared with 51% in GMM, with parallel differences in the inanimate-head condition (Mann-Whitney test for by-phrase omissions in animate-heads: $z=-2.63$, $p=.009$; Mann-Whitney test for inanimate-heads: $z= -2.52$, $p= .01$). Nevertheless, there was a significant difference in *by*-phrase omissions across the animacy conditions in the present experiment (as there was in GMM) (only reliable in the by-participant analyses, Wilcoxon Signed Ranks Test: -2.83, $p=.005$). This suggests that despite the overall reduction of agent omissions compared to GMM, there was still a tendency to omit agents when two animate nouns were being talked about. This is consistent with a role for similarity-based interference in production, which is nevertheless modulated by contextual availability.

### *2.2.2 Similarity and structure choice*

To test the key prediction of the current study that similarity-based interference influences structural choices, we correlated the proportion of passive structures produced for each item with the similarity ratings representing agent-patient similarity in the animate condition. We found that the higher the similarity rating, the greater the proportion of passive structures produced ($r = .47$, $p < .01$; see Figure 4, left panel). This suggests that despite the overall increased availability of alternative structures and the general mention of agents encouraged in the experimental context, semantic similarity played a role in structure choice: The more similar the animate nouns in a relative clause were, the more likely they were to be mapped into passive structures—arguably due to similarity-based interference.

## 2.3. Discussion

Although the design of our study cannot conclusively establish which specific aspect of the context modulated production—either the repeated mention of agents or the increased availability of active object relatives, or both, may be the source of the contextual facilitation— the results of Study 1 suggest that similarity-based interference influenced active vs. passive structural choices for animate-head clauses when alternative structural frames and agent mentions were made available in the experimental context. This contrasts with previous studies showing only animacy effects in active vs. passive choices and late similarity effects in *by*-phrase omissions (GMM). The present results therefore suggest that active vs. passive choices, reflected in function assignment and verb morphology, are driven not only by categorical animacy distinctions, but also depend on the similarity between the entities being talked about and contextual characteristics. The role of context found here is consistent with corpus studies indicating that although few in number, object relative clauses are produced when the second noun is highly salient in the discourse or given (Fox &

Thompson, 1990; Roland, Mauner, O'Meara, & Yun, 2012). This is why the most acceptable and easy-to-understand object relative clauses are those in which a pronoun is used, as in *the student I examined* (Warren & Gibson, 2002), suggesting that contextual availability of entities in the discourse modulates both production choices and comprehension difficulty.

Importantly, the role of similarity was graded so that more similar animate entities were produced as passives, whereas less similar ones were mapped onto active object structures, as indicated by our correlation result. This indicates that contextual characteristics exerted little influence on high-similarity cases, whereas they promoted active structures for low-similarity cases. The presence of a highly similar competing agent when retrieving the animate head-noun led to stronger interference and possibly stronger inhibition of the competitor after selection of the head noun. Strongly suppressed agents were then not accessible for mention right after the head noun and resulted in the production of passive structures, as suggested in GMM. In such cases, contextual characteristics were not strong enough to alleviate or release interference during planning. However, for less similar competing agents, interference was weaker and contextual properties sufficiently boosted their activation or accessibility so that they could be produced after the head noun in active object relatives. We hope to shed more light on these processes in relative clause production in Study 3.

In sum, Study 1 and previous work suggest that similarity-based competition plays a role in the planning and production of a relative clause: The more similar the entities to be talked about, the stronger the tendency to produce passive structures and omit agents, even when the context promotes the use of agent mentions and alternative structures. Conversely, less similar entities are more likely to be mapped into active structures when the context makes these structures or agent mentions more available. Thus structural choices, including function assignment and the planning of verb morphology, are influenced by the degree of similarity between the entities being talked about. In the next study, we explore whether similar semantic interference effects can be observed in the comprehension of relative clauses.

## 3. Study 2: Similarity-based interference in relative clause comprehension

Previous comprehension studies have already established that animate-head active relatives like (2) are more difficult to comprehend than inanimate-head active relatives and subject relatives (Gennari & MacDonald, 2008; Gordon et al., 2006; Mak et al., 2006; Traxler et al., 2002). These studies typically involve on-line measures such as word-by-word reading times or eye-tracking fixation measures while reading relative clauses in larger sentential contexts (e.g., *The senator that the journalist attacked was angry*). They typically report increases in fixation durations, reading times and eye-movement regressions at critical points of the sentence and in particular verbs (relative clause verbs or main sentential verbs). This is consistent with the idea that the relative clause nouns are interpreted at the point where their semantic roles and syntactic functions must be established relative to a verb. Before a verb is encountered, readers might predict a semantic or syntactic role for the nouns, in particular for the first noun, which is typically the agent of an upcoming verb. However, such predictions must be revised when the verb is encountered, as the first noun is not the agent or subject of the relative clause verb. This therefore suggests, as argued here, that determining the appropriate semantic-syntactic role of the animate nouns in an active relative clause engenders competition because the two nouns are equally good candidates for an agent or a patient role as well as alternative syntactic functions.

Although the existing evidence is suggestive, similarity manipulations in reading studies of relative clauses have been limited so far. Effects of similarity have been obtained by manipulating the type of referential expressions used within the relative clause. For example, Gordon and colleagues (Gordon et al., 2001, 2006) compared object and subject relative clauses where the noun within the relative clause is a definite description (e.g., *the banker*) or a proper name (e.g., *Sophie*). They found that the magnitude of the difference between object and subject relatives is much reduced when a proper name is used in the object relatives (e.g., *The banker that Sophie praised climbed the mountain* vs. *The banker that praised Sophie climbed the mountain*). Proper names single out a specific individual in the referential domain, thus making this referent more distinctive than that of a definite description such as *the banker.* In another study, the magnitude of the difference between object and subject relative clauses was increased when readers were asked to maintain similar nouns in memory through a secondary task (e.g., maintaining occupation nouns or proper names) (Fedorenko, Gibson, & Rohde, 2006). These results support the idea that less similar animate entities engender less interference and are easier to process than highly similar entities, as argued in memory approaches to sentence comprehension (Lewis, Vasishth, & Van Dyke, 2006; Van Dyke et al., 2014).

Importantly, previous studies have not systematically manipulated the degree of conceptual similarity between the nouns of the relative clause and have not compared object relatives to meaning-matched controls. Indeed, comparing object relatives to subject relatives is not ideal because object and subject relatives have different meanings: the head noun of a subject relative (e.g., *The banker that praised the barber*) is the agent of the relative clause verb, whereas in an object relative (e.g., *The banker that the barber praised*) the head noun is the experiencer of the relative clause verb. A better matched comparison for object relatives, which keeps the meaning and the semantic roles the same across conditions, would be a passive subject relative (e.g., *The banker that was praised by the barber*).

In the present study, we aim to establish whether there are similarity-based effects in comprehension that parallel those reported in production. To this end, we systematically manipulated conceptual similarity as in Study 1 and compared object relative clauses to passive counterparts. Specifically, we used a picture based comprehension task using the materials from Study 1. In this task, participants saw a picture with a highlighted entity as in Figure 2, read a description of this entity, and decided whether the description was accurate. This task requires mapping a linguistic structure onto a depicted event, as in the production task, and furthermore, provides reaction time measures indexing difficulty. Moreover, the use of descriptions rather than full sentences, as in previous reading studies, does not include the additional cost of integrating the relative clause content with a main verb (e.g., with the verb *climbed* in *The banker that the barber praised climbed the mountain*). Arguably, therefore, we can obtain relatively clean measures of the cost of computing the semantic-syntactic roles with respect to the relative clause verb. Hence, in addition to an overall effect of head-animacy, as previously shown, we expect to observe degrees of processing cost as a function of semantic similarity in animate-head active structures.

## 3.1 Methods

*3.1.1 Participants*: 40 native English speakers from the University of York participated in this study for course credit.

*3.2.1 Materials*: The same 42 pictures as in Study 1 were used. Each picture had an animate and an inanimate version in which the relevant entity was highlighted with a red square. Each picture was paired with four description types as in (1)-(4): animate-head active and passive relatives, and inanimate-head active and passive relatives. These descriptions were obtained by using the most common words provided by speakers in describing the pictures, as established in Study 1.

*3.2.3 Design and procedure*: The 42 pictures paired with four descriptions each (a total of 168 unique cases) were allocated to four different lists. In each list, a given picture only appeared once in association with one of the description types, but across all pictures, all conditions were shown to all participants. Each participant therefore experienced either 10 or 11 cases from each of the four conditions (animate-head passive, animate-head active, inanimate-head active, inanimate-head passive). In addition, 42 foil trials were included in which an inaccurate description of the highlighted entity was shown. This thus elicited a roughly equal number of *yes* and *no* responses in the task. These foil trials were similar to the experimental set (there were at least two entities interacting), and contained a highlighted entity that did not match the description provided. To encourage understanding of the roles in the events, the foil descriptions were constructed in such a way that either the semantic roles were reversed (e.g. *The shark that the fish is eating* when the correct description would be *The fish that the shark is eating*), or the verb or the nouns were incorrect (e.g. *The man that is lifting the woman* when the correct description would be *The man that is lifting the weights*; or *The bride that is being tickled by the princess* when the correct response would be *The bride that is being kissed by the princess*).

Participants were instructed to indicate as quickly and as accurately as they could whether the description provided was an accurate description of the highlighted entity in the picture. In each trial, a picture without any red squares was presented in the center of the computer screen for 3000ms. This provided ample time to apprehend the events in the picture. After this period, one of the characters of the picture was highlighted with a red square for 1000ms. The picture then disappeared and a description was provided in the center of the screen until a button box response was recorded. Previous practice familiarized participants with the location of the *yes* and *no* buttons, and they were told to keep their hand on the button box throughout the experiment.

Only responses to the experimental *yes* trials were analyzed. We used the raw reaction time data in the analyses, as well as residual reading times, which were calculated by regressing out phrase length in number of characters to account for the fact that passive structures are longer than active structures. Analyses conducted on both of these measures were comparable in every respect, so only the raw reaction time data is reported. To test whether the time taken to process the relative clauses is related to agent-patient similarity, a correlation was performed with the RT data and the similarity ratings for the animate-head active items.

3.2. Results

*3.2.1 Accuracy:* The average accuracy was high across all conditions (see Table 2). As can be seen in in Table 2, animate-head clauses were generally more difficult, and thus were less accurate than inanimate head clauses for both active and passive structures. Statistically the only reliable difference was for active structures (Wilcoxon signed ranks tests (1-tailed): by-participants: $z = -1.82$, $p = .03$, by-items: $z = -1.67$, $p = .06$). This is suggestive of more difficulty in active relative clauses with two animate nouns, where the wrong semantic roles may have been assigned to the nouns, and thus the description was rejected when in fact it was correct.

Table 2: Accuracy results from Study 2

| | Relative clause structure | |
|---|---|---|
| Head-Noun | Passive | Active |
| Animate | 92.62% (7.98) | 91.43% (1.57) |
| Inanimate | 95.24% (7.73) | 95.24% (7.40) |

*3.2.2. Comprehension decision times:* For the reaction time data, only accurate responses were analyzed. Responses that deviated 3SD from the mean or fast outliers were excluded. This resulted in the removal of fewer than 3% of responses. The average reaction times across items for each condition are shown in Figure 3. A repeated-measures ANOVA with condition (animate- or inanimate-head) and structure (active or passive) as factors showed a significant main effect of condition ($F_1(1,39) = 18.74$, $p < .001$; $F_2(1,41) = 9.57$, $p < .01$) and structure ($F_1(1,39) = 29.66$, $p < .001$; $F_2(1,41) = 19.53$, $p < .001$), and a significant interaction ($F_1(1,39) = 20.40$, $p < .001$; $F_2(1,41) = 8.76$, $p < .01$). As predicted, t-tests showed that responses were slower for the animate-head actives than the inanimate-head actives ($t_1(39) = 6.54$, $p < .001$; $t_2(41) = 4.03$, $p < .001$). However, there was no significant difference between the passive conditions ($t_1(39) = .89$, ns.; $t_2(41) = .53$, ns.). The results therefore suggest that decisions for animate-head active clauses (with two animate nouns, e.g. *the man that the woman is punching*) were more costly than any other conditions, suggesting competition in establishing the semantic or syntactic roles of the nouns within the relative clause.

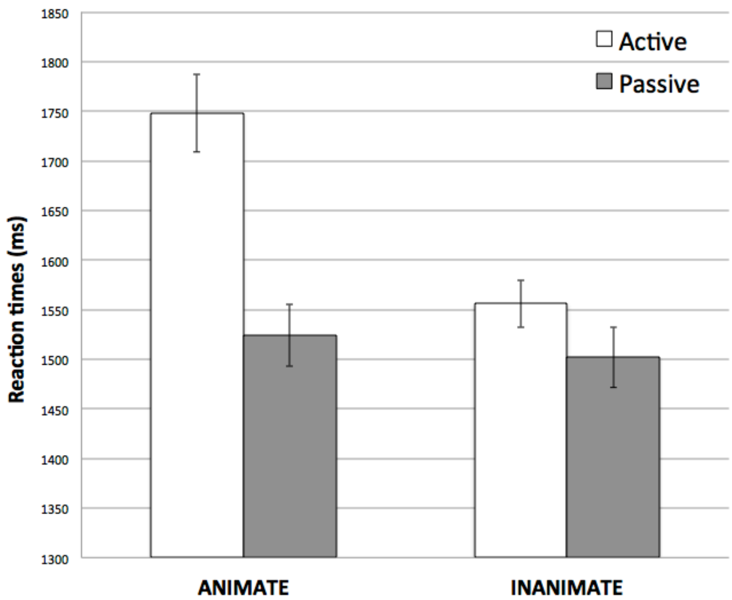


Figure 3: Response times in Study 2 as a function of animacy and structure condition. Error bars represent standard errors.

*3.2.3 Similarity and processing cost*: If processing difficulty is related to similarity-based competition between the concepts to be interpreted relative to the verb, then a correlation should be found between the degree of agent-patient similarity in the animate-head active items and the response time data. Indeed, a significant positive correlation was found ($r$ = .35, $p$ =.02; Figure 4, right panel) for active structures, whereas no such correlation was observed for animate-head passive structures.

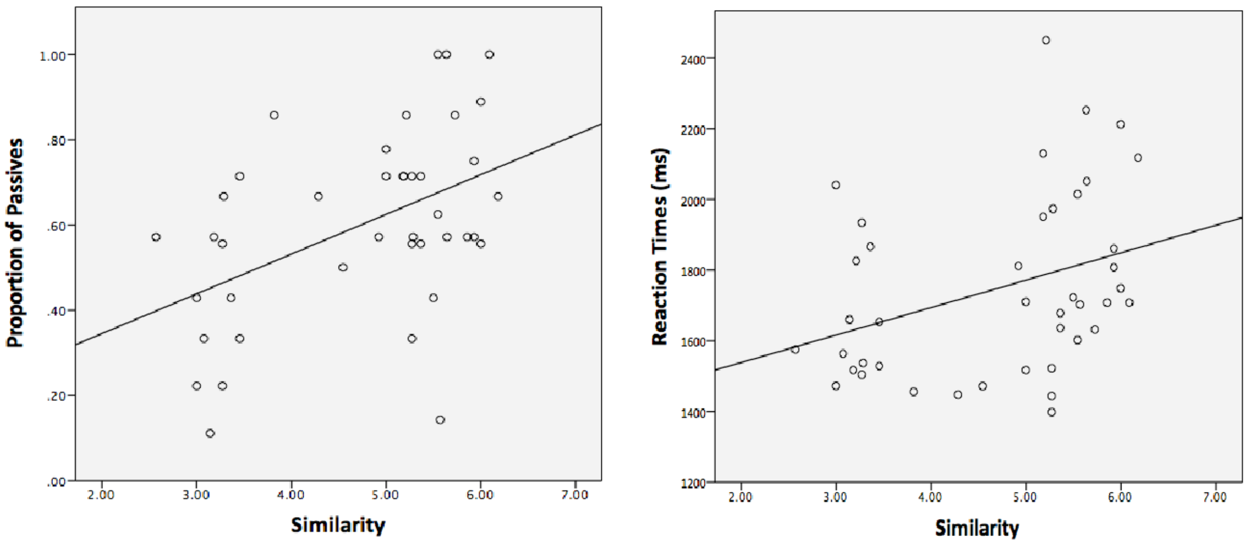


Figure 4: Correlations between similarity and the proportion of animate-head passives produced in Study 1 (left panel) and the response times to animate-head actives in comprehension in Study 2 (right panel).

3.3. Discussion

This study aimed to investigate whether similarity-based competition occurred in relative clause comprehension using a task comparable to that in production as in Study 1. The results confirmed the presence of similarity-based competition. Comprehending animate-head active clauses such as *the man that the woman is punching* was more difficult than inanimate-head clauses, and the degree of agent-patient similarity in animate-head active clauses predicted processing cost. We expected to find a competition effect in animate-head active clauses and not inanimate-head ones or passives because competition should influence the interpretation of the nouns when both must be maintained in memory and considered nearly simultaneously at the point the verb is encountered. Thus, semantically similar nouns that share likely features of agents and patients (and can equally be objects or subjects) induce processing cost. In contrast, the lexical meaning of inanimate-head clauses unambiguously signals a patient/theme role (object function) relative to the upcoming verb and thus these structures elicit no difficulty. Similarly, passive structures do not elicit competition because the head-noun can be interpreted as the subject of the verb bearing a patient role at the verb phrase (e.g., *being punched*), due to the incremental nature of processing (e.g. Kamide, Altmann, & Haywood, 2003). Thus the processing of the two nouns does not overlap in time. These findings therefore add to the existing accounts of sentence processing and provide strong evidence for the role of graded semantic similarity as a major influence on relative clause comprehension (Fedorenko et al., 2006; Gordon et al., 2006; Lewis & Vasishth, 2005; Van Dyke & McElree, 2006).

## 4: Study 3: Eye-movements during the production of relative clauses

Studies 1 and 2 provided evidence for the role of semantic similarity in production and comprehension of relative clauses. The higher the semantic similarity between the two animate nouns being planned or comprehended, the more likely those nouns are to be produced within passive structures, and the harder they are to comprehend when within active structures. Both these processes have in common that the two animate nouns are planned or interpreted close in time, and thus the processing load is higher than in simple structures. However, whereas it is clear that in comprehension similarity-based competition is manifested at verb positions (Gennari and MacDonald, 2008; Gordon et al., 2001, 2006), in production, to our knowledge, there is no evidence of similarity effects when verbs are planned or encoded. We have seen in Study 1 that the retrieval or selection of the head noun may interfere with potential competitor agents in the visual context and the event being described. Indeed, the contextual manipulation in Study 1 indicated that more passive structures were produced for more similar entities, suggesting a role of similarity-based interference in reducing the availability of competitor agents despite facilitation in the overall production of actives. Moreover, Study 1 suggested that functional choices for subjects or objects are modulated by similarity, with low-similarity but not high-similarity entities resulting in actives due to contextual facilitation. This provided preliminary evidence that directly or indirectly the planning of the verb and its passive morphological form would also be subject to the effect of semantic similarity. But is there a role for similarity in production, particularly, when establishing syntactic functions at the verb, as in comprehension?

In the present study we aim to shed some light on these issues by examining the time course of relative clause production using eye-tracking measures in a picture-based task. Because this is the first eye-tracking study to investigate relative clause production, and the design and task follow the same naturalistic question-answer procedure as in GMM for comparison, the study is by necessity exploratory in nature and not perfectly controlled. For example, comparing fixations across animate and inanimate entities may be not entirely informative because in some cases animate entities are larger than inanimate entities, and thus have a larger area within which fixations can land on. For this reason, we have drawn our conclusions from the fixation differences between animate entities, which were more comparable in size in our picture stimuli. Similarly, participants were not instructed to use a pre-determined relative clause frame, and therefore, their productions display greater variability in structures and length (e.g., *the man wearing boots being punched on the chin* vs. *the man being punched* for Fig. 1). Such variability may influence what the eyes are doing at a given point. We have taken some of this variability into account by aligning and averaging fixations to the beginnings and ends of critical words on a trial-by-trial basis.

As discussed in Study 1, the accessibility of entities and their names can be modulated by contextual factors, such as the presence of competitors. In the present paradigm, speakers first inspected a visual scene like Figure 1 for three seconds and then heard a question. Before hearing the question, participants may apprehend the events depicted, but given that the aspect of the scene being queried is only established at the question, participants are unlikely to have planned a specific answer or have assigned a specific syntactic function to characters during this period (experimental and filler questions queried objects, agents, patients, locations and actions in the scene). Thus, in grammatically encoding the answer to a question such as *who is bald?,* speakers must retrieve an appropriate term (and formulate an appropriate phrase) to characterize the target character as distinct from other potential competitors, such as the other men in the scene and the agent of the event, which nevertheless needs to be considered in formulating the relative clause (the distinctive feature of the man is that he is being punched by the woman). This process thus requires entertaining very specific and distinctive representations of the entities being talked about, and is susceptible to interference by semantically similar entities (Arnold & Griffin, 2007; Fukumura et al., 2011). Indeed, eye-tracking studies of object naming suggest that although targets are fixated much more relative to other entities in the picture before speech, there are more fixations on entities sharing visual or semantic features with the target relative to unrelated objects, even if there are no naming time differences between competitive and non-competitive situations (Huettig & Hartsuiker, 2008).

Based on previous research and the inherent high accessibility of animate entities, we therefore reasoned that animate-head targets like the bald man in Figure 1 would dominate attention before speech while retrieving an appropriate term for it. Moreover, because the agent woman is also inherently accessible and a relevant competitor sharing features with the target, one could expect fixations on this character to some extent, and in particular, more fixations than unrelated objects. In contrast, when planning inanimate-head clauses like *the sandbag the woman is punching* in response to *what is orange?,* we expect more fixations on the agent of the corresponding event (the woman in purple) compared to the agent of animate-head targets (the woman in yellow). This is because (a) inanimate targets are inherently less salient and in the scene they interact with more salient agents that attract fixations, (b) agents do not share semantic features with the inanimate target and thus do not interfere with each other, and (c) in response to the question, the way to successfully identify the inanimate target is to specifically attend to the event in which it is taking part.

These predictions are consistent with previous research and widely assumed accessibility considerations. The question now is what we should expect the eyes to do during the production of the relative clause. Previous sentence production studies have already established that as speakers utter a word, they are simultaneously lexically encoding subsequent words (Griffin & Bock, 2000; Konopka & Meyer, 2014). The process of lexical encoding is reflected in the fixations to the appropriate referent (e.g. Griffin & Bock, 2000). For example, when producing simple sentences like *the dog is chasing the man*, fixations on the depicted man peak during the utterance of the subject noun, because the lemma *dog* has already being selected and during its articulation speakers lexically encode the upcoming words. Thus, if there is any evidence of similarity effects beyond that occurring when the first head-noun is selected, we would expect that such evidence should occur when the first head-noun is being uttered. This should be particularly the case in passive structures because these structures are more likely produced when the entities are most similar, i.e., the strongest competition takes place, as suggested in Study 1. Thus, while the first noun is being uttered (e.g., *the man* in *the man being punched by the woman*), the verb phrase is being lexically encoded. During this process function assignment should come into play, because the verb morphology should agree in form and meaning (*being punched*, rather than *is punching*) with the head of the relative clause. GMM indeed suggested that competition between syntactic roles of the two animate entities may occur (as it appears to do in Spanish) because the two animate lexical concepts and their features are consistent with both agent and patient roles and do not unambiguously indicate a subject or object function. In contrast, no such competition occurs in inanimate-head cases because their lexical concepts already indicate an unambiguous semantic role in the event.

If this is the case, then, what pattern of fixations should we expect as evidence of competition *during* production of passives, and particularly at verb encoding? Assuming that the extent to which competitors are fixated reflects their degree of activation in the speaker's mind, as implicitly assumed in eye-tracking studies (Huettig & Hartsuiker, 2008; Wolter, Gorman, & Tanenhaus, 2011), we would expect relatively fewer fixations on agents in animate-head passives compared to inanimate-head passives, due to similarity-based mechanisms in the former but not the latter. That is, in competitive situations (animate-head targets), the agent is relatively less activated due to interference with the target, compared to non-competitive situations (inanimate-head targets), although it should still be more active than unrelated entities, as utterance-relevant characters are. More importantly, whatever the relative pattern of fixations, we would expect that the extent to which agents are fixated in the animate-head condition should negatively correlate with the similarity between the two animate entities involved because after target selection, more similar competing alternatives should be less accessible (or more inhibited) than less similar alternatives. For example, compared to the case of the man and the woman in Figure 1 (left panel), where both characters are equally likely agent subjects, in an item asking for *the baby being held by the woman* (Figure 1, right panel), the semantic features of *woman* and *baby* both would cue agent subject roles to some extent, but the woman concept (and associated world knowledge) would more strongly cue the agent subject of holding. Thus the competition between the two would be less strong, attention on the target would be easier to maintain, and the target's patient role relatively easier to indicate with the appropriate verb morphology. This should result in more fixations on the woman holding the baby relative to the woman punching the man. Thus, less similar animate agents, which interfere less with the target, should be fixated more compared to highly interfering entities. This result is critical if indeed the degree of similarity between the entities plays any role at the point of encoding the verb. This result then would nicely parallel the competition effects observed in comprehension.

4.1 Methods

*4.1.1. Participants:* A total of 42 undergraduate students at the University of York participated in the study for course credit or payment. They were native speakers of English, with no reported speech, language or reading problems, and normal to corrected-to-normal vision. Three participants were excluded from the analyses below as their productions included fewer than 50% of relevant sentence structures (see below for details).

*4.1.2. Materials:* 21 scenes previously used in GMM were used (half of the pictures in the Supplemental Materials). Each scene had the structure of Figure 1, where two different events and corresponding contrasting entities were represented. Note that across the 21 scenes used, there was variability in the similarity between the two animate entities involved in the events (GMM). The procedure for obtaining similarity ratings is provided in GMM.

Each scene was paired with 3 questions, one about the animate patient (e.g., *Who is bald?*), another about the affected entity (*What is orange?*) and another about either of the agents in the scene (e.g., *Who is wearing yellow/purple?* for the women in Fig. 1). As in Study 1, the first two questions lead to the production of an animate-head or inanimate-head relative clause about the patient or the affected object respectively, thus eliciting a choice between an active object relative and a passive relative. The third question was included to add variability to the produced structures and elicit subject relatives (e.g., *The woman who's punching a bag*). In addition to experimental scenes, we included 40 filler scenes paired with questions eliciting different structures (e.g. *What is the man with the glasses doing? What did the girl just do?*).

*4.1.3. Design:* Experimental scenes and questions were arranged in 3 lists. Each list presented all scenes once. Among the experimental scenes, 7 were paired with animate-patient queries, 7 with inanimate-themes queries, and 7 with animate-agent questions. Each participant thus only saw each experimental scene once and answered only one question about it. The filler scenes were the same across the three lists.

*4.1.4. Procedure:* Participants were seated in front of a 22-inch display monitor, with their eyes approximately 60 cm away from the monitor. Their eye movements were recorded by an Eye Link II head-mounted eye tracker, sampling at 250 Hz. Participants' verbal responses were digitally recorded using a microphone positioned in front of the monitor. The images were 640 x 480 pixels, presented at the center of the display screen (with a 800 x 600px resolution). The questions were presented auditorily via two loudspeakers located at each side of the display monitor.

Each trial started with a drift correction dot presented in the center of the screen. After the participant looked at the dot, it was replaced by the image. The question was played over the loudspeaker 3 s after the image onset. The image stayed on the screen for 6 s after the offset of the question. Every 5 trials, a standard 9-point calibration procedure was performed to avoid drift. There were 4 practice trials before the main experimental block. The entire session lasted approximately 40 minutes, including 5 to 10 minutes of introduction and calibration plus 30 minutes for the experimental session. Instructions indicated that participants would be presented with a picture, and after inspecting it carefully, they had to answer the questions played on the speakers. As in GMM, we discouraged the use of descriptions based on position or color, such as *The man on the right*, or *The man in blue* and suggested that the best way to provide a clear identification was to describe the action that the characters were engaged in (GMM).

*4.1.5. Verbal responses coding and analysis:* Verbal responses were transcribed by research assistants blind to the purpose of the study. For the experimental pictures in which an agent was queried, participants routinely produced subject relative clauses; these responses are not discussed further here. For the items querying an animate or inanimate affected participant, responses were coded by structure: active object relative, passive relative (with and without the *by*-phrase) or other structures. Any response containing the verb *to be* or *get* with a past participle was considered a passive structure irrespective of whether it contained a relative pronoun (*who(m), that*) or an agent by-phrase (e.g., *The man being hit, The man getting hit by the woman, The man who/that is being hit*). Analyses reported below exclude incorrect responses (identification of an entity other than that being intended, e.g., *The woman's hair* in response to *What is orange?* in Figure 1), irrelevant responses (those that fail to uniquely identify a character in the scene, e.g., *The man* in Figure 1) or responses with a structure other than the targeted passive or active relative clause (e.g., *The man fighting with the woman*, *The man wearing boots,* in Figure 1). Some responses of this type are inevitable in a relatively unconstrained production task like this, where we did not directly instruct participants to produce relative clauses. There was no significant difference in the average proportion of excluded responses in the two animacy conditions (animate: 9%, inanimate: 14%; Wilcoxon's signed-ranks tests by participants: $z$ = -1.55, p = .12; by items: $z$ = -.57, p = .57). Participants who had more than 50% irrelevant responses in the experimental items were excluded from the analysis, resulting in the removal of three participants. One experimental item was also removed for eliciting more than 50% of irrelevant responses. On average, there were 12 responses with targeted structures per participant in the analyses. For each participant and each experimental item, we computed the proportion of passive relative clauses produced in each animacy condition out of all included responses.

In each spoken response, a trained research assistant blind to the conditions of the study marked onsets and offsets of relevant elements in the structure using Praat (Boersma, & Weenink, 2015). The markers indicated the onset and the offset of critical words in the sentence if present (first determiner, first noun, relative pronoun, verb auxiliary, main verb, *by*-preposition, second determiner, second noun). Timings were also obtained for the queries in each trial. Specifically, in each question the onset and the offset of the word providing the critical information to identify a character (the question critical word) was identified. For example, for *Who is bald?* the critical word is *bald*, whereas for inanimate queries, this word would typically be a color word (*orange*). The timings of all these markers were then aligned with the eye-tracking data starting at the beginning of the trial.

*4.1.6. Eye-movement data analyses:* The analyses of eye movements included fixations to different entities in the scene, which were drawn on the image files on a pixel-by-pixel basis covering the whole of the entity. In each scene and for each query type (questions about animate or inanimate entities), the *target entity*, the *contrast entity* and the *agent* of the event were defined. For example, for Figure 1 and for the animate query (*Who is bald?*), the target entity was the bald man, the agent was the woman punching the man, and the contrast entity was the men in the background. For the inanimate query (*What is orange?*), the target was the orange sandbag, the agent what the woman punching this sandbag, and the contrast entity was the blue sandbag.

To tests the hypotheses outlined above, we focused on animacy effects and on specific time windows before and after speech initiation. (For an overview of fixations over the pre-question 3secs period, see Appendix 1). Computed from the onset of the question critical word, participants took 2821ms on average to start speaking (ranging from 1340ms to 13388ms). Due to the high variability of speech initiation latencies, we examined pre-speech fixations aligning them in two ways. First, we examined fixations aligned to the question critical word, thus starting from the onset of the question critical word and up to 2000ms post critical word onset. These after-question fixations included the entities that were fixated soon after the target entity being asked about could be identified, and therefore reflect aspects of utterance planning. Second, a 1000ms window before the onset of the head noun (i.e., excluding determiners or disfluencies if any) was chosen because previous results suggest that fixations on the character to be named peak within this time window (Griffin & Bock, 2000). Analyses in this window therefore aimed to examine the pattern of fixations while encoding the head noun of the relative clause. Note that these two analyses include overlapping data points for those participants who started speaking before the 2000ms window from the question critical word has elapsed, but should provide an overall picture of fixations before the head noun is uttered.  Finally, we examined fixations during speech, by synchronizing them with onsets and offsets of each component phrase on a trial-by-trial basis for each participant to establish the time course and distribution of fixations on targets and agents during utterance production.  Similarity ratings obtained by GMM were used in these analyses to evaluate whether the similarity between the two animate entities involved in animate-head structures correlated with the likelihood of fixating the agents in the scene while uttering the head noun (i.e. during verb planning).

For statistical analyses, across participants and items we transformed the number of trials in which a region of interest was fixated at a given time window onto the log odds scale using an empirical logit transformation as follows: empirical logit = (number of trials + .05 / total number of cases in a condition – number of trials + .05) (Barr, 2008). Log odds transformations result in unbounded distributions centered on 0 that can be analyzed with standard statistics. We then used standard statistical analyses across participants and items to compare average logits across conditions for the pre-defined regions. For display purposes, figures show untransformed proportions of fixations. For the analyses before the head-noun onset, we averaged the fixations of all inanimate-head structures, whether passives or actives. For the analyses during speech, we only examined animate- and inanimate-head passive structures, given that in object relatives the word order implies different words being planned at different time points. Finally, because comparison across passive structures with and without agent-omissions involved relatively few data points, all passive structures within a condition were analyzed together, regardless of agent omissions.

## 4.2 Results

### *4.2.1. Relative clause structure choice and by-phrase omissions*

In the animate-head condition participants mostly produced passive relatives (94%), and in the inanimate-head condition the responses were split between passive (55%) and active (45%) relatives. Wilcoxon's signed-ranks test confirmed that participants produced significantly more passives in the animate- than the inanimate-head condition (by-participants: $z$ = -4.63, p < .001; by-items: -3.94, p < .001), thus replicating the findings from Study 1 and previous research. As before, there was also a higher proportion of agent omissions in animate-head (37%) than inanimate-head (17%) passive relatives (by participants: $z$ = -4.53, p < .001; by items: -1.96, p = .05). The rate of agent omissions correlated with the similarity ratings for the two animate entities within the relative clause (*r* = .42, p = 0.03 (1-tailed)). Together, these results replicate those previously reported and confirm that the animacy and the similarity between the animate nouns in relative clauses influence structure choices in production (Gennari & MacDonald, 2009; Gennari, Mirkovic, & MacDonald, 2012).

*4.2.2. Eye movements before production*

The key contribution of Study 3 is in assessing the time-course of production processes, and specifically the points in the structure where similarity-based effects occur. For illustration purposes only, Figure 5 shows the average proportion of fixations on each relevant character for each condition (animate or inanimate query), as measured from the question critical word onset up to the beginning of speech. At each time bin lasting 200ms, the average proportion was obtained over a 500ms window (500ms windows were sampled every 200ms, resulting in partially overlapping windows, e.g. at time 0, the average of the 0 to 500ms is shown, at time 200, the average of 200ms to 700ms are shown, and so on[1]). The figure shows averages across different types of inanimate-head clauses (actives and passives) because there was no difference between these structures in the averaged proportions targeted here. *Targets* represent fixations on the queried entity in each condition, for example, the bald man or the orange sandbag in Figure 1. *Agents* represent fixations on the corresponding agent in the event (the woman in the yellow shirt or the woman in the purple suit in Figure 1), and *contrasts* represent fixations on contrasting entities for each corresponding events such as the men in the background or the blue sandbag. *Other* represents the entities that were irrelevant

[1] This procedure was adopted simply to make easily apparent the pattern of fixations, as in cumulative proportion plots often used in eye-tracking studies. Computing fixations, say, in each 200ms window with no overlap between them, makes the graph noisier due to great variability in the data set, as different participants started speaking at different times. See the range of speech onset times below.

for the given description being planned (e.g., for animate targets, this category includes inanimate objects and the woman in purple in Figure 1). Computed from the onset of the critical word in the question, the average speech onset time for the animate-head passive relatives was 2646ms (range 1199-13388ms), and the average speech onset time for active and passive inanimate-head structures 2822ms (range 1340-11222ms). Note that the entities that are irrelevant for the target event are typically fixated less than the relevant ones, suggesting goal-directed behavior, independently of other inherently salient animate entities in the picture. For an overview of fixations during the pre-question 3sec period, see Appendix 1.

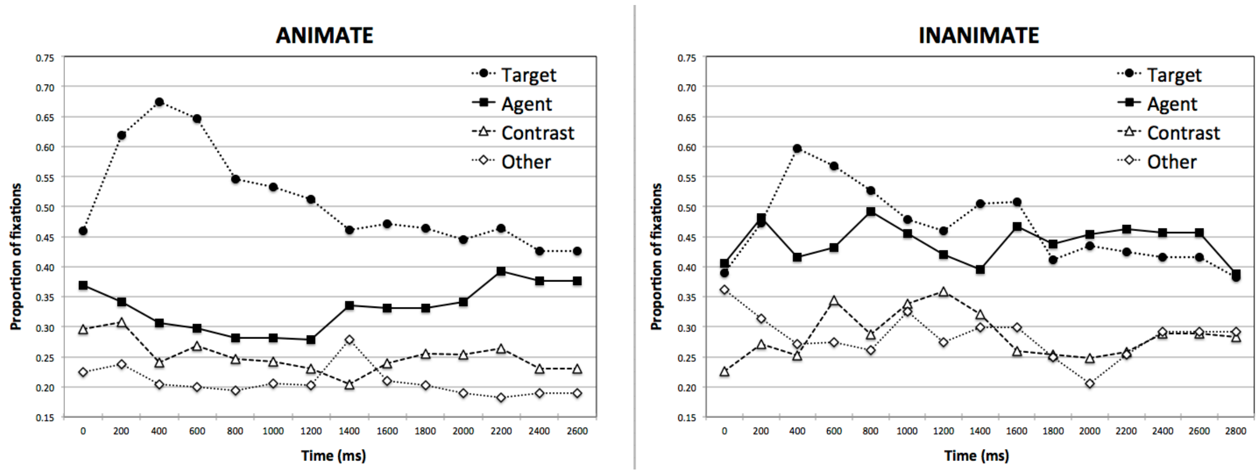


Figure 5: Fixation proportions on scene entities averaged over 500ms windows every 200ms as a function of time. Fixations are synchronized to the onset of the critical word in the questions up to the average speech onset time in each condition. Questions queried either an animate or inanimate target entity (e.g., *who is bald/what is orange?* For Figure 1, left panel). Inanimate-head actives and passives are collapsed together.

*4.2.2.1 2000ms post question critical word onset*: To examine the full fixation pattern as participants initiate utterance planning, we compared fixations on target and agent entities across animate and inanimate conditions (animate-head vs. all inanimate-head structures) averaged across the 2000ms window, although our main predictions refer to fixations on agents. As can be gleaned from Figure 5, comparisons of averaged logits across animacy conditions indicated that fixation odds for the target entity were higher for the animate than the inanimate condition ($t_1(74)= 5.73$, $p<.0001$; $t_2(37)=2.31$, $p=.03$). These differences across targets can potentially be due to size differences between animate and inanimate targets. However, as predicted, fixations on agents showed the reverse pattern: agents were more likely to be fixated in the inanimate than the animate condition, although marginally across items ($t_1(74)= 2.32$, $p=.02$; $t_2(35)=-1.73$, $p=.09$)[2]. Nevertheless, as shown in Figure 5, fixations on agents of animate target events were more likely than those on other objects such as contrasts ($t_1(34)= 4.07$, $p <.001$; $t_2(17)= 2.99$, $p =.008$) suggesting enhanced activation of the agent compared to other entities. More importantly, within the animate condition, the agent-target similarity negatively correlated with the odds of fixating the agents ($r= -.610$, $p=.004$) but there were no such correlations with the fixations on the target. Thus agents that were less similar to the animate target (e.g., dog or baby) tended to be fixated more than those agents that were more similar to the target. This suggests that similarity-based mechanisms played a role already at this early stage of planning, with reduced fixations (reflecting the level of activation) on highly similar competitors.

*4.2.2.2 1000 ms before head noun onset*: We next examined the pattern of fixations immediately preceding the head noun onset (see Figure 6). These fixations reflect the encoding of the head noun. Inanimate-head structures are collapsed together because at this point, there was no difference between them in proportion of fixations. Similar to the time window analyses presented above, animate targets were more likely to be fixated than inanimate targets ($t_1(73)=3.03$, $p=.003$, $t_2(37)=2.04$, $p=.05$), and agents were more likely to be fixated in the inanimate than the animate condition ($t_1(70)= -2.61$, $p=.01$; $t_2(33)= -2.78$, $p=.009$). As before, fixations on agents in the animate condition were nevertheless more likely

[2] The degrees of freedom may vary across different analyses because of missing values, i.e., trials in which an entity was not fixated in that particular time window.

than on other objects in the scene such as the contrast entity ($t_1(19)=3.42$, p=.004; $t_2(11)=2.81$, p=.009), suggesting some degree of activation of the agents, yet less so than agents in the inanimate condition. Critically, fixation odds for agents in the animate target condition negatively correlated with their similarity to the target (r=−.47, p=.05), suggesting that the less similar the animate entities involved in the event being talked about, the more likely the agent was fixated.

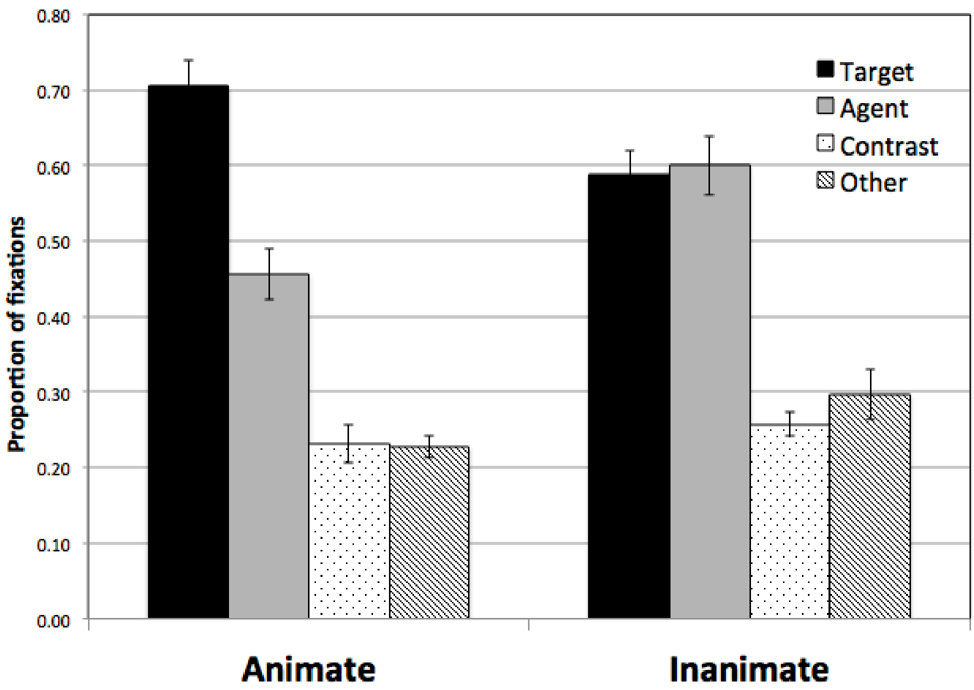


Figure 6: Average proportion of trials with fixations to scene entities within 1000ms before the onset of the head noun referring to the animate or inanimate target entity. Error bars indicate standard error.

*4.2.3 Fixations during speech*

Figure 7 shows the overall pattern of fixations for each constituent phrase during the utterance of inanimate-head and animate-head passives. For statistical analyses we only provide comparisons across passive structures, given that active structures have different word orders and by the time the first noun is uttered different upcoming words are being planned. All passives with or without *by*-phrases were included in the analyses up to the end of the verb phrase. Fixations during the *by*-phrase only represent those utterances in which there was a *by*-phrase (i.e., analyses and plots are contingent on the spoken words). Contrast entities and other entities are included in the figure for illustration (under the *Other* category), but were not statistically compared. As reported for the analyses before speech onset, Figure 7 also illustrates that during speech the relevant entities for describing the targeted event were fixated more than the irrelevant ones, particularly in the animate head condition.

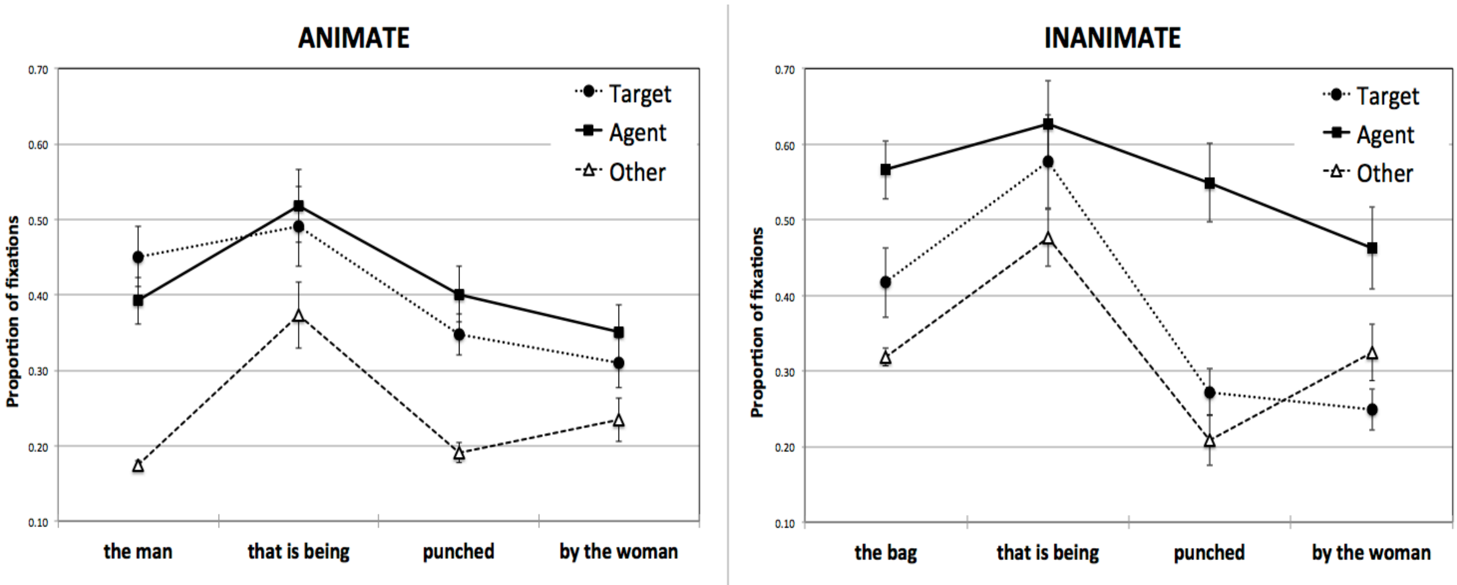


Figure 7: Average proportion of trials fixated on scene entities as critical phrases were uttered. Error bars indicate standard error.

4.2.3.1 *Noun phrase:* The average length of this window was 740ms. During the production of the determiner and the first noun, the upcoming verb phrase (with its auxiliary verb) is being planned and thus evidence of competition is expected, as reasoned in the introduction of this study. At this point in the utterance there were no differences in fixation odds across animate and inanimate targets. However, as predicted, fixations were more likely for agents of inanimate-head structures than those of animate-head structures (by-participants: t(70)=-

2.60, p=.01; by-items: t(35)= -2.07, p<.05). This suggests that in the animate-head condition, fixations on agents are fewer because agents are less accessible and show a reduced level of activation, despite being more active than other entities such as contrasting entities (agents vs. contrasts in the animate head condition: $t_1(39)= 4.62$, p<.009; $t_2(25)=3.83$; p=.001). Importantly, clear evidence indicative of semantic similarity effects comes from a significant negative correlation across items between agent fixations and agent-target similarity ($r=-.58$, p=.009), indicating that across agents of animate-head cases fewer fixations are directed to agents that are highly similar to the animate target and elicit the strongest competition (see Figure 8). No correlations were observed between similarity and target fixations. Together, these findings might suggest that the speaker experiences interference or competition when planning the upcoming verb phrase and indicating the syntactic and semantic role of the two animate nouns of the animate-head condition (the correct verb morphology). Interestingly, fixations on agents but not targets are modulated by similarity, suggesting that the target maintains its level of activation independently of its similarity to the competitor agent, but the degree to which the agent interferes with current processes does vary with similarity.

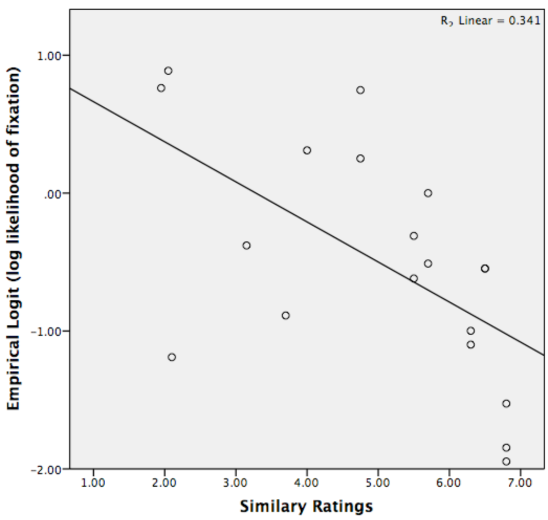


Figure 8: Correlation between the agent-target similarity ratings and the log likelihood of fixating the agent during verb encoding, i.e., while the head noun is being uttered.

4.2.3.2 *Verb phrase region*: Although our predictions do not specifically refer to this and subsequent windows, we report them here for completeness. Fixations during the utterance of the verb-phrase should reflect the planning of the upcoming *by*-phrase if any (recall all structures with and without *by*-phrases were analyzed together up to the verb). Within the region including the relative pronoun (if any) plus the auxiliary verb (e.g., *that is being, that's being, being*), a region averaging 527ms in duration, there were no differences in the likelihood of fixating the targets or the agents across inanimate and animate-head conditions. However, at the main verb region (a window averaging 450ms), the agents in the animate-head condition continued to be less likely to be fixated than in the inanimate-head condition ($t_1(156)=-2.16$, $p=.04$; $t_2(35)= -2.54$, $p<.02$). This might be due to the fact that in inanimate-head passives, the upcoming agent *by*-phrase was being planned at the time of uttering the verb (recall that 83% of inanimate-head passives were full passives). In contrast, in animate-head passives, this was less likely to be the case (63% of passives were full passives). This interpretation receives preliminary support from the comparison across items between agentless and full passives at the main verb time window, indicating higher likelihood of fixating on the agent for those main verbs that were continued with a *by*-phrase compared to agentless passives (animate-heads: $t(7)= 3.64$, $p=.008$; inanimate-heads: $t(5)=2.29$, $p=.07$), although there were very few agents fixated for the few agentless passives produced in the inanimate-head condition.

Finally, there were still fixations on agents and other characters during the production of the *by*-phrase, perhaps reflecting wrap-up effects, continuing encoding or monitoring processes to confirm or check the response. Whatever the case, while producing the *by*-phrase (a window averaging 856ms), fixations on the agent were marginally less likely in the animate-head condition compared to the inanimate-head condition ($t_1(50)=-1.83$, $p=.07$;

$t_2(33) = -2.35$, $p<.03$), possibly reflecting the fact that *by*-phrases were more likely to be produced in inanimate-head than animate-head passives, as reported above. Importantly, for animate-head passives, the likelihood of fixating the agent during this portion of the utterance negatively correlated with the similarity ratings ($r=-.63$, $p=.007$), suggesting that agents were more likely to be fixated when they were less similar to the target entity, and hence were more likely to be mentioned in a by-phrase. That is, those agents that were fixated were mentioned, and these tended to be less similar to their targets. The pattern of fixations during the verb phrase and the mentioned *by*-phrases therefore might suggest the continuing effect of interference for high-similarity agents in the animate-head condition, resulting in their omission.

### 4.3. Discussion

This study aimed to determine the time course of similarity-based competition during phrase planning and passive utterances as reflected in the eye-movement patterns. Despite all animate agents being inherently accessible, agents of inanimate targets were more likely to be fixated than those of animate targets, although the latter were nevertheless more likely to be fixated than other entities in the scene. These findings demonstrate different degrees of agents' accessibility or salience as a function of their respective contexts (animate or inanimate targets) due to interference between highly similar animate entities. Thus, when planning an animate target, speakers entertained the agent of the event to some extent in comparison to unrelated entities, but they did so to a lesser extent than the agents of inanimate targets. Importantly, the similarity between the animate entities in the relative clause being planned played a role at early stages of planning (before speech and before target onset): the more similar the two animate entities, the less likely the fixations on the agent. These results are consistent with similarity-based interference studies and suggest that attention to and selection of the target interfered more with highly similar agent entities making them less likely to be fixated and less available for mention right after the head noun, leading to passive structures.

Importantly for the main purpose of this article, we expected that if there was any comparable process associated with verbs in production as in comprehension, there should be similarity influences reflected in the likelihood of fixating agents at the point of encoding the verb phrase (during the production of the head noun phrase). At this point, fixations on agents were indeed more likely for inanimate than animate targets, as in pre-speech fixations, suggesting that agents of animate targets continued to be relatively less accessible than agents of inanimate ones, due to their reduced activation resulting from competition or interference. Critically, the accessibility of animate agents (as reflected in the fixations) was negatively correlated with their similarity to the targets thus providing evidence of similarity-based mechanisms. Specifically, within animate-head items, the likelihood of fixating the agent decreased as similarity increased, suggesting that highly competitive agents were relatively less attended to and less accessible. Assuming that this similarity modulation reflects the encoding of the upcoming verb phrase, semantic similarity must therefore play a role in indicating the semantic/syntactic roles of the nouns, because at this point the verb (*is being punched* vs. *is punching*) must be decided, which explicitly indicates whether the head noun is the patient or agent of the action. It is thus likely that the semantic features of the entities, which cue both the agent role and the subject function, elicited competition between alternative syntactic functions and verb forms. This is particularly so if one considers that the most frequent and automatic role-to-function mapping in English is from agent to subject, rather than patient to subject.

The results nevertheless suggest that semantic similarity played an early role in first noun encoding (reducing the availability of agents) and continued to do so during verb planning and beyond. As previously reported, *by*-phrase omissions were also modulated by similarity, and thus the fixations on agents during the *by*-phrase planning found here are likely to reflect the mention of the *by*-phrases. Interestingly, fixations on the animate targets were unaffected by similarity at all points in which fixations on their agents were, from question offset to *by*-phrase planning. This may suggest that maintaining the target representation in the utterance plan and indicating its function interfered with high-similarity agents and their function throughout the animate-head passive production. As a result, the incremental encoding of each word within the animate-head passives showed correlations between agent fixations and similarity. Thus, semantic interference already apparent at head noun encoding played a role at verb phrase encoding, and later determined whether the *by*-phrase was lexically encoded, given their optionality.

Overall, the results are consistent with the view that similarity-based competition occurs at the point in which entities must be integrated or linked to syntactic functions and verbs must be selected to indicate a relationship between them. This therefore parallels comprehension studies, which have indicated that difficulty in understanding object relative clauses like (2) typically manifests at verb positions, when the nouns already in working memory must receive functional and semantic roles from the input verb (Gennari & MacDonald, 2008, 2009; Gordon et al, 2001, 2006; Mak et al., 2002). Of course, many of the processes implied by the present study are task-specific. In production, highly accessible entities and their competitors, if any, may guide the process of mapping concepts into words, particularly first uttered nouns. Similarly, speakers may choose to mention elements or not (e.g., *by*-phrases), and generally, they can control the timing of their utterance to alleviate cognitive effort. Despite a role for similarity in these production processes, none of this seems to occur in comprehension. In comprehension, difficulty precisely manifests at places where sequences of unrelated words, which already increase working memory load, must be mapped onto an event meaning. Nevertheless, the fact that verb encoding and verb comprehension, i.e., the mapping of an event meaning into verb syntactic roles and vise versa, are associated with similar competitive processes suggests that similarity-based mechanisms may be shared across tasks.

## 5. General discussion

The main aim of the work presented here was to examine the relationship between production and comprehension processes, and specifically the role of semantic similarity when producing and comprehending relative clauses. The results of Study 1 have shown that semantic similarity plays a role in determining the structure of the relative clause (passive or active) and consequently the functional role of the head-noun, at least when the experimental context makes agents and active structures more available for use or mention compared to neutrally designed studies. In particular, high-similarity animate entities were mapped onto passive structures, even when alternative structures were made more available via contextual facilitation. Conversely, less similar animate entities were more susceptible to contextual facilitation, and thus more likely to be mapped onto active object structures compared to previous studies. These results were paralleled in comprehension in Study 2, which used the same picture stimuli as Study 1 and investigated descriptive phrases containing the relative clause, rather than full sentences as in previous studies. Response times indicated that active object relatives such as *the man that the woman is punching* were the most difficult to comprehend, with little difference between other structures. Critically, this effect was modulated by the similarity between the two animate entities with longer response times for more similar entities. As previously suggested, this indicates that determining a functional and semantic role for the nouns of the relative clause in comprehension involves competition, which is stronger when nouns are highly similar. Finally, Study 3 examined the time course of competition in the production of passive relatives. Because high similarity nouns tend to be mapped onto passive structures, we expected similarity-based competition to be reflected in these structures at two main points: during initial planning and first noun retrieval or selection (when targets and competitors might interfere with each other), and when the verb is planned (when the event syntactic and semantic roles are morphologically indicated as in comprehension). The results largely supported these predictions. Competitor agents of animate targets were generally fixated less than non-competitor agents of inanimate targets, but were fixated more relative to other entities throughout pre-speech planning and formulation. This suggests a reduced level of activation or accessibility for competitor agents, which were nevertheless viable candidates in the formulation process compared to unrelated entities. Importantly, the likelihood of fixating the competitor agent of animate targets—both during head-noun and verb planning—negatively correlated with similarity, indicating that more similar entities induced fewer fixations on competitors compared to less similar ones, due to their relatively lower level of accessibility.

Taken together, these findings suggest that encoding or determining the semantic and syntactic roles of animate entities elicits similarity-based competition between the animate concepts entertained during relative clause verb planning or understanding. In particular, it is the temporal overlap in entertaining these concepts in the utterance plan or working memory that results in competition (Smith & Wheeldon, 2004). Indeed, the planning of simple subject-verb-object sentences with two animate nouns does not appear to elicit competition because speakers can follow a well-rehearsed strategy and attend to the nouns one at a time (Griffin &

Bock, 2000). Thus, the relative clause structure constrains speakers to attend to the two competing nouns near simultaneously: As the first noun is uttered in an animate-head clause, the indication of its syntactic and semantic role with respect to the relative clause verb involves considering the role of the competing entity, which is also involved in the event being talked about. Therefore, more cognitive effort needs to be allocated to the process of determining the event roles, compared to simple structures.

This process clearly parallels those operating in relative clause comprehension according to memory-based accounts. Animate-head active relatives are more difficult to comprehend not simply because two nouns precede the verb—two nouns also precede the verb in inanimate-head relatives—but because the semantic roles of these nouns should be simultaneously entertained with respect to the verb. Semantic similarity makes this process difficult. Indeed, the results from Study 2 indicate that animate-head active relatives are more difficult to comprehend than inanimate-head ones. Inanimate-head clauses do not challenge comprehenders because inanimate objects are from the start assumed to play a patient-theme role (Gennari & MacDonald, 2008; Trueswell, Tanenhaus, & Garnsey, 1994). Similarly, passive relatives do not elicit similarity-based competition in comprehension because the first noun can be incrementally integrated—the passive verb form clearly specifies the role of the first noun.

These findings indicate that in both production and comprehension similarity-based competition is most likely to occur in structures and contexts in which the syntactic and semantic roles of similar animate entities must be considered close in time, suggesting common event-based role competition. The more similar the two concepts are, the more likely they are to have similar roles in an event and sentence structure. Their semantic features and associated world knowledge cue similar semantic and syntactic roles at the point a verb is processed, and therefore more processing resources are needed to distinguish between the nouns' agent and patient roles. This interpretation is consistent with many previous studies indicating that noun features, in particular agent- or patient-like features, cue the event roles that the nouns play (Ferretti & Gagne, 2006; Ferretti, McRae, & Hatherell, 2001; McRae, Ferretti, & Amyote, 1997). For example, inanimate nouns, or animate nouns with patient-like feature (as in *the young naïve gambler* followed by the verb *manipulated*…), are immediately understood as patients of an upcoming verb because they cue the appropriate patient role. In contrast, agent-like animate nouns (e.g., *the man examined by the doctor*…) are initially interpreted as agents, leading to garden-path effects (McRae et al., 1997; Trueswell et al., 1994). Similarly, the relatively equal likelihood with which two nouns fulfill the agent or the patient role of a verb (e.g., *the tiger was killed by the lion*) leads to interpretation errors in aphasics, suggesting that world knowledge associated with the nouns also cues event roles (Berndt, Mitchum, & Haendiges, 1996). Thus, conceptual knowledge activated for the nouns during reading comprehension and message planning is linked to the potential roles they can play in an event, leading to competition when noun features greatly overlap and cue the same syntactic or semantic role.

This provides an explanation as to why less similarity between the animate entities does not engender strong competition. A woman and a baby in *the baby the woman is holding*, by their semantic features and the world knowledge associated with them, bring to mind different roles, because a small baby is more likely to be the receiver of an adult's action than an agent. It is therefore likely that the relevant conceptual knowledge in semantic memory is retrieved or transiently activated both in production and comprehension when verbs are processed. Competition might even occur at the same semantic/conceptual level of linguistic representations during processing, before a specific word form has been chosen in production or after it has been recognized in comprehension. At the very least, the comparable similarity-based processes in production and comprehension that we have revealed here are consistent with this possibility, because both involve functionally determining the roles in the structure at hand.

Additional evidence for common competition mechanisms in production and comprehension comes from numerous neuropsychological and brain imaging studies suggesting that semantic competition occurs in overlapping brain regions and across a variety of tasks, including verbal comprehension and production (Jefferies & Lambon Ralph, 2006; Kerns, Cohen, Stenger, & Carter, 2004; Miller & Cohen, 2001; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001). These studies argue that similar cognitive control mechanisms are recruited when semantic processing becomes more difficult, and in particular, when prepotent automatic interpretations or plans must be inhibited according to task demands, or when interpretations or plans are under-determined due to equally likely alternatives (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Such mechanisms tend to elicit more activity in the prefrontal cortex, and Broca's area in particular (or left inferior frontal gyrus). This region is recruited by a variety of cognitively demanding tasks, including the Stroop task, syntactic ambiguity resolution, and competition in name selection (January, Trueswell, & Thompson-Schill, 2009; Novick, Trueswell, & Thompson-Schill, 2010; Spalek & Thompson-Schill, 2008). Patients with circumscribed lesions in this region also show parallel impairments across these tasks (Novick et al., 2009). Such findings therefore highlight the role of Broca's area in competition resolution in both production and comprehension (Thompson-Schill, Bedny, & Goldberg, 2005).

More relevant to the present studies, several imaging studies have directly compared the brain networks involved in both sentence production and comprehension, and their findings indicate overlapping activity in prefrontal brain areas (e.g. Menenti et al., 2011). Critically, when processing relative clauses, which increase processing load and engender competition, both production and comprehension show common prefrontal activity within Broca's area, in addition to task-specific regions (Humphreys & Gennari, 2014). Specifically, when comparing production and comprehension of relative clauses with the methods and materials used in Studies 1 and 2, semantic similarity was found to positively correlate with the activity within this region (Humphreys & Gennari, 2014; Humphreys, 2012). Together, these findings suggest that production and comprehension recruit shared neural resources, in particular when producing and comprehending relative clauses involving similarity-based competition.

### 5.1. Implications for models of production and comprehension

The behavioral and neuroimaging evidence discussed above suggest that current models of sentence production and comprehension should accommodate the commonalities observed across production and comprehension in competition resolution more generally and similarity-based competition in particular. Note that the existence of a common knowledge base (e.g. lexical and world knowledge) for language, including expectations about the roles that animate and inanimate characters may play in the event does not appear to explain the specific processes observed here. Although world knowledge and noun features cue specific syntactic and semantic roles, as argued above, such cuing in itself does not explain the processing cost, the structural choices, and the pattern of fixations observed here. Two similar animate concepts may cue the same semantic or syntactic role with relatively equal strength, and therefore, additional mechanisms must be brought into place during processing to favor one alternative over the other (competition or conflict resolution). In comprehension of object relative clauses, this entails the maintenance of the head-noun in memory when intrusive competitors are encountered, the selection of the appropriate noun or concept for the agent role, and the suppression of more frequent prepotent interpretations, e.g., that the first noun is the agent of the event. In production, this may entail the maintenance of the head noun concept in the utterance plan in the face of intrusive competitors, and the selection of the appropriate noun/concept for the patient or subject role. The way in which competition is resolved is specific to the information temporarily activated in the context and the task demands at hand, and general world knowledge predictions would typically not explain how a specific instance of competition will resolve over time. Such predictions do nonetheless influence the extent to which a particular syntactic or semantic role may be activated during planning or comprehension, and therefore, the extent to which targets will interfere with competitors, as our similarity effects suggest. Therefore, whereas a common knowledge base provides the relevant information that will be temporarily activated during processing, the specific interplay and competition between these transient activations play out *outside* the common semantic memory store at some other level of representation during processing. Some researchers may view this processing space as working memory, others may view it as a network of interactive activations.

The present results are consistent with current proposal arguing for a common role of learning and linguistic experience in production and comprehension (Chang et al., 2006; Gennari & MacDonald, 2009; Macdonald, 2013; Mirkovic & MacDonald, 2013). For example, MacDonald (2013) and Gennari and MacDonald (2009) argue that production constraints shape the form of sentences that comprehenders are exposed to and expect through learning, leading to difficulties when expectations are not met—the Production-Distribution-Comprehension account. Indeed, the tendency to avoid or alleviate competition in production leads speakers to produce passive relative clauses when two animate nouns are being talked about, so active object relatives are rarely experienced by comprehenders. Thus, when comprehenders encounter active object relatives, they expect different structures and interpretations (e.g., the first animate noun is initially understood as an agent), and experience difficulty in establishing who is doing what to whom in the event.

However, the present results go beyond a common knowledge base resulting from learning and suggest that conflicting semantic information leads to similarity-based competition in both production and comprehension during verb processing, due to similar influences of world knowledge (semantic memory) during syntactic and semantic role assignment. Competitive mechanisms have been the main topic of a large research tradition in comprehension, with emphasis in ambiguity resolution more generally and object relative clauses in particular (Gennari & MacDonald, 2008; MacDonald et al., 1994). In contrast, competitive mechanisms in sentence production have received much less attention, despite some advances in agreement production (Haskell & MacDonald, 2003; Mirković & MacDonald, 2013). The present work therefore highlights the competitive nature of sentence production processes when more complex structures and difficult planning processes are targeted. As suggested by neuroimaging research, it is indeed likely that some processes in competitive environments are shared across production and comprehension.

More generally, the current findings indicate that somewhere along the mapping from message to words in production and words to interpretation in comprehension, information processing undergoes similar competition mechanisms when conflicting cues need to be resolved. Although neuro-cognitive models propose a common brain space such as Broca's area where conflict resolution occurs across a variety of tasks (Binder & Desai, 2011; Thompson-Schill et al., 2005), current psycholinguistic models have not explored this possibility. Nevertheless, it is possible that competitive mechanisms do not take place in a single brain region but rather take place across more complex brain networks, and information flow within the network may be partially shared across various tasks. Indeed, several studies suggest more than one brain region might be involved in competition resolution (Duncan, 2010; Fuster, 2001; Gennari, MacDonald, Postle, & Seidenberg, 2007; Humphreys & Gennari, 2014). Although more research is surely needed to understand the cognitive and neural processes instantiating competition resolution, the present results suggest possible competitive mechanisms shared across tasks and new avenues for further research.

## 5.2. Conclusion

Although it is difficult to conclusively demonstrate that competition mechanisms share cognitive architecture across different tasks, the present findings suggest that similar competitive mechanisms operate in both production and comprehension at points in which event roles are considered nearly simultaneously for alternative nouns, i.e., during verb encoding or interpretation. Moreover, current neuroimaging evidence suggests that overlapping neural resources are recruited to resolve such competition. This possibility does not preclude the existence of task-specific processes but highlights processing and architectural commonalities that deserve further study.

## References


Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Linguistics and Language Compass*, *4*(4), 187–203. http://doi.org/10.1111/j.1749-818X.2010.00193.x

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*, 521–536.

Baddeley, A. (1986). *Working Memory*. Oxford, UK: Clarendon Press.

Baddeley, A., & Hitch, G. (1974). Working Memory. *Psychology of Learning and Motivation*, *8*, 47–89.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*(13), 2883–2901. http://doi.org/10.1016/j.neuropsychologia.2007.06.015

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*, 457–474. http://doi.org/10.1016/j.jml.2007.09.002

Berndt, R. S., Mitchum C. C., & Haendiges, A. N. (1996). Comprehension of reversible sentences in "agrammatism" a meta-analysis. *Cognition*, *58*, 289–308.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527–536. http://doi.org/10.1016/j.tics.2011.10.001

Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *12*, 575–586.

Bock, J. K. (1987). An effect of the accessibility of word forms on sentence structure. *Journal of Memory and Language*, *26*, 119–137.

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. Academic Press.

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., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, *21*, 47–67.

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*(3), 437–458.

Boersma, P. & Weenink, D. (2015). Praat: doing phonetics by computer. Retrieved from from http://www.praat.org/

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict Monitoring and Cognitive Control. *Psychological Review*. http://doi.org/10.1037//0033-

295X.108.3.624

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming Syntactic. *Psychological Review*, *113*(2), 234–272.

Clark, H. H. (1965). Some structural properties of simple active and passive sentences. *Journal of Verbal Learning & Verbal Behavior*, *4*(5), 365–370. http://doi.org/http://dx.doi.org/10.1016/S0022-5371%2865%2980073-1

Conway, a R., & Engle, R. W. (1994). Working memory and retrieval: a resource-dependent inhibition model. *Journal of Experimental Psychology. General*, *123*(4), 354–373. http://doi.org/10.1037/0096-3445.123.4.354

Costa, A., Alario, F. X., & Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bulletin & Review*, *12*(1), 125–131.

De Smedt, K. J., & Kempen, G. (1987). Incremental sentence production, self-correction and coordination. In G. Kempen (Ed.), *Natural language generation: New results in artificual Intelligence, psychology and linguistics* (pp. 365–376). Dordrecht: Kluwer.

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172–179. http://doi.org/10.1016/j.tics.2010.01.004

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language*, *54*, 541–553. http://doi.org/10.1016/j.jml.2005.12.006

Ferretti, T. R., & Gagne, C. L. (2006). The recovery of thematic role structure during noun-noun interpretation. *Psychonomic Bulletin & Review*, *13*(3), 423–428.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role. *Journal of Memory & Language*, *44*(4), 516–547.

Fox, B. A., & Thompson, S. A. (1990). A discourse exaplanation of the grammar of relative clauses in English conversation. *Language*, *66*(2), 297–316.

Freedman, M. L., Martin, R. C., & Biegler, K. (2004). Semantic relatedness effects in conjoined noun phrase production: Implications for the role of short-term memory. *Cognitive Neuropsychology*, *21*, 245–265.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of Experimental Psychology. General*, *133*(1), 101–135. http://doi.org/10.1037/0096-3445.133.1.101

Fukumura, K., & van Gompel, R. (2010). The effects of animacy in the choice of referring expressions. *Language and Cognitive Processes*, (June 2015), 37–41. http://doi.org/10.1080/01690965.2010.506444

Fukumura, K., van Gompel, R. P. G., Harley, T., & Pickering, M. J. (2011). How does similarity-based interference affect the choice of referring expression? *Journal of Memory and Language*, *65*(3), 331–344.

Fuster, J. M. (2001). The prefrontal cortex--an update: time is of the essence. *Neuron*, *30*(2), 319–333.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory & Language*, *58*, 161–187.

Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, *111*(1), 1–23. http://doi.org/http://dx.doi.org/10.1016/j.cognition.2008.12.006

Gennari, S. P., MacDonald, M. C., Postle, B. R., & Seidenberg, M. S. (2007). Context-dependent interpretation of words: evidence for interactive neural processes. *Neuroimage*, *35*(3), 1278–1286.

Gennari, S. P., Mirkovic, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: a cross-linguistic investigation. *Cognitive Psychology*, *65*, 141–176.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76.

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

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, *118*(13-42).

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory Interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cogntion*, *27*(6), 1411–1423.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*(1), 97–114.

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *32*(6), 1304–21. http://doi.org/10.1037/0278-7393.32.6.1304

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science : A Journal of the American Psychological Society / APS*, *11*, 274–279. http://doi.org/10.1111/1467-9280.00255

Haskell, T. R., & MacDonald, M. C. (2003). Conflicting cues and competition in subject–verb agreement. *Journal of Memory & Language*, *48*(4), 760–778.

Huettig, F., & Hartsuiker, R. J. (2008). When you name the pizza you look at the coin and the bread: eye movements reveal semantic activation during word production. *Memory & Cognition*, *36*(2), 341–360. http://doi.org/10.3758/MC.36.2.341

Humphreys, G. F. (2012). *Linking sentence production and comprehension: the neural mechanisms underlying production and comprehension control processes*. University of York.

Humphreys, G. F., & Gennari, S. P. (2014). Competitive mechanisms in sentence processing: common and distinct production and reading comprehension networks linked to the prefrontal cortex. *NeuroImage*, *84*(0), 354–66. http://doi.org/10.1016/j.neuroimage.2013.08.059

January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of stroop and syntactic ambiguity resolution in Broca's area: implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience*, *21*(12), 2434–2444. http://doi.org/10.1162/jocn.2008.21179

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain* , *129* (8 ), 2132–2147. http://doi.org/10.1093/brain/awl153

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156. http://doi.org/http://dx.doi.org/10.1016/S0749-596X%2803%2900023-8

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science: A Multidisciplinary Journal*, *11*, 201–258.

Kerns, J. G., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2004). Prefrontal cortex guides context-appropriate responding during language production. *Neuron*, *43*(2), 283–291.

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, *73*, 1–40. http://doi.org/10.1016/j.cogpsych.2014.04.001

Levelt, W. J. M. (1989). *Speaking: From intention to articulation. ACL-MIT Press series in natural-language processing*. Cambridge, MA: The MIT Press.

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences*, *6*(7), 299–305. http://doi.org/10.1016/S1364-6613(02)01923-X

Lewis, R. L., & Vasishth, S. (2005). An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval. *Cognitive Science*, *29*(3), 375–419.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454.

http://doi.org/http://dx.doi.org/10.1016/j.tics.2006.08.007

Macdonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*(April), 226. http://doi.org/10.3389/fpsyg.2013.00226

MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). Lexical Nature of Syntactic Ambiguity Resolution. *Psychological Review*, *101*(4), 676–703.

Mak, W. M., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, *54*(4), 466–490.

McDonald, J. L., Bock, J. K., & Kelly, M. H. (1993). Word and world order: Semantics, phonological, and metrical determinants of serial position. *Cognitive Psychology*, *25*, 188–230.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic Roles as Verb-specific Concepts. *Language and Cognitive Processes*, *12*(2/3), 137–176.

Menenti, L., Gierhan, S. M. E., Segaert, K., & Hagoort, P. (2011). Shared Language: Overlap and Segregation of the Neuronal Infrastructure for Speaking and Listening Revealed by Functional MRI . *Psychological Science* , *22* (9 ), 1173–1182. http://doi.org/10.1177/0956797611418347

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Mirković, J., & MacDonald, M. C. (2013). When singular and plural are both grammatical: Semantic and morphophonological effects in agreement. *Journal of Memory & Language*, *69*(3), 277–298.

Montag, J. L., & MacDonald, M. C. (2014). Visual Salience Modulates Structure Choice in Relative Clause Production. *Language and Speech*, *57*, 163–180. http://doi.org/10.1177/0023830913495656

Novick, J. M., Kan, I. P., Trueswell, J. C., Thompson-Schill, S. L., Novick, J. M., Kan, I. P., … Thompson-Schill, S. L. (2009). A case for conflict across multiple domains: memory and language impairments following damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology*, *26*(6), 527–567.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca' s Area and Language Processing: Evidence for the Cognitive Control Connection. *Language and Linguistics Compass*, *4*, 906–924. http://doi.org/10.1111/j.1749-818X.2010.00244.x

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

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*(3), 348–379.

http://doi.org/http://dx.doi.org/10.1016/j.jml.2007.03.002

Roland, D., Mauner, G., O'Meara, C., & Yun, H. (2012). Discourse expectations and relative clause processing. *Journal of Memory and Language*, *66*(3), 479–508. http://doi.org/10.1016/j.jml.2011.12.004

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, *29*, 86–102.

Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1503–1514. http://doi.org/10.1037/a0024350

Smith, M., & Wheeldon, L. (2004). Horizontal Information Flow in Spoken Sentence Production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 675–686. http://doi.org/http://dx.doi.org/10.1037/0278-7393.30.3.675

Spalek, K., & Thompson-Schill, S. L. (2008). Task-dependent semantic interference in language production: an fMRI study. *Brain & Language*, *107*(3), 220–228. http://doi.org/10.1016/j.bandl.2008.05.005

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*(1), 71–86. http://doi.org/10.1016/j.cognition.2010.04.002

Thompson-Schill, S. L., Bedny, M., & Goldberg, R. F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology*, *15*(2), 219–224.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(26), 14792–14797.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, *47*(1), 69–90.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic Influences on Parsing: Use of Thematic Role Information in Syntactic Ambiguity Resolution. *Journal of Memory and Language*, *33*, 285–318.

Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *33*(2), 330–407.

Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition*, *131*(3), 373–403. http://doi.org/10.1016/j.cognition.2014.01.007

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, *55*(2), 157–166. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=psyc5&AN=2006-09177-001

Van Dyke, J. A., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, *65*(3), 247–263. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=psyc7&AN=2011-19529-004

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), B61–B69. http://doi.org/http://dx.doi.org/10.1016/S0010-0277%2802%2900107-5

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*(2), 329–338.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*(1), 79–112.

Wolter, L., Gorman, K. S., & Tanenhaus, M. K. (2011). Scalar reference, contrast and discourse: Separating effects of linguistic discourse from availability of the referent. *Journal of Memory and Language*, *65*(3), 299–317. http://doi.org/10.1016/j.jml.2011.04.010

Acknowledgements: we thank Xierong Liu and Sinead Murphy for help with the eye-tracking data analysis. Moreover, we thank the reviewers of this article, in particular, Zenzi Griffin, for extremely valuable comments.