Northumbria University
NEWCASTLE

Attention and memory play different roles in syntactic choice during sentence production

Andriy Myachykov, Simon Garrod, and Christoph Scheepers

Author note

Andriy Myachykov, Department of Psychology, Northumbria University, Newcastle-upon-Tyne. Centre for Cognition and Decision Making, National Research University Higher School of Economics, Russian Federation.

Simon Garrod, Institute of Neuroscience and Psychology, University of Glasgow.

Christoph Scheepers, Institute of Neuroscience and Psychology, University of Glasgow.

Correspondence concerning this article should be addressed to Andriy Myachykov, Department of Psychology, Northumbria University, Northumberland Building, Newcastle upon Tyne, NE1 8ST, United Kingdom, Tel.: +44-191-227-31-58, Fax: +44-191-227-45-15, e-mail: andriy.myachykov@northumbria.ac.uk

Acknowledgements: This work has been supported by the Russian Academic Excellence Project '5-100'

Word count: 4500 words, excluding abstract, figures, and tables.

Abstract

Attentional control of referential information is an important contributor to the structure of discourse (Sanford, 2001; Sanford & Garrod, 1981). We investigated how attention and memory interplay during visually situated sentence production. We manipulated speakers' attention to the agent or the patient of a described event by means of a referential or a dot visual cue (Posner, 1980). We also manipulated whether the cue was implicit or explicit by varying its duration (70 ms versus 700 ms). Participants used passive voice more often when their attention was directed to the patient's location, regardless of whether the cue duration. This effect was stronger when the cue was *explicit* rather than *implicit*, especially for passive-voice sentences. Analysis of sentence onset latencies showed a divergent pattern: Latencies were shorter (1) when the agent was cued, (2) when the cue was explicit and (3) when the (explicit) cue was referential; (1) and (2) indicate facilitated sentence planning when the cue supports a canonical (active voice) sentence frame and when speakers had more time to plan their sentences; (3) suggests that sentence planning was sensitive to whether the cue was informative with regard to the cued referent. We propose that differences between production likelihoods and production latencies indicate distinct contributions from attentional focus and memorial activation to sentence planning: While the former partly predicts syntactic choice, the latter facilitates syntactic assembly (i.e., initiating overt sentence generation).

231 words

Keywords: attention, memory, priming, sentence production, syntactic choice

Attention, Memory, and Syntactic Choice in Sentence Production

Classical psycholinguistic studies of discourse (e.g., Sanford and Garrod, 1981) show how interlocutors' linguistic choices reveal referents' activation status resulting from constant shifts of attention between these referents (also, Sanford, Moxey, & Paterson, 1996). Hence, for linguistic interaction to succeed speakers' and listeners' need to maintain adequate attention to and memory for, discourse relevant information. In this paper, we will address the questions of how (1) the speaker's changing attentional state and (2) the ability to access semantic information about one of two interacting referents prior to describing an event affect the choice of syntactic alternatives and the time course of overt sentence production.

Evidence that the speaker's attention may play a special role in syntactic choice during sentence production comes primarily from research using variants of the *perceptual priming* paradigm: Speakers describe visual events while their attention is manipulated toward one of the event's referents. Visual cues used in these studies differ but they are typically unrelated to (and, therefore, uninformative about) the cued referent (Posner, 1980). Arguably, the first demonstration of attentional effects on syntactic choice is Tomlin's Fish Film study (Tomlin, 1995). In this study, English speakers reliably selected between active and passive voice frames after their visual attention was directed (via an explicit arrow cue) to either the agent or the patient of the event to be described. Subsequent research provided further evidence about the role of visual attention in syntactic choice (Tomlin & Myachykov, 2015, for a recent review). A recurrent finding is that explicit (Myachykov, Garrod, & Scheepers, 2012) and implicit (Gleitman, et al., 2007; Myachykov, Garrod, & Scheepers, 2011) cueing of visual attention predicts the choice of syntactic structure in sentence production by assigning a prominent grammatical role (e.g., subject) to the cued referent. Importantly, this coupling between visual attention and grammatical processing during production is not limited to the choice of global syntactic frames or to grammatical subject assignment alone. A similar link has been demonstrated for other grammatical phenomena, such as grammatical gender (Caffarra, Pesciarelli, Cacciari, 2013), relative clause production (Montag & MacDonald, 2013), and compound sentences (Coco & Keller, 2015). Moreover, syntactic choice has been shown to depend on attentional orienting for both monolingual and bilingual speakers (Antón-Méndez, Gerfen, & Ramos, 2015). Finally, a number of neuroimaging studies provide direct evidence of cross-talk between brain circuits involved in visual attention and sentence production (Cai, Van der Haegen, Brysbaert, 2013; Iwabuchi, et al., 2013).

Previous evidence for the influence of attention on syntactic choice came from studies using a *referential priming* paradigm. These studies typically use a *referent preview* instead of a referentially uninformative visual cue to bias syntactic choice. Such referential priming studies also demonstrate that the previewed referent is more likely to appear as the subject of a subsequent sentence, thereby triggering the corresponding syntactic choice (e.g., Clark & Chase, 1972; Prentice 1967; Turner & Rommetveit, 1968). One of the earliest studies using the referential priming paradigm (Prentice, 1967) was motivated by an even earlier finding (Carroll, 1958) that the likelihood of using active versus passive voice descriptions of transitive events varied as a function of an interrogative cue – a question about the agent or the patient of the event. Prentice used a set of cartoon pictures portraying simple transitive interactions between two characters (e.g., *fireman kicking cat*). The event pictures were presented after cue slides depicting one of the event's characters: The agent or the patient. The participant then had to describe the event. He found that speakers tended to make the pre-viewed referent the subject of their descriptions. Arguably, referent preview produced an attentional orienting bias toward the primed referent, but it also could have increased the referent's *conceptual accessibility* or the ease of retrieving semantic and lexical information about the primed referent from working memory (Bock & Warren, 1985). Subsequent research suggests that conceptual accessibility indeed contributes to syntactic choice – for example, by modulating the strength of *syntactic priming* effect (Bock, Loebel, & Morrey, 1992; Gámez & Vasilyeva, 2015; Gennari, et al., 2012; Griffin & Weinstein-Tull, 2003).

Although attentional focus and working memory activation are closely related (e.g., Erickson & Kintsch, 1995; Kastner & Ungerlieder, 2000), they are nevertheless distinct cognitive processes supported by distinct neuroanatomical networks. Attentional and memorial contributions to sentence comprehension were also shown to demonstrate a certain degree of independence in psycholinguistic research (e.g., Foraker & McElree, 2007; McElree, 2001). Hence, at least theoretically, both can contribute to syntactic planning – the former by prioritizing the cued referent for attentional processing and the latter through preferential access to information about a referent in memory (conceptual accessibility). These two processes can also interact: The discourse-related information that is continuously maintained in attentional focus does not need reactivation in working memory, whereas the memory retrieval involves more active deployment of attention (e.g., Givon, 1992). Some existing data support this claim. For example, a study by Arnold & Griffin (2007) examined how speakers choose between pronouns and proper names during sentence production. The results of this study demonstrated that the probability of selecting a pronoun over a proper name depends on both conceptual accessibility of the target referent and whether or not it is in the focus of attention. Although not immediately targeting syntactic choice, this finding suggests that focus of attention and memory access interact to influence linguistic choices during sentence production.

However, this proposal is somewhat controversial. In fact, the “mainstream” production models assign a relatively minor role to what they label as “conceptual” (or message-level) effects on the choice of structure in sentence production arguing instead that, for the most part, sentence planning follows linguistic cues, available to the speaker primarily during *lemma* selection and grammatical encoding (e.g., Bock & Ferreira, 2014). These stages arguably specify the lexical and morphosyntactic parameters of retrieved words and their ordering in the unfolding syntactic structure. Although there are few studies that directly examine the role of attention and memory in syntactic choice, a recent *explicit* priming study (Myachykov et al., 2012) set out to distinguish the effect of conceptual accessibility and visual attention to syntactic choice by using both perceptual and referential priming manipulations. The perceptual cue in this study was a pointer to the subsequently presented referent's location while the referential cue was a picture of a cued referent (i.e., referent preview) presented in the same location. Crucially, while both cues directed the speaker's attention to the referent's location, only the latter cue increased the referent's conceptual accessibility. The study replicated previously reported perceptual priming effects (the tendency of the cued referents to be selected as the sentential subjects, triggering the associated syntactic choice), but found no interaction between cue type (perceptual vs referential) and cue location (agent vs patient). This pattern suggests that visual attention affects syntactic choice independent of access to information about referents and that conceptual accessibility alone (e.g., without accompanying syntactic priming manipulations) plays little role in this process. A recent study by Vogels, Krahmer, & Maes, (2013) corroborates this view by showing that visual salience influences the global interpretation of the described scene but not the referents' accessibility status.

Together the research reviewed above demonstrates that (1) visual attention predicts syntactic choice independent of conceptual accessibility and (2) that a similar role cannot be unequivocally attributed to conceptual accessibility, as the evidence about its exact role in syntactic choice is still controversial. At the same time, research in object naming shows convincingly that speakers benefit from increased conceptual accessibility by demonstrating faster production times (Schotter, 2011, for a review) and that this preview benefit can be achieved after very short previews (Dobel, et al., 2007; Hafri, Papafragou, & Trueswell, 2013) and is independent of preview durations (Schotter, et al., 2014).

This suggests the following possibility: Whereas visual attention is important for grammatical role assignment and/or selection of syntactic frames, a referent's activation in working memory (conceptual accessibility) facilitates naming that referent during overt sentence production. This assumption predicts differential effects of attention and memory on syntactic choice (i.e., the probability of selecting one syntactic frame over the other) and on the timing of sentence production (i.e., the speed with which the sentence is overtly produced). Below, we report the results of a study where we tested this hypothesis by analysing the effects of implicit and explicit visual and referential cues on (1) the choice between active and passive frames and (2) the time course of sentence generation. First, we aimed to replicate findings from Myachykov, et al., (2012) and register a comparable effect on syntactic choice from both referential and perceptual cues regardless of cue duration. Second, we predicted that while the Cue Type would not affect syntactic choice, availability of referent preview would lead to faster sentence onsets, possibly in both explicit and implicit cue conditions.

Experiment

Design

Three factors were independently manipulated at two levels each: (1) Cue Location (Agent/Patient), (2) Cue Type (Referent/Dot), and (3) Cue Duration (70ms/700ms). The specific contrast between the two cue durations was chosen based on previous studies on control of visual attention (e.g., Gibson & Bryant, 2005; Hommel, et al., 2001). All manipulations except Cue Type were within-subjects and within-items. Each target picture appeared both in Cue-Dot and the Cue-Referent condition. This was to ensure variation of syntactic choice as a function of both referentially informative and referentially uninformative cues when describing the same event. The dependent variables were (1) the proportion of Passive Voice sentences produced and (2) sentence onset latencies.

Participants

Twenty-four native English speakers (Glasgow University undergraduates; 14 female) with normal or corrected-to-normal vision took part. They either received course credits or £6 subject payment. The mean age of participants was 19.8 years.

Materials

The target pictures consisted of 24 black-and-white line drawings portraying simple transitive events (see example in Figure 1). These were selected from the Edinburgh picture set previously used in numerous studies on sentence production (e.g., Branigan, Pickering, Cleland, 2000; Pickering & Branigan, 1998). Experimental materials employed four different event types (*pull*, *punch*, *push, and shoot*). There were 12 interacting referents rotated across these four events (*boxer, chef, clown, cowboy, doctor, monk, nun, policeman, soldier, swimmer, teacher,* and *waitress*). These referents were rotated across 24 individual items in such a way that each referent was equally likely to be either agent or patient of the event, appear on the left or on the right, and co-appear in the same event. In addition, the materials were controlled for size, animacy, and colour.

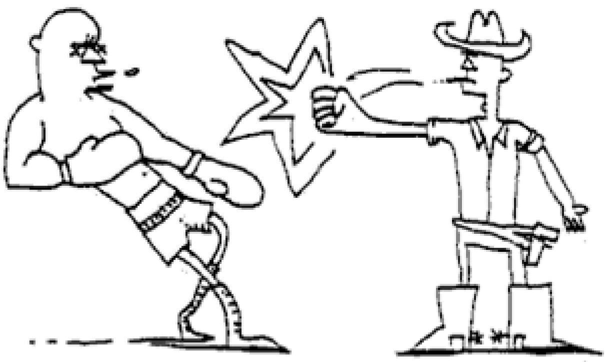

Figure 1. Target event example.

In addition, each participant engaged in a practice session before the beginning of the experiment. Since the images and their names were not controlled for general recognisability or corpus frequency, it was deemed important that participants would familiarise themselves with all the characters and events they would encounter in the main experiment. The practise session consisted of 8 trials in pseudo-random order – two per each event type (see Procedure for further details).

Finally, we included a set of 98 filler pictures of intransitive events (e.g., *a skier sleeping*). The experimental trials were pseudo-randomized individually with four fillers at the beginning of each session and each target trial preceded by two filler trials.

Apparatus

The experiment was implemented in *SR-Research Experiment Builder*. An EyeLink II head-mounted eye-tracker monitored participants' eye movements in order to ensure the efficiency of the cueing manipulation.[1]. In approximately 97% of the *explicit* (700 ms duration) and

[1] In order to analyse initial fixations on visually cued versus non-cued referents, the pictures were pre-coded to include separate areas of interest: one for each referent (agent and patient)

92% of the *implicit* (70 ms duration) Cue trials, presenting the cue led to the execution of a saccade to the cued location. The small number of trials where participants did not respond to the cue was excluded from analysis. The experimental materials were presented on a 17" CRT monitor of a DELL Optiplex GX 270 desktop computer running at a display refresh rate of 75 Hertz. Also connected to the PC was a pair of stereo speakers. A SONY DAT recorder was used for speech recording. The audio clips were later uploaded onto a PC and analysed with the help of Adobe Audition 2.0.

Procedure

Participants were positioned approximately 60 centimetres from the monitor. Viewing was binocular, but only the participant's right eye was tracked. Before the main experimental session, each participant was run through a practice session during which they saw the pictures of the referent characters that would later be presented in the target trials and sample pictures of both the target and the filler materials. The referents appeared one at a time in the centre of the screen simultaneously with their names. Participants were instructed to read out the referent names and to remember them for the following tasks. Also, each participant had to describe eight sample event pictures (one for each event type) during the practice session. The pictures of the events were presented in the middle of the screen. Practice events included different referents from those used later in the experimental session but the same type of events. No specific instruction as to how to describe these event pictures was given to

---

and one for the background. The referent areas included the referent itself plus a surrounding area of about two degrees of visual angle. Both Dot and Referent cueing manipulations were highly effective in attracting initial visual attention to the cued location.

participants, except that participants should always make reference to the event’s name and both interacting characters.

The instruction to the participants was to describe a picture as soon as they could and in a single sentence. Participants were unaware of the nature of the experimental manipulations, any difference between target and filler trials, or the exact purpose of the study. They were told that the study was concerned with speaking about what they see on the computer screen. Figures 2 and 3 illustrate typical target trial sequences in Dot-Cue and Referent-Cue conditions, respectively.

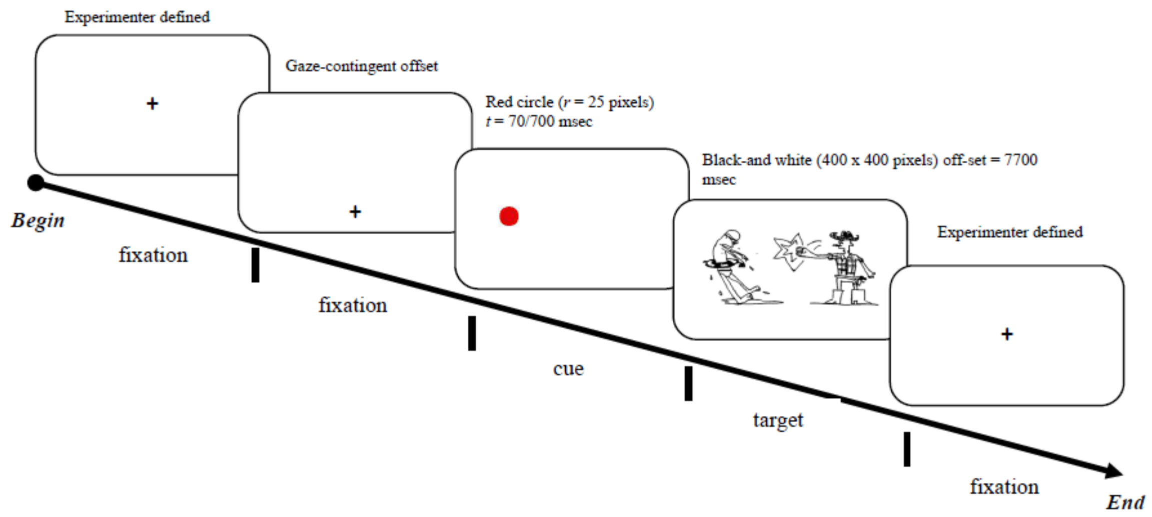

Figure 2. Example of a Dot-Cue trial.

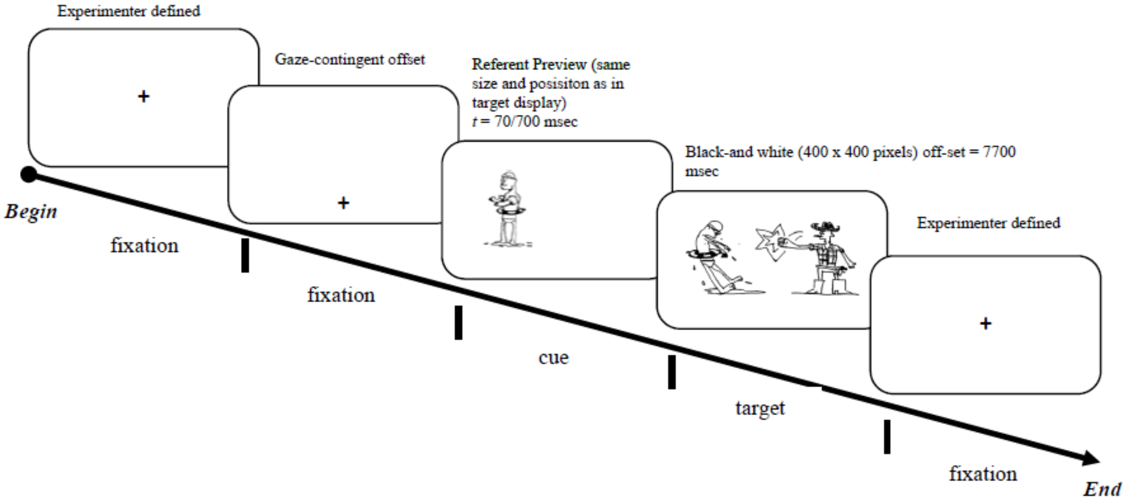

Figure 3. Example of a Referent-Cue trial.

Each target trial began with the presentation of the central fixation cross. Once a participant successfully located and fixated the central fixation cross it was replaced by a dislocated fixation cross ensuring that participants would not be looking at the centre of the screen at the time of cue presentation and that they would always have to make a saccade to the cued location or, if the cue was overlooked, to another location in the target picture once it appeared on the screen. The dislocated fixation cross was equidistant from the cued locations. The presentation of the cue was contingent on fixating the dislocated fixation cross for a minimum of 200msec, after which either a Dot Cue or a Referent Cue screen was displayed. The Dot Cue was a red circle (25 pixels in diameter) displayed in the approximate centre of one of the subsequently presented referents (agent or patient). The Referent Cue was operationalized via previewing one of the event referents (agent or patient) prior to the target display presentation. The previewed referent always appeared in neutral posture, preventing any thematic role (agent or patient) suggestibility. Referent Cues also appeared in the locations corresponding to their locations in the subsequently presented target displays. Crucially, Dot Cues only provided location information whereas Referent Cues provided both location and identity information about the subsequently presented referents. Cue duration was either 70 ms or 700 ms regardless of the type of cue, ensuring either implicit or explicit cue apprehension (Posner, 1980). Participants were instructed to locate the cue as soon as they could, direct their gaze to it, and wait for the target event. Then, the target picture appeared on the screen accompanied by an auditory marker – a single 200 ms 44.1 kHz beep with a sharp and clearly defined onset (the auditory marker allowed us to synchronise the sound recordings per trial). Participants were instructed to ignore the beep and describe the target picture in a single sentence, then press the space bar to move on to the next trial. In case the participant did not respond, the picture disappeared from the screen after 7700msec. Filler trials employed a comparable presentation sequence: The trial would begin with a central fixation mark, after which a dislocated fixation mark appeared, followed by a visual cue (identical to the procedure in the target trials), and finally, the presentation of the target display.

Results

Target Structure

Target responses were coded by a condition-blind coder as Active Voice, Passive Voice, or Other. To be coded as Active Voice, the description had to employ a transitive verb referring to the depicted event, a subject NP referring to the agent, and a direct object NP referring to the patient (e.g. *The cowboy is punching/punched the boxer*). To be coded as Passive Voice, the description had to employ a passivized transitive verb referring to the depicted event, a subject NP referring to the patient, and a by-phrase referring to the agent (e.g. *The boxer is/was* [*being*] *punched by the cowboy*). Truncated passives (not including a by-phrase) were rare since they were explicitly discouraged during the practise session. All remaining responses (including missing responses) were coded as Other (e.g. *The boxer and the cowboy are fighting*). The latter accounted for less than 2% of the data and will not be considered further.

Our first set of analyses focused on proportions of Passive Voice descriptions. Statistical analyses were performed using *Generalized Linear Mixed Effects Models* (GLMM), which are part of the lme4 package in R. Since the dependent variable of interest was binary (Passive Voice description: True or False), a binary logistic model (binomial(logit)) was specified in the family argument of the glmer() function. The model included a full-factorial cue_location * cue_type * cue_duration fixed effects design (all predictors were centered around the grand mean using deviation-coding), as well as various crossed random effects to allow for simultaneous generalization of fixed effects results across participants and items. Following the recommendations in Barr, Levy, Scheepers, and Tily (2013), the maximal random effects structure justified by the design was used. Since all experimental variables were both within-subjects and within-items, the model therefore not only included by-subject and by-item random *intercepts*, but also by-subject and by-item random *slopes* for every main effect and higher-order interaction in the fixed effects design. The only deviation from the 'standard' maximal approach was that we excluded any random correlations from the model (inclusion of random correlation terms caused severe convergence problems). The simulations in Barr et al. (2013) suggested that models without random correlations are unlikely to increase the risk of a Type I error. P-values were obtained via *Likelihood Ratio* model comparisons rather than the "*t* as *z*" approach (the former tends to be less anticonservative than the latter, cf. Barr et al., 2013). The results are summarized in Table 1.

Effect	$_{LR}\chi^2(1)$	*P*
Cue-Location (L)	11.942	.001
Cue-Type (T)	0.240	.624
Cue-Duration (D)	0.208	.649
L x T Interaction	1.575	.210
L x D Interaction	3.557	.059
T x D Interaction	0.210	.647
L x T x D Interaction	0.502	.479

Table 1. Results from binary logistic GLMM analyses modelling proportions of passive voice responses as a function of Cue Location (L), Cue Type (T), and Cue Duration (D). The reported test statistic is *Likelihood Ratio Chi-Square*.

The intercept of the model was estimated as −3.112 ± 0.313 (μ ± SE) logit units which is substantially smaller than zero (i.e., smaller than 0.5 in probability space). This is because passive-voice responses were greatly outnumbered by active-voice responses overall. This finding is in line with existing corpus-analysis data (e.g., Roland, Dick, & Elman, 2007; Svartvik, 1966) as well as previous findings using visual cueing and referent preview paradigms (Myachykov, et al., 2012). There was a reliable main effect of Cue Location, with more passive-voice responses in the Patient-Cued than in the Agent-Cued condition, replicating previous findings whereby attentionally focused referents tend to appear in

Subject position in an English transitive sentence (Tomlin & Myachykov, 2015, for a recent review). Crucially, there was neither a main effect of Cue Type, nor or any interaction between Cue Type and other factors. There was, however, a marginal interaction between Cue Location and Cue Duration (see Figure 4): In the *patient-cued condition*, participants were about 8% more likely to produce passives when the cue was explicit (700 ms) rather than implicit (70 ms), and the corresponding Cue Duration simple effect (at Cue Location = Patient) was indeed significant[2]: ${}_{LR}\chi^2(1) = 7.083$; $P = .008$; by contrast, in the *agent-cued condition*, the Cue Duration contrast pointed in the opposite direction and was not reliable: ${}_{LR}\chi^2(1) = 1.520$; $P = .218$.

Together these findings suggest (1) that attention affects the selection of a syntactic frame independently of a referent's conceptual accessibility (cf. Myachykov, et al., 2012) and (2) that the strength of attentional cueing (i.e. explicit vs implicit) matters more if it encourages a normally less preferred syntactic frame: Cue duration mattered more when the cue was on the patient, which supported the activation of the passive frame.

[2] Simple effects were determined via dummy-coding of a reference predictor. For example, when Cue Location is dummy-coded as 1=Agent, 0=Patient, the resulting 'main effect of Cue Duration' in the model actually refers to the *simple effect of Cue Duratio*n given Cue Location = 0 (Patient).

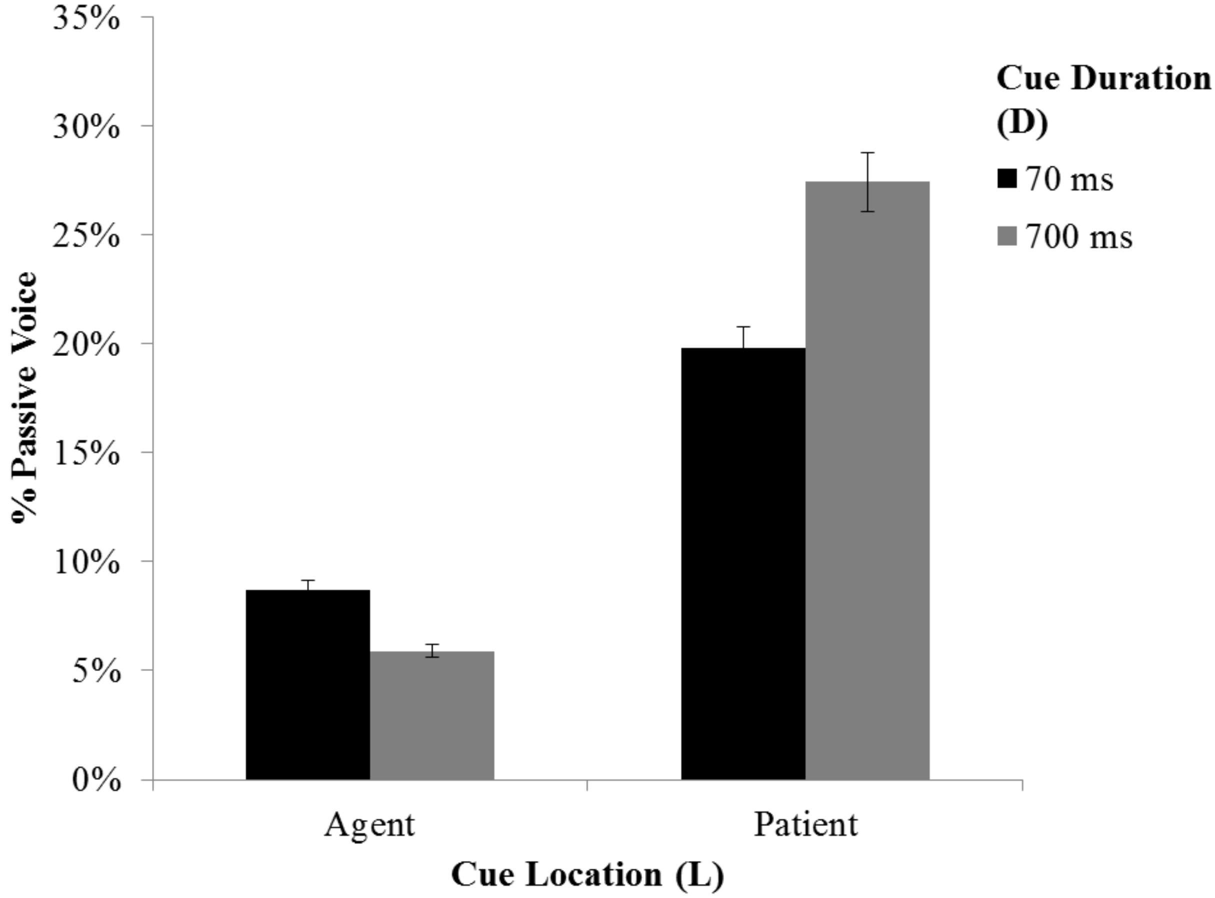

Figure 4. Mean percentages (error bars represent standard errors) showing the Cue Location x Cue Duration Interaction on proportions of Passive Voice responses.

Sentence Onset Latencies

We also performed analyses of sentence onset latencies (SOL). For this purpose, participants' verbal responses were coded as follows. First, we identified the beep onset corresponding to the onset of the target event per trial. Next, we identified the sentence onset corresponding to the onset of the Subject NP. Potential hesitations and filled pauses prior to the onset of the Subject NP were coded as parts of SOLs.

The analyses were based on GLMMs that were identical in all respects to those reported in the previous section, except that the family argument was now set to its default gaussian(identity)specification, since we were dealing with continuous instead of binary data here. Again, random correlations were excluded to avoid convergence problems, and p-values were obtained from *Likelihood Ratio* model comparisons.

The corresponding inferential results are presented in Table 2, and Figure 5 illustrates the only interaction we found.

Effect	$_{LR}\chi^2(1)$	*P*
Cue-Location (L)	18.027	.001
Cue-Type (T)	0.052	.820
Cue-Duration (D)	17.825	.001
L x T Interaction	0.397	.529
L x D Interaction	1.672	.196
T x D Interaction	10.712	.001
L x T x D Interaction	0.007	.932

Table 2. Results from Linear Mixed Model analyses on sentence onset latencies as a function of Cue Location (L), Cue Type (T), and Cue Duration (D). The reported test statistic is *Likelihood Ratio Chi-Square*.

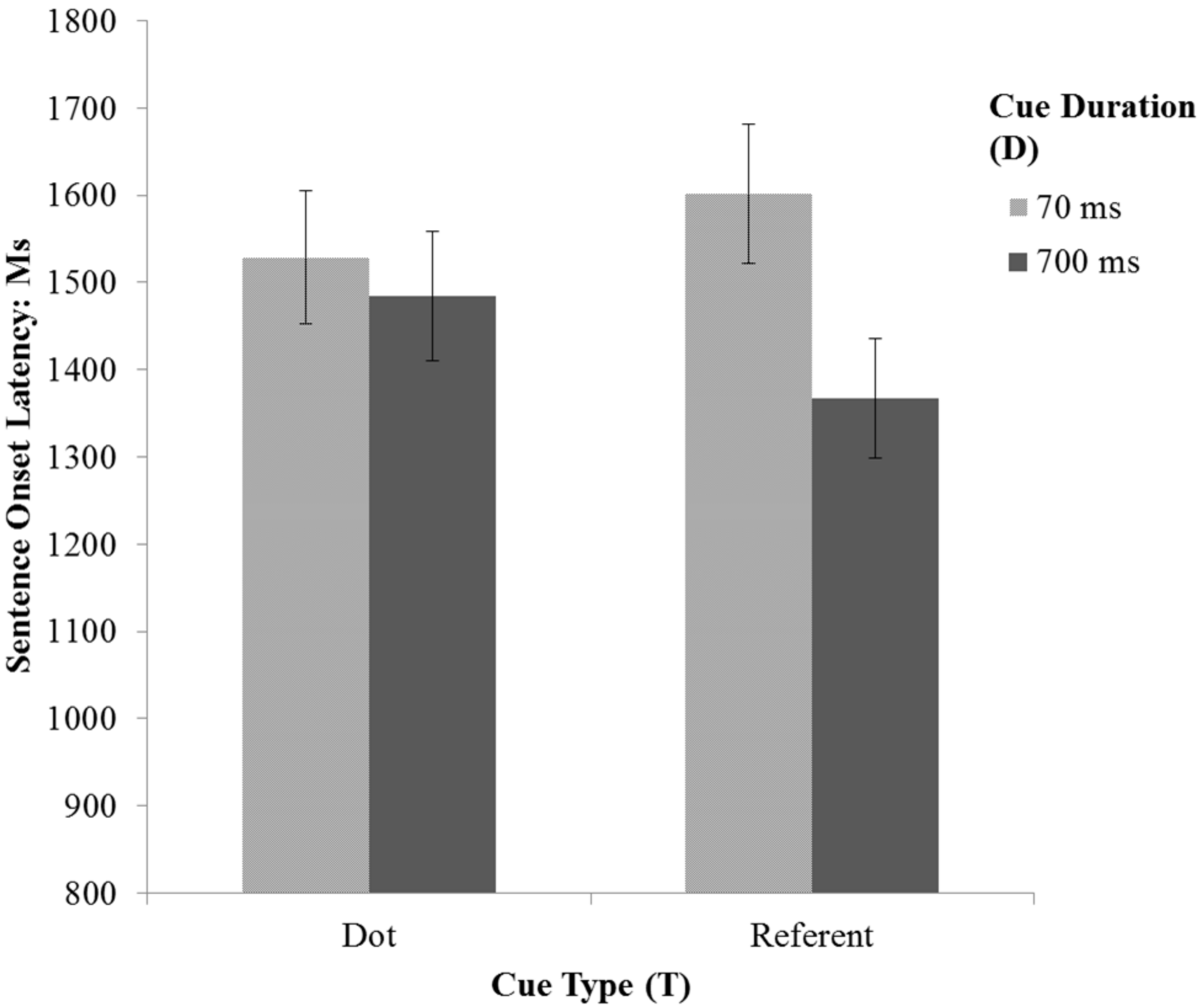

Figure 5. Mean sentence onset latencies (error bars represent standard errors) for the Cue Type x Cue Duration Interaction

The grand average intercept was estimated as 1497 ± 45 ms (μ ± SE). The analyses revealed a main effect of Cue Location: Sentences were initiated faster when the cue was on the agent (1420 ms) than when it was on the patient (1574 ms). This finding suggests that it was easier to plan and produce *canonical* SVO sentences, in which the agent acts as the Subject of the sentence (cf. Hwang & Kaiser, 2014; Myachykov, et al., 2013; Sevilla, et al., 2014). Our analysis also revealed a main effect of Cue Duration: Sentences were initiated faster when the cue was explicit (1426 ms) than when it was implicit (1568 ms). This finding suggests that participants are more prepared to initiate overt sentence production after their attention was explicitly anchored to one of the referents' locations. Finally, we registered a reliable interaction between Cue Type and Cue Duration (cf. Figure 5). Simple effect analyses showed that this interaction was driven by a significantly stronger Cue Duration effect in the Referent Cued condition. Specifically, event descriptions were initiated ca. 238 ms faster with an explicit (700 ms cue duration) rather than implicit (70 ms cue duration) Referent cue, yielding a significant simple effect of Cue Duration: ${}_{LR}\chi^2(1) = 28.220$; $P < .001$. With a Dot cue, by contrast, the simple effect of Cue Duration was not reliable (${}_{LR}\chi^2(1) = 1.223$; $P = .269$) as event descriptions were initiated only about 44 ms faster in the explicit (700 ms) rather than implicit (70 ms) Cue Duration condition.

Discussion

This study investigated the independent contributions of visual attention and memory activation to sentence production when describing a visual scene. Our main results show that visual attention (the result of cueing the location of one of the referents in the scene) and the referent's memory activation (the result of referent preview) make distinct contributions to the sentence production process: The former biases the choice of syntactic structure while the latter facilitates incremental sentence generation.

Our analysis of the proportion of passive-voice responses confirmed that cueing the location of one a referent generally increases the chances of this referent being selected as the subject of the sentence affecting the choice between available syntactic structures. The finding is consistent with previous studies showing that attentional focus plays a role in Subject assignment (and corresponding structural choice) in visually situated sentence production (e.g., Gleitman, et al., 2007; Myachykov, et al., 2011; Tomlin, 1995). That syntactic choice was not affected by Cue Type gives further evidence for the *special* role of attention in syntactic choice: Contrary to studies showing that conceptual accessibility plays an independent role in determining Subject assignment (e.g., Arnold & Griffin, 2007; Prat-Sala & Branigan, 2000), the was no differential effect of cuing either the referent or its location on structural choice (cf. Myachykov, et al., 2012). A reliable interaction between Cue Location and Cue Duration indicates that the strength of attentional cueing (i.e. explicit vs implicit) is particularly important for the selection of subdominant (non-canonical) syntactic alternatives (here, passive voice): A stronger visual cue to the patient further facilitated activation of the passive frame over the active frame.

Analysis of sentence onset latencies provided new and intriguing results. First, sentence initiation was faster when the cue was on the agent than on the patient suggesting that speakers find it easier to generate canonical active voice sentences with the agent in the subject role than passives. Second, the main effect of Cue Duration confirms that sentence initiations were faster when the cue was explicit than when implicit. Presumably this is due to explicits cues leading to stronger attentional focus than implicit cues hence facilitating overt sentence production. Although there was no main effect of Cue Type, Cue Type did interact with Cue Duration: Sentence onset latencies were further facilitated when the cue was both informative (Referent but not Dot) and Explicit while there was no difference between sentence onset latencies in Dot condition. This finding supports our hypothesis about a distinct role of conceptual accessibility in overt incremental sentence generation. Presumably, accessing and naming the lemma of the initial sentence constituent is affected by conceptual accessibility. Another important finding is that SOLs differed in the explicit cue but not in the implicit cue condition where a short 70 ms cue duration made the effect of both types of cue comparable. Hence, short preview is not sufficient for accessing conceptual information about the referent to facilitate sentence generation (cf. Schotter, et al., 2014). That there is no such interaction in the analysis of syntactic choice suggests that whereas referent preview preceding overt sentence generation leads to faster sentence assembly (perhaps faster lemma selection and lexical access) it does not affect syntactic choice. This suggests that syntactic choice is a distinct operation independent of referential analysis and naming.

Put together, our results demonstrate that the choice of sentential Subject and the resulting syntactic choice strongly rely on the allocation of the attentional focus to one of the interacting referents. Our results also demonstrate that this mapping mechanism is (1) categorical with the structure planned as a whole (cf. Myachykov, et al., 2013) and (2) independent from the incremental assembly of the sentence. Finally, our results support the idea that conceptual accessibility supports lemma access and incremental generation of constituents' names as components in a previously selected syntactic structure.

References

Antón-Méndez, I., Gerfen, C., & Ramos, M. (2015). Salience Effects: L2 Sentence Production as a Window on L1 Speech Planning. *Journal of psycholinguistic research*, 1-16.

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

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

Bock, J.K. & Ferreira, V. (2014). Syntactically speaking. In: M. Goldrick, V. Ferreira, & M. Miozzo, (eds.) *The Oxford Handbook of Sentence Production*, 21-46. New York, NY: Oxford University Press.

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

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75(2),* B13-B25.

Caffarra, S., Pesciarelli, F., & Cacciari, C. (2013). The interaction between language and visual spatial attention systems in grammatical gender processing. An N2pc study. *Cognitive neuroscience, 4(3-4),* 217-224.

Cai, Q., Van der Haegen, L., & Brysbaert, M. (2013). Complementary hemispheric specialization for language production and visuospatial attention. *Proceedings of the National Academy of Sciences*, 110(4), E322-E330.

Clark, H.H. & Chase, W.G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*, 472-517.

Coco, M. I., & Keller, F. (2015). Integrating mechanisms of visual guidance in naturalistic language production. *Cognitive processing, 16(2),* 131-150.

Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica, 125(2),* 129-143.

Ericsson, K.A. & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211-245.

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language, 56(3),* 357-383.

Givon, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics*, 30, 5-55.

Gámez, P. B., & Vasilyeva, M. (2015). Exploring interactions between semantic and syntactic processes: The role of animacy in* syntactic priming. *Journal of Experimental Child Psychology*, *138*, 15-30.

Gennari, S. P., Mirković, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: a cross-linguistic investigation. *Cognitive psychology, 65(2),* 141-176.

Gibson, B. S., & Bryant, T. A. (2005). Variation in cue duration reveals top-down modulation of involuntary orienting to uninformative symbolic cues. *Perception & Psychophysics, 67(5),* 749-758.

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of memory and language, 57(4),* 544-569.

Griffin, Z.M. & Weinstein-Tull, J. (2003). Conceptual structure modulates structural priming in the production of complex sentences. *Journal of Memory and Language*, *49,* 537-555.

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General, 142(3),* 880.

Hardin,J., & Hilbe,J. (2003). *Generalized Estimating Equations.* London: Chapman andHall/CRC.

Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science, 12(5),* 360-365.

Iwabuchi, T., Inui, T., Ohba, M., & Ogawa, K. (2013). A functional MRI study of a picture–sentence verification task: evidence of attention shift to the grammatical subject. *NeuroReport, 24(6),* 298-302.

Kastner, S. & Ungerleider, L.G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review Neuroscience*, *23,* 315-341.

Lipsitz, R.S., Kim, K., Zhao, L. (1994). Analysis of repeated categorical data using generalized estimating equations, *Statistics in Medicine, 13*(11), 1149–1163.

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27(3),* 817.

Montag, J. L., & MacDonald, M. C. (2013). Visual salience modulates structure choice in relative clause production. *Language and speech*, *57*, 163–180.

Myachykov, A., Garrod, S., & Scheepers, C. (2011). Perceptual priming of structural choice during English and Finnish sentence production. In R.K. Mishra & N Srinivasan (Eds.), *Language and Cognition: The State of The Arts.* (pp. 53-71). Lincom Europa, Munich.

Myachykov, A., Garrod, S., & Scheepers, C. (2012). Determinants of structural choice in visually-situated sentence production. *Acta Psychologica, 141(3)*, 304-315. DOI: 10.1016/j.actpsy.2012.09.006

Myachykov, A., Thompson, D., Garrod, S. & Scheepers, C. (2012). Referential and visual cues to structural choice in sentence production. *Frontiers in Psychology, 2*:396. DOI: 10.3389/fpsyg.2011.00396

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: The case of English and Russian. *Quarterly Journal of Experimental Psychology, 66(8)*, 1601-1619.

Olson, D.R. & Filby, N. (1972). On the comprehension of active and passive sentences. *Cognitive Psychology, 3*, 361-381.

Pickering, M.J., Branigan, H.P., 1998. The representation of verbs: evidence from syntactic priming in language production. *Journal of Memory and Language 39*, 633–651.

Posner, M.I. (1980). Orienting of attention, *Quarterly Journal of Experimental Psychology, 32*, 3-25.

Prat-Sala, M. & Branigan, H.P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English & Spanish. *Journal of Memory & Language, 42,* 168-182.

Prentice, J.L. (1967). Effects of cuing actor vs. cuing object on word order in sentence production. *Psychonomic science, 8,* 163-164.

Sanford, A.J. (2001). Context, attention, and depth of processing during interpretation. *Mind and Language, 17(1),* 188-206.

Sanford, A.J. & Garrod, S.C. (1981). *Understanding written language: Explorations beyond the sentence.* Chichester: John Willey & Sons.

Sanford, T. , Moxey, L. , Paterson, K. (1996). Attentional focusing with quantifiers in production and communication. *Memory and Cognition, 24(2),* 144-155.

Schotter, E. R. (2011). Eye movements as an index of linguistic processing in language production. *Journal of Psychology and Behavior, 9,* 16–23.

Schotter, E. R., Jia, A., Ferreira, V. S., & Rayner, K. (2014). Preview benefit in speaking occurs regardless of preview timing. *Psychonomic bulletin & review, 21(3),* 755-762.

Tomlin, R.S. (1995). Focal Attention, Voice, and Word Order. In P. Dowing & M. Noonan (Eds.) *Word Order in Discourse* (pp. 517-552). Amsterdam: John Benjamins.

Tomlin, R. S. & Myachykov, A. (2015). Attention in the organization and development of grammar. In E. Dabrowska & D. Divjak (Eds.) *Handbook of Cognitive Linguistics.* (pp. 31-53). De Gruyter Mouton.

Turner, E.A. & Rommetveit, R. (1968). Focus of attention in recall of active and passive sentences. *Journal of Verbal Learning and Verbal Behavior, 7(2),* 543-548.

Vogels, J., Krahmer, E., & Maes, A. (2013). Who is where referred to how, and why? The influence of visual saliency on referent accessibility in spoken language production. *Language and cognitive processes, 28(9),* 1323-1349.