Integrating Perceptual, Semantic and Syntactic Information in Sentence Production

Andriy Myachykov

Department of Psychology

University of Glasgow

Submitted for the Degree of Ph.D. to the Highers Degree Committee of the Faculty of Social Sciences, University of Glasgow

August, 2007

Acknowledgements

First, I would like to extend my deepest gratitude to my supervisor and captain, Simon Garrod, for his excellent and indispensable guidance throughout the time of my Ph.D. research and the preparation of this thesis. May you always have seven feet under the keel and favourable winds!

I am also grateful to Christoph Scheepers for his extensive advice on methodological and statistical issues.

This research would not be possible without the help in data collection from Jukka Hyona, Seppo Vainio, Andrej Kibrik, Mira Bergelson, Viktor Shklovsky, Maria Ivanova, and Stephanie Coquillon.

Special thanks to Mike Posner, Russ Tomlin, and Svitlana Zhabotinska for advising me in my previous studies, putting me on the path of academic career, and encouraging me throughout my Ph.D. work.

Finally, this work would have never been accomplished without the support, the encouragement, the patience, and the grace of my family – my wife, Hanna and my children, Ivan and Lisa. You were always there for me to remind me that I could do it. Looks like you were right after all…

Declaration

I declare that this thesis is my own work carried out under normal terms of supervision.

To my Mother, who would be proud of me…

Abstract

The experimental work and the theoretical model presented in this thesis explore the behavior of the sentence production system in perceptually, conceptually, and syntactically changing environments across languages. Nine experiments examine how speakers of different languages integrate available perceptual, conceptual, and syntactic information during production of sentences. Such integration occurs under the global control of canonical causality and automated syntax. Analysis of speakers' performance in perceptually manipulated setting demonstrated that perceptual motivations for word order alternation are relatively weak and limited to the initial event apprehension. In addition, salience-driven choices of word order are realized differently in different syntactic structures and in languages with different grammatical systems. Combining perceptual and conceptual priming paradigms did not substantially improve cueing efficiency. Contrastingly, early availability of lexical and syntactic information led to the most consistent alternation of the word order.

I conclude that the uptake of perceptual information does not directly influence structural processing. General cognitive processes, such as attentional control and higher memorial activation actively contribute to the concept's accessibility status, but the syntactic organization of a spoken sentence constitutes a relatively independent psychological reality that can be realized partially as a product of the aforementioned operations but does not directly depend on them.

Table of Content

Introduction. Executive summary

The major undertaking of psycholinguistic research is to improve our knowledge about the psychological processes involved in acquisition of linguistic knowledge, cognitive organization of the language faculty in the human brain, and specifics of real-time linguistic communication. The studies discussed in this thesis relate to the latter research question.

Two major theoretical questions motivate the research in this thesis. The first question relates to the linguistic component. It has to do with the fact that when people talk they need to properly organize grammatical properties of the discourse in order to facilitate communication. One part of this process is being able to identify, track, and bind the concepts that discourse is about. In addition, any given language has a variety of grammatical means to express the same semantic content. For example, a speaker of English could describe the event portrayed in Figure 1 by producing one of the two following sentences:

1. A policeman is kicking/kicks a boxer.
2. A boxer is being kicked/was kicked/got kicked (by a policeman).

What determines which of these possible syntactic variants a native speaker of English will use in a given discourse environment? This question is by no means trivial. Even using the most superficial analysis, one finds that this question can be further subdivided in two more specific questions. The first question is what factors make the elements to appear first, second, or last in the sentence? The second question is what factors determine the choice of the grammatical structure to convey the kicking interaction between the policeman and the boxer? These two questions may look similar, but in fact they address quite different phenomena. The factors determining the ordering of words (positional factors) and the assignment of the grammatical relations between them (structural factors) may be related, but they may also be distinct and independent processes that rely on contributions from different sources of information.

Undoubtedly, perceptual, conceptual, and linguistic factors together influence the positioning of the constituents in a spoken sentence. For example, it is well known that speakers tend to place perceptually salient information before the rest of the discourse material. Also, previously mentioned information tends to appear before new information in discourse. Finally, speakers actively recycle previously used or encountered syntax in their speech, putting additional constraints on the two previously mentioned forces. It is unclear, however, whether and how these separate forces interact in their ability to bias speaker's grammatical choices. One purpose of my research is to investigate both structural and positioning effects in sentence production through priming different elements within the discourse environment – perceptual, conceptual, and syntactic.

The second theoretical question I ask in this thesis is largely psychological. It hinges on our understanding of the mind's ability to simultaneously rely on two sets of computations: the computations that belong to the general processing domains, such as memorial activation and attentional control, and the computations that are bound to linguistic processing. Together, the domain-general and the language-specific operations allow processing of the perceived reality into conceptual content, encoding of such content into lexical material, and positioning this lexical material in a sentence according to the constraints of the language grammar.

Although it seems natural to assume that language processing relies on both domain-specific and domain-general processes, little is known about exactly what types of linguistic operations share resources with the rest of human cognition or processing of which linguistic units is more independent than others. The continuum of theoretical views on the topic ranges from extremely modular (e.g. Chomsky, 1965) to extremely interactive (e.g. Lakoff, 1987; Vygotsky, 1962; Whorf, 1965). As with all complex mental operations, the truth probably occupies the middle ground: While some of the psychological material language is cut from is borrowed from general cognition, other layers of language processing are specific to linguistic operations (Jackendoff, 2002).

Such co-dependence prompts the existence of a persistent, ubiquitous, and language-specific communication interface between *"general"* and *"special"* components of language processing (Jackendoff, in press). The role of such an interface is to ensure successful mappings from event semantics onto abstract units of phonology and syntax in accordance with interlocutors' intentions or event's salience map. Many theorists observe this dichotomy and agree that "the chief issues in language production include how and when the processing system retrieves different kinds of linguistic knowledge,…how the system interrelates linguistic and non-linguistic knowledge, and how the system is organized within and constrained by, human cognitive capacities " (Bock & Huitema, 1999, p.204). The degree of domain specificity involved in real-time language processing and the regularities in the interface between language and cognition that support such processing are the two major psychological questions motivating this thesis.

The methodological framework for the research discussed in this thesis reflects real-time world-situated discourse and is usually referred to as Visual World paradigm (Tannenhaus, et al., 1995). As such, it allows investigations of real-time linguistic performance; therefore the analyzed discourse typically proceeds extemporaneously and unfolds online. Also, this paradigm is limited to a language processing that is visually mediated; thus the analyzed sentences are always about the events and entities immediately observed by the speaker. Because of this, there is always a strong perceptual component involved in both production and comprehension studies using Visual World paradigm. Although I will make some general theoretical inferences about the perceptual biases on sentence production, one caveat has to be observed. One should be cautious when interpreting results of the Visual World studies as such results are limited to the visually-mediated language. Whether processing language about abstract concepts and the displaced events follow the same principles is largely unclear (but see Altmann, 2004; Barsalou, 1999; Glenberg & Kaschak, 2002; Richardson, et al., 2003; Spivey and Geng, 2001).

Also, although linguistic communication typically involves at least two parties – a speaker and a listener – the research reported in this thesis is limited to the *speaker's* performance; thus, the role of the listener in biasing speaker's choices is not discussed.[1] Finally, current studies are based on *single sentence* production, therefore contextual

[1] Some theories suggest that the speakers plan and deliver the semantic content and the syntactic structure of the sentences partially tuned to the listeners' knowledge and expectations (e.g. Clark & Murphy, 1983; Bell, 1984; also Garrod & Anderson, 1987) while others draw somewhat more egocentric picture of speakers' performance (e.g. Brown & Dell, 1987; Ferreira & Dell, 2000; Keysar, et al., 1998). Current research, however, is based on single sentence production without an addressee in sight. Therefore, I do not make any claims about the role of the audience on speaker's linguistic choices.

factors, which undoubtedly make their own contributions to the organization of sentences, are discussed very briefly and indirectly.

The thesis is organized into 8 Chapters. Chapter 1 discusses existing research using perceptual priming paradigm[2] and the evidence for the idea that a speaker's syntactic choices may be motivated by the distribution of the attentional resources within the described array. Typically, the idea proposed by the advocates of such a view follows from the proposal that the perceptually salient entities tend to appear in syntactically prominent positions. According to this hypothesis, the item currently in the focus of the speaker's attention acquires preferential accessibility status, and, therefore, is more likely to be entered first into the frame of the sentence. On the other hand, an opposition to this claim suggests that a bottom-up link form attention to syntax is unlikely; instead, it is what people choose to say that determines the distribution of attention among the event's referents. The limited existing research using perceptual priming paradigm uses data from English speakers. The lack of studies using languages with a different grammatical structure motivates the research reported in Experiments 1-5 of the thesis.

Chapter 2 reviews referential priming studies and evidence that conceptual forces related to referential access also play a role in the formulation of syntax. For example, some experimental evidence suggests that the references associated with given or old discourse status tend to precede the new information in spoken discourse (e.g., Bock, 1982). Other semantic factors that demonstrate to influence order of mentioning include

[2] In this thesis, I use two terms that many researchers treat as synonymic in certain experimental contexts: *priming* and *cueing*. To prevent readers' confusion I need to explain how these terms are different. I use the term *priming* when referring to *the experimental paradigm* used in a particular study (e.g., perceptual, referential, or structural priming). I use the words *cue* and *cueing* only when talking about *visual* materials used in experimental protocols to achieve priming effect. Hence, the use of these two terms is limited to priming studies that employ (1) strictly perceptual cues (e.g., an arrow) and (2) cues that combine perceptual and conceptual properties (e.g., single referent preview). When priming effect is established using lexical and/or syntactic materials, I refer to such materials as *primes*.

animacy, definiteness, and humanness. Unfortunately, both perceptual and referential priming studies often confound perceptual and referential components. It is important to separate these effects in order to be able to arrive at a more valid and comprehensive picture of the information integration processes during the formulation of sentences. In Chapter 3, I compare perceptually and semantically-motivated accounts of sentence production and outline possible ways to combine both types of manipulations in the same experimental setting. This makes it possible to generate specific hypotheses about possible conflicts and interactions between perceptual and referential factors during sentence production. These hypotheses are later tested in Experiments 6 and 7.

Chapter 3 draws attention to the idea of syntactic or structural priming being a strong determinant of the speaker's grammatical choices. Such grammatical biases result from preceding discourse and lead to a strong tendency of interlocutors to repeat previously encountered syntax (Bock, 1986). The basis of syntactic priming is an ongoing debate in psycholinguistic research. While some maintain an autonomous-syntax explanation (Bock, 1989; Ferreira, 2003), others suggest that the power of syntactic priming may vary dependent on the degree of conceptual or lexical-semantic overlap between the prime and the target sentences (Branigan, et al., 2000; Cleland & Pickering, 2002; Corley & Scheepers, 2002; Pickering & Branigan, 1998). Regardless of the theoretical assumptions, the experimentation on the topic has never involved producing primed sentences in a perceptually manipulated environment. This empirical gap motivates Experiments 8 and 9 reported in the current thesis. Chapter 3 finishes discussion of background evidence for the research described in Chapters-5-7.

Chapter 4 discusses sentence processing models that afford different amounts of interaction between domain-specific and domain-general operations during generation of sentences. These models are roughly divided into two types: modular and interactive. Proponents of the first type usually argue for a certain degree of separation between processing modules so that computations occurring at each stage are relatively encapsulated and the processing generally proceeds in feed-forward manner (e.g. Bock & Levelt, 1994). The second type permits parallel computations on the inputs at different stages of sentence generation and allows constant update of the processed material via a series of loops. The parallelisms are usually realized as spreading activations that occur almost simultaneously or in a cascading fashion (e.g. Bates & MacWhinney, 1982; Dell 1986). I suggest that modular and interactive approaches are not mutually exclusive, and that, with a certain amount of creativity, features of both can be accommodated within a single model. I propose one possible design of such symbiotic processor, spell out some predictions about the model's behaviour, and propose a set of relevant measurements for the assessment of speakers' behaviour at each processing stage.

Chapter 5 begins discussion of the experimental data. Experiment 1 uses the Fish Film protocol (Tomlin, 1995) in order to investigate Russian speakers' performance on an attentionally cued sentence production task. The methodological limitations of the Fish Film paradigm are overcome in Experiments 2-5 by using novel materials, eye-tracking methodology, and a wider range of behavioral data, such as, speech-related and gaze-related reaction times and the eye-voice spans associated with the incremental construction of the sentential frame. Experiments 2 and 3 provide norming data for English and Russian sentence production with a minimal manipulation set. Experiments 4 and 5 investigate the performance of Finnish speakers on two experiments using implicit cueing of attention. Experiment 4 addresses the production of transitive, and experiment 5 – the production of ditransitive sentences. In addition to perceptual effects on the word order choice, an effect of the notional verb preview is discussed in both studies.

Chapter 6 continues the discussion of the experimental data. Experiments 6 and 7 introduce a more complex methodology combining perceptual and referential priming with a set of visual constraints on how freely the speakers can interrogate the described scene. The notional verb preview remains an additional factor of the analysis. This complex methodology allows insights into the interactions between perceptual, conceptual, and lexical information available to the speaker in the course of sentence production.

In Chapter 7, two more experiments (Experiments 8 and 9) using a combination of perceptual, referential, and syntactic priming are reported. These studies test a set of hypotheses about possible interactions between perceptual, conceptual, and syntactic forces in their ability to influence the formation of new syntax during the production of spoken sentences. Chapter 8 provides thesis conclusions and suggests directions for further research.

Chapter 1. Perceptual effects in sentence processing

1.1. Category of Attention

A large portion of this thesis discusses the contribution of perceptual, or more narrowly, *attentional* effects to the speakers' choice of word order. Thus, it is important to outline some general properties of the human attentional system that are relevant to the current research. I will only discuss here such basic notions of the attentional system as attentional networks and attentional shifts. Also, I will introduce the most relevant features of attentional manipulations in psychological experiments, such as (1) explicit and implicit cueing of attention and (2) the difference between overt attention, which correlates with the deployment of visual focus and implicit attention, which happens outside of the visual focus.

What people see is largely determined by what they happen to attend to. In the situation when the environment contains excessive perceptual information available for processing, attention facilitates selection of the information relevant to making behavioral decisions (Chun & Wolfe, 2001). Many psychologists underline the prominence of attention among other cognitive operations. For example, Tichener (1908) called attention "the heart of the psychological enterprise." (cited from Posner, et al., 2007). In his classical definition, William James suggested the following understanding of what attention is.

> "Everyone knows what attention is. It is the taking possession of the mind in clear and vivid form of one out of what seem several simultaneous objects or trains of thought". (James, 1890, p.402).

Recent definitions are still quite subjective. For example, Corbetta (1998) defines attention in the following terms.

> "Attention defines the mental ability to select stimuli, responses, memories, or thoughts that are behaviorally relevant, among the many others that are behaviorally irrelevant" (Corbetta, 1998, p. 831).

Posner and colleagues (e.g., Fan, et al., 2002; Posner & Fan, 2004; Posner & Petersen, 1990; Posner & Raichle, 1994; Posner & Rothbart, 2007a) suggested a very detailed and systemic view of attention. Using the Attention Network Test (ANT) they were able to evaluate three anatomically and behaviourally distinct brain areas or networks responsible for *alerting, orienting,* and *executive control* of attention. The alerting network supports achieving and maintaining an alert state; orienting helps selection of information from sensory input; and executive attention provides resolution of a conflict among responses (Fan et al., 2002). These networks have distinct anatomical, developmental, and genetic organization (Fan, et al, 2002; Posner & Fan, 2004; Posner & Rothbart, 2007b, Posner, Rothbart, & Sheese, 2007). Trying to adapt attentional theory for language research, Tomlin proposed a single notion of *attention detection* (Tomlin, 1995, 1997). Attention detection collapses operations from Posner's three-partite set. According to Tomlin, an attentionally detected entity is the one currently in the speaker's attentional focus. Although slightly underspecified psychologically, the notion of attention detection proved adequate for visually cued sentence production tasks, in which speakers' attention is manipulated explicitly so that the visual focus always correspondence to the attentional focus. Although a finer-grained attentional manipulation may be necessary to separate contributions of separate attentional networks to the organization of particular linguistic phenomena, a single attentional operation approach is adequate for the first approximation analysis.

One more operational category is important for the current research. It is the notion of *attentional shifts* (Posner & Petersen, 1990). These shifts include (1) disengaging from some current focus and (2) moving the index of attention toward to a new area or a stimulus. One recent study implicated the posterior parietal lobe in particular as a center for controlling the deployment of attention to locations in space (Yantis, et al., 2002). In terms of the precise function of this area, two hypotheses are considered. According to the "sustained hypothesis", activity in parietal lobe *maintains the locus* of attention. According to the "transient hypothesis", this area is associated with the act of *switching attention* from one location to another. The difference between the two hypotheses is in the predicted time course of neural activity and not in the spatial distribution. The results of the study conducted by Yantis and colleagues speak in favor of the "transient hypothesis".

The control of attention in visual experimental tasks is usually achieved by means of cueing paradigm (Posner, 1980), which differentiates between *exogenous* and *endogenous, explicitly* and *implicitly* driven, and *overt* and *covert* deployment of attention (Posner & Raichle, 1994). Exogenous cues are the salient features in the outside world, which draw the perceiver's attention to a particular location. Endogenous cues originate from within the perceiver's mind and are guided by his internally generated plans in carrying out cognitive tasks. One study (Corbetta, et al., 2002) suggested a model of partially segregated networks of brain areas that support cognitive (endogenous) and sensory (exogenous) orienting systems. One system, which includes parts of the intraparietal cortex and superior frontal cortex, is involved in preparing and applying top-down selection for stimuli and responses. The other system, which includes the temporoparietal cortex and inferior frontal cortex, and is largely lateralized in the right hemisphere, is specialized for the bottom-up detection of behaviorally relevant stimuli, particularly when they are salient or unexpected. The facilitation produced by the sensory cues appears more rapidly (within 50 ms.) than that produced by cognitive cues.

Attention can be captured by means of either explicit or implicit cueing. An explicit cue is a clearly visible and therefore consciously processed marker that attracts attention to a location or an entity. An example of such cue is an arrow pointing toward a certain location on the screen presented long enough (e.g., 500 msec.) for a participant to notice it and (likely but not inevitably) direct her gaze toward it. An implicit cue directs attention in a more subtle manner. Such a cue is usually presented very briefly, for duration much smaller than would be necessary for conscious processing of the cue (e.g., 50 msec.). Although the presentation of an implicit cue is typically not noticed by a participant, its brief display may be efficient in attracting attention and directing the gaze toward the cued location.

However, eye-movements do not necessarily accompany attentional shifts although they typically follow the allocation of attention (Fischer, 1998, for a review). An overt attentional shift occurs when the eyes move to align the visual focus with the attended object. A covert shift directs the focus of attention outside of the visual focus making the two foci dissociable. In a way, it is similar to “looking out of the corner of your eye”. Posner and colleagues (Posner, 1980; Posner, Nissen, & Ogden, 1978) developed the spatial cueing task to measure covert shifts of visual attention. In this task, observers are required to respond to a peripherally presented target, which is preceded by a cue that serves to direct covert visual attention to a particular location. A typical finding of spatial cueing studies is one of more efficient processing of targets appearing in the cued location compared to a non-cued location. A processing advantage at the cued location, such as faster detection of the target, is assumed to reflect a shift of covert attention. It was also shown (Shulman, Remington & McClean, 1979) that response times to probes at intermediate locations were enhanced at intermediate times as though attention actually moved through the space and that it was possible to prepare to move the eyes to one location while moving attention covertly in the opposite direction. Such dissociation between the attentional and the visual foci is notoriously difficult to elicit, and it rarely occurs outside the experimental settings. Whether attention in fact moves through the intermediate space and how free covert attention is from the eye movement system are still disputed matters.

1.2. Attentional system and language processing

The psychological literature provides (1) structural, (2) developmental, and (3) behavioural support for the idea that linguistic performance relies upon allocation of attentional resources.

The structural argument proceeds from the evidence that the human brain is flexibly organized so that the same cortical region often supports a variety of mental operations. For example, neuroimaging studies in reading inform us about brain areas involved in chunking visual letters into words, associating letters with sounds and providing entry into a distributed lexicon of semantics. Chunking visual letters into words takes place in a posterior visually specific area of the left fusiform gyrus (McCandliss, et al., 2002). In the right hemisphere, similar areas are involved in the perception and individuation of faces (Kanwisher, et al., 1997). While these areas were first thought to be word and face specific, more recent conceptualizations argue that they are more related to process of chunking of visual elements or individuation of complex forms, which can be performed on other inputs, for example on dogs or horses if one has become expert enough to individuate them (Gautier, et al., 1999). This same principle of localization of mental operations rather than domain specific representations may explain why Broca's area seems important for some forms of non-speech motor activity. For example, structural ERP research has shown a large area of activation in the anterior cingulate gyrus during lexical search (Abdulaev & Posner, 1998; Raichle, et al., 1994). The same area is known to be involved in conflict resolution and executive attention (Fan, et al., 2002; Posner & Petersen, 1990). An fMRI study (Newman, et al., 2001) revealed that syntactic violations elicit significantly greater activation in superior frontal cortex – the area largely involved in attentional control. Other neuroimaging studies revealed a strong attentional component in syntactic processing. Violations of syntactic structure in the studies using Event-Related Potentials of the scalp (ERP) give rise to quite distinct wave forms (Kutas & Van Petten, 1994, for a review). As a result, a typical EEG response consists of an early left anterior negativity (LAN) (e.g. Osterhout & Mobley, 1995) and/or a late positive wave with a peak at 600 ms (P600) (e.g. Hagoort, et al., 1993). Hahne and Friederici (1999) hypothesized that the early left anterior negativity is a highly automated process, whereas the P600 involves more attention. Hanne & Friederici tested this hypothesis in a study manipulating the proportion of correct sentences and sentences with structural violations in them. Syntactically incorrect sentences appeared in a low (20% violation) or a high (80% violation) proportion conditions. Both conditions led to the elicitation of the LAN effect, while only low proportion of incorrect sentences resulted in P600. These results support the idea that early left negativity is an automated first-pass sentence parsing mechanism invariably observable as a correlate of syntactic processing. The P600 component relates to a second-pass parsing that requires a larger allocation of attention and a deliberate deployment of executive attention. All of these neuroimaging findings point to the fact that the brain localizes *processes* or *mental operations* not particular types of representation either linguistic or non-linguistic and that the sharing processing regions may lead to sharing resources between domain-specific and domain-general operations computed in the same area.

Research in infant development suggests that attentional amplification of the visual input is actively used by the caretakers during the early stages of language development. For example, experiments show that children follow the interlocutors' gaze when learning the meanings of words (Carpenter, et al., 1998; Carron, et al., 2002). Also, joint attention between mother and child appears to be a major determinant of language learning (Baldwin, 1995; Dominey & Dodane, 2004). The importance of attentional control in language development suggests an early and a potentially strong coupling between the distribution of attention in the environment and the organization of the language about this environment. In adult performance the link between attending to objects and acting on them remains strong. People tend to look at objects of their actions regardless of whether they linguistically describe their actions on these objects or not

(Ballard, et al., 1997). If one were to understand linguistic processing as a subsystem of other behavioural tasks, a similar link can be expected between attending to objects and naming them in a sentence.

Assessment of linguistic behaviour in adults also led many researchers to conclude that attentional performance correlates with linguistic processing at a variety of levels. Claims about perceptual processing being somehow represented in the syntactic system were made by a number of theoreticians. For example, Landau and Jackendoff (1993; also Jackendoff, 1996) suggested that how representing objects are in the human mind (*what*) and locations (*where*) maps directly onto the distinction between nouns and prepositions. Such consistent mapping from perception to conception to grammar is believed by many to be the moving force of language acquisition. For example, Mandler (1992) proposed that, after initial analysis, perceptual information in child's mind is represented in a form of image schemas that support development of more abstract conceptual representations and derived thematic relationships.

With regard to attentional control, research in both early and late bilinguals showed that the constant management of and switching between, two languages correlates with enhanced executive attention (Bialystok, 2001; Yang & Lust, 2005) and a better maintenance of the alert state (Costa, et al., 2007). At the same time, better performance on Stroop task and to lesser extent on the span task predicts more successful L2 vocabulary learning (Michael & MacWhinney, 2003; Miller & Kroll, 2002). Existing research with adult monolinguals further supports the existence of a systematic relationship between the distribution of attention to the elements in the described scene and the organization of speech that describes it (see next Chapter). The most important assumption behind perceptually-driven sentence production proposals is that the form of a given grammatical structure may to some extent depend on operations within the domains shared by language with other cognitive processes. This leads to a functional and/or cognitive view of grammar, according to which linguistic structures reveal the cognitive processes involved in the preparation and production of sentences. Basic cognitive operations such as memory retrieval and attentional tracking of entities therefore become important phenomena underlying aspects of grammar and lexicon.

1.3. Perceptual manipulations in psycholinguistics

On the psycholinguistic arena, some of the early motivations for the idea that the attentional processing of the cognized world may somehow be reflected in how people organize their production and comprehension of sentences come from studies by Osgood and Bock (1977) and MacWhinney (1977). The former study explicitly suggested that the referents' salience status (*vividness*) acting as an exogenous determinant of the distribution of speaker's attention should promote the referents currently in focus to the prominent positions in a spoken sentence. The second study presented a theoretical platform known as the *Starting Point* hypothesis. Although it is not specifically geared toward sentence production, the Starting Point framework predicts that one of the main factors determining the assignment of the prominent positions in a sentence is the interlocutor's *perspective* or attentional focus.

1.3.1. Perceptual component in sentence comprehension

It is only natural to assume that when people have to comprehend the discourse about visually perceived events, the distribution of attentional foci will somehow correspond to the elements of the scene being currently processed. A number of *comprehension* studies supported existence of a tight linkage between perceptual properties of the world and the organization of sentences about it. In many of these studies, attentional focus appeared among factors that influence referential access and anaphora resolution (Myachykov & Posner, 2005 for a recent review), facilitated activation of lexical items in comprehension tasks and acted as an important resource in organizing the structure of dialogue (Garrod & Pickering, 1999; Sanford, 2001; Sanford & Garrod, 1981, Sanford, et al., 1996). A more recent experimental paradigm studying visual effects in comprehension is *visual world* paradigm (Tannenhaus, et al., 1995). In a typical Visual World task, speakers have to process simultaneously linguistic and visual inputs and map the converging message onto the event semantics. Analysis of eye movements in these studies provides accurate time-locked scanning signature of what information becomes available to the speaker at any given time during the task and when this information becomes available (Meyer, 2004). Eye movements and fixations in visual world studies are usually interpreted as correlates of attentional foci. As such, they were observed to accompany retrieval of semantic information during production of single words (Levelt & Meyer, 2000; Meyer et al.,1998; van der Meulen, et al., 2000) and reflect syntactic interpretation of sentences with visually co-present (Tanenhaus, et al., 1995) as well as displaced events (Altman & Kamide, 2004; Altman & Kamide, in press). A study by Altmann & Kamide (2004) demonstrated that the time necessary to launch and plan a saccade in Visual World tasks may reflect the access to the information about the referent when relating auditory information to the visual display. Finally, although contribution of strictly perceptual factors seems to be ubiquitous in facilitating our comprehension of visually mediated speech, resolution of the mapping between the linguistic input and the described scene seems to depend on a combination of the perceptual *and* semantic information from the observed scene and the information contained in the initial elements of the comprehended sentence (Knoeferle, et al., 2005).

1.3.2. Perceptual priming in sentence production

Another question is whether systematic variations of the syntactic structure during *the production of sentences* also depend on attention to individual referents? This question is related to the understanding of the structural organization of a spoken sentence as a correlate of the perceptual details in the scene this sentence describes. Some models propose that preferential positioning of the constituents in the upcoming sentence may directly depend on the distribution of the speaker's attentional resources among the event's referents (e.g. Tomlin, 1995). The attended object, in such bottom-up system, is either promoted as a higher accessibility item within the conceptual and then the linguistic representation (cascaded view) or it is repeatedly reactivated (serial processing view). For example, if the attention of a native speaker of English is directed to the Agent of a transitive event like the one portrayed in Figure 1, he may be more likely to describe such event with a sentence like *"A policeman is kicking a boxer"*. If, on the other hand, the attentional focus is on the Patient, a sentence like *"A boxer is being kicked by a policeman"* seems more likely. In other words, focusing attention on the referent may lead to an earlier commitment to a sentential starting point and the distribution of the syntactic roles in a spoken sentence

In fact, there is some experimental evidence that confirms such scenario (Myachykov, et al., 2005, for a recent review). For example, speakers tend to inspect visually presented scenes before they talk about them. Such visual inspection was found to (1) precede the production of simple noun phrases (Meyer, et al., 1998), (2) precede and mirror the order of mentioning when describing arrays of multiple objects (van der Muelen, 2001, 2003), and (3) full-fledged sentences (Griffin & Bock, 2000). However, while naming studies observed that speakers tend to start naming the objects incrementally before scanning the whole array, sentence production studies reveal a non-incremental phase of rapid apprehension that precedes incremental formulation of a sentence (Griffin & Bock, 2000, van der Muelen, 2001, 2003).

Unfortunately, existing research on perceptually-motivated sentence production is very limited. Nevertheless, one of the most prominent attempts to unravel the relationship between attention and the assignment of syntactic roles in a sentence was undertaken by Tomlin (Tomlin, 1995; 1997). Tomlin's experimental program stems from functional linguistic tradition, and it came as an opposition to a view commonly shared in linguistics, which is well illustrated by the following extract:

> The (sentence) constituents move to certain positions because of their discourse function interpretation (King, 1995, 63).

This and similar views are based on the idea that the assignment of the syntactic positions in a clause is based on the functional opposition between clause-level *theme* and *rheme*, *topic* and *comment,* and/or the referents' semantic roles, like *agent* and *patient*.

Traditionally, (Daneš, 1970; Firbas, 1965; Halliday, 1985; Mathesius, 1929) the notion of theme is associated with the element “one is talking about, the topic,” while the rheme is “what one says about it, the comment” (Daneš, 1970). It is widely assumed that the subject of a sentence frequently acts as the syntactic counterpart of the theme or topic of the utterance although other structures have been known to highlight the theme through means different than subjecthood, for example clefts and dislocations. Functional interaction between discourse entities is usually realized as the hierarchy of semantic roles. For example, the term *agent* is traditionally used to identify an acting instigator of the action while *patient* is referred to as an experiencer of the agent’s action (Fillmore, 1968). It was suggested the semantic agent is the most likely candidate to take the position of syntactic subject of a sentence. Some psycholinguistic studies (e.g., Kako, 2005) demonstrated that speakers tend to perceive sentential subjects as “agent-like”, while objects are rated as more “patient-like”. Interestingly, the latter study reports the existence of such link using rating tasks using both real language and nonce language materials. Although the controversy about the relation of structural mappings and the correspondence of the syntactic slots to the thematic roles continues (e.g., Bock & Loebell, 1990; Bock, Loebell, & Morey, 1992) some studies using syntactic priming paradigm (Bock, 1986b) confirmed the existence of such link (Chang, Bock, & Goldberg, 2003).

In response to this common theoretical platform, Tomlin conducted a set of experiments, in which he used a computer animation program called “The Fish Film”. Participants viewed and described an unfolding engagement of two fish, which resulted in one fish eating the other (see Figure 2). In each trial, a visual cue in form of an arrow directed participants' attention toward one of the two fish. The cue appeared at the beginning of a trial and later disappeared, finally reappearing again right before the target event. The experimental instruction was to treat the arrow as an explicit cue and to direct the gaze to the cued fish whenever the cue appeared on the screen. Apart from being outwardly *explicit*, such cue is also *mixed*, as it employs both endogenous (participants are told to treat the arrow as an attentional cue) and exogenous properties (the cue draws their attention to the target by virtue of pointing to it) (Posner, 1980). Also, such cue is *constraining* (as the instruction does not permit switching of attention to the non-cued fish when the cue is displayed).

Descriptions of the eating event were analyzed for their syntactic structure. The results demonstrated that participants consistently varied the assignment of the syntactic subject and the grammatical voice dependent on which fish was cued: When the agent fish was cued the participants produced active voice sentences, when the patient fish was cued they produced passive voice sentences. This was true in virtually 100% of all experimental trials. Based on his results, Tomlin drew a strong conclusion that the grammatical subject in English may be consistently assigned to the referent that is currently in the speaker's attentional focus.

One potential reason for such a powerful effect in Tomlin's experiment is probably the nature of the employed cueing procedure. The cue used by Tomlin was quite strong because it was explicitly presented with the stimulus and because it combined both exogenous and endogenous properties. Both the cueing procedure and the repetitive nature of Fish Film paradigm received criticism from some psycholinguists for being "too brutal" (Bock, et al., 2004) or crude and suggestive about the experimenter's goal

(Gleitman, et al., in press). From methodological point of view, such criticisms are, at least partially, justified. First, although the experimental instructions did not say anything about how to treat the cue in relation to the choice of event description, it considerably constrained their attentional focus to the cued referent making it not only perceptually but also conceptually more accessible. Second, although the cue itself did not provide any semantic information about the target (cf. Olson & Filby, 1972), its coupling with the continuously presented stimulus probably enhanced the conceptual prominence effect. Third, Fish Film protocol instructs participants to both view and describe continuously *all* the interactions between the fish, including those preceding the target event. This inevitably increases the discourse status of the both fish, more so for the fish that is cued making it the old or given discourse element (cf. Bock, 1982, Chapter 2 of this thesis). Finally, the repetitive nature of the target event and the lack of interrupting filler materials make effects of syntactic priming a possible concern (but see Chapter 5). For the time being, it is important to note that whether explicit manipulations of attention lead to a stronger alternation of the produced syntax than the implicit manipulations do is a valid empirical question. Experiments 6 and 7 of this thesis, among other things, cast some light onto this issue. Also, Experiments 7 and 9 introduce a combination of perceptual and semantic manipulations in order to arrive at a clearer picture of how these different sources of information are used by the speaker during preparation and execution of the syntactic plan.

Tomlin's findings received support in studies using other syntactic structures (e.g., Forrest, 1997), in languages other than English (Diderichsen, 2001), and for linguistic forms which are not present in English, such as Japanese *wa* (Hayashi, et al., 2002). For example a study by Forrest (1997) explored perceptually primed production of locative events. The experimental protocol used in Forrest's experiment was methodologically more advanced than "brutal" cueing force used by Tomlin in his studies. This time, speakers' attentional focus was manipulated prior to the target event presentation, which separated attentional cueing from conceptual analysis of the described event. Also, a masque was used between each target trial in order to minimize both visual and linguistic priming from trial to trial. The experimental materials were simple line drawings of locative events, for example a star left of a heart. The visual cue cued the location of either the start or the square prior to target display presentation. As a result of this perceptual manipulation, speakers tended to produce sentences like *A star is left of a heart* when the cue was in the left part of display, and *A heart is right of a star* when the cue was in the right part of display. Although the perceptual priming effect was not as large as in Tomlin's studies, it was still quite big to support Tomlin's account or perceptually-motivated grammatical role assignment in English sentence production.

However, It is important to draw a cautious demarcation line between interpretations of the visually-cued production studies that use *syntactic alternation* tasks (e.g. Tomlin, 1995, this thesis) and those using *starting point* tasks (e.g. Forrest, 1997; Gleitman, et al. in press). Performance on the syntactic alternation tasks not only requires speakers to choose one referent over the other for the initial position in a sentence; this preferential treatment inevitably leads to making choices between two or more structural variants equally applicable to the portrayed event (e.g., active/passive voice). Starting point studies are interested strictly in the speakers' choice of what to start the sentence with in the environment where no syntactic alternation is necessary (e.g., locative phrases). The reason for such contrast is primarily theoretical: While both types of studies usually make claims about the speakers' syntactic choices, in fact only the tasks used in the former group assess both positional and structural effects in sentence production. The use of starting point tasks rules out structural inferences leaving the researcher only with positional interpretation of the results. All experiments reported in this thesis use word order alternation tasks.

One recent study provided important support to perceptual accounts of word order alternation by (1) testing a whole array of structures and (2) using implicit attentional cueing protocol (Gleitman, et al., in press). Sentences with the verbs of perspective (*give/receive*), conjoined noun phrases (*The boy and the girl/The girl and the boy*), voice alternating transitive sentences, and symmetrical predicates (*The boy meets the girl/The girl meets the boy*) were elicited with the help of static pictures presented on a computer screen. Similarly to Tomlin's experiment, participants described visually presented events, but their attention was manipulated immediately before the target picture presentation. It was done by flashing a black square in the location of one of the event's referents for brief 75 msec. right before the picture of an event appeared on the screen. Once the picture was on the screen, participants extemporaneously described the presented event without any further manipulations of attention. The employment of such implicit cue was quite effective in directing attention to the cued referent. Also, an early preference to visually interrogate the visually cued referent observed in this study suggested initial syntactic biases toward the perceptually promoted structure. However, the resulting syntactic alternations were not nearly as strong as those reported by Tomlin. In all the tested structures the cued referent was indeed more likely to claim the starting position within the produced structure, but the magnitude of the structural alternation was rarely higher than 10%. In other words, the speakers' reliance on canonical event causality and the corresponding grammar was so strong that even when attention was directed to the non-preferred starting point, the likelihood of producing the less frequent structure starting with this referent was only 10% higher than regularly.

An opposition to the attentionally-driven proposals of word order alternation suggests that the production of names and structures does not immediately rely on the allocation of attention among the referents (e.g., Griffin, 2004a,b; Griffin & Bock, 2000). Griffin & Bock (2000), for example, revealed that speakers were sensitive to the perceptual information contained in the scene only when they prepared to engage in a linguistic task. On one hand, participants in this study were more likely to fixate the salient referents (as early as 300 msec) if they were going to describe the event. On the other, no such effect was found when the participants were instructed to silently search for one of the referents.

One conclusion from the latter study was that a direct bottom-up link from attention to language is unlikely. Instead, there is a reverse structure-dependent tendency for speakers to visually interrogate the scene as they describe it. According to this view, the extraction of the gist of the event, or *rapid apprehension* happens prior to incremental assignment of the slots in a sentence. Rapid apprehension typically mirrors the canonical event causality (e.g., agent-action-patient) and maps the resulting syntactic structure accordingly. The tendency to order information according to some "natural" pattern was found in tasks using non-linguistic communication channels, such as gesture (Gershkoff-Stowe & Goldin-Medow, 2002). Once rapid apprehension is complete, it is the commitment to a particular structure that predicts how speakers will interrogate the scene as they describe it. This pattern of results suggests a reverse link from language to perception. Whereas the mechanism put forward by Tomlin can be simply described as *"attend – assign subject – construct word order"*, the mechanism that can be derived from the results of Griffin and Bock is *"assign syntactic roles – attend as you construct the word order"*. In other words, it is not allocation of attention to the referents that influences positioning of sentence constituents; rather, it is the assignment of the sentential roles that drives preferential attention to the referents. Griffin (2004a) presented more evidence for the idea that syntactic production may occur relatively independent of direct attention to the elements in the described scene. One task in her study was to describe pictures of transitive events. Salience of one of the referents was manipulated by means of changing their sizes. Performance on this task demonstrated that, the participants (1) actively and rapidly examined the scene prior to producing a sentence about it; (2) the production of the sentence elements was preceded and mirrored by the speakers' gazes toward relevant entities; (3) finally, factors like animacy and thematic roles of the objects biased the speakers' preference to look at these objects and to use them as the grammatical subjects of the corresponding sentences. However, preferential looks to the salient referents were observed later in the task, during the linguistic preparation phase.

4.1. Conclusions. Chapter 1

The research reviewed above provides evidence for a regular link between visual attention and the syntactic organization of human discourse. In general, speakers' performance on visually mediated psycholinguistic tasks suggests that distribution of attention (1) may influence positioning of referents in a sentence, and (2) it can experience influence of the committed structure in return. However, perceptually-driven and structurally-driven accounts of interactions between attention and the assignment of grammatical roles in a spoken sentence are not as mutually exclusive as it may seem. The model proposed in Chapter 1 predicts a bidirectional set of interactions: The cognition-to-language pathway allows perceptual factors to bias initial apprehension of the observed event while the language-to-cognition pathway biases looks to the referents once the sentence production began. The research question then gravitates from *what comes first?* to *how operations at different processing levels interplay?*

Of course, in a regular discourse situation, the speaker has to take into account not only what visually stands out from the environment, but also the preceding discourse specifics. For example, increase in general availability status of the lexical material employed in the discourse inevitably leads to promotion of this material to a higher prominence status in subsequent sentences (see Chapter 2). How such promotion interplays with the preference to visually interrogate salient elements of the scene is largely unknown?

Finally, the attentional system of the human brain does not differ between speakers of different languages. Therefore for example, a longer exposure to the cue should lead to a stronger cueing effect on the choice of word order regardless of the language in use. However, the grammatical organization of the languages we speak may differ. As a result, one could expect that the ways attention interplays with the linguistic code may depend on the particular organization of the latter. For example, the susceptibility of an English sentence to perceptual priming may be different from that of a Russian sentence. This difference may come as a result of the fact that once the syntactic subject position in an English sentence is assigned to the focally attended discourse entity, the binary choice of grammatical voice can finalize the organization of the syntactic pattern toward either the active or the passive voice. Also, absence of active case marking in English grammar makes it easier for speakers to switch between the two available word orders once the subject of the sentence is assigned. However, perceptual priming may result in quite different pattern of interaction in the languages with flexible word order and active case marking. Experiments 1 and 3-5 will test the results obtained by Tomlin using languages dramatically different from English, providing better control of the allocation of attention by using eye-tracking methodology, and reporting a much wider variety of behavioral measurements.

Chapter 2. Conceptual effects in sentence processing

What determines the choice of a structure and the ordering of the sentence constituents during the production of utterances? The converging evidence discussed thus far points to the role of salience and the distribution of attention in planning and formulation of sentences. Speakers seem to actively take into account the attentional status of the referents in the scene when they decide what to say first and, depending on that, what structure to use to organize the sentence. However, a variety of other factors have also been found to influence the accessibility of words and, therefore, their ordering in sentences. Among those factors are those related to the referent's conceptual status: Novelty in discourse, animacy, definiteness, imageability, concreteness, and prototypicality. A typical way to investigate the role of one of these factors is to manipulate the corresponding property (e.g., animate vs. inanimate) of the referents expecting to observe a positional effect in the elicited sentences. Sometimes such manipulation is achieved using the *lexical priming*, for example, by presenting written or spoken words or sentences containing these words that promote one of the referents through repetition or association. *Referential priming,* on the other hand employs non-linguistic materials, such as pictures of entities and events in order to achieve the same effect.

All the studies discussed in detail below regularly report the existence of an independent tendency to assign referents of a more prominent referential status (old, animate, definite) to prominent syntactic positions, e.g., Subject NP. In order to formalize what makes one referent more prominent than another, many authors resort to the concept of *conceptual accessibility* (Bock & Warren, 1985). This concept becomes extremely important when one attempts to disentangle the positional (word order) and the structural effects resulting from manipulations of the prominence status of one of the event's referents.

2.1. Conceptual accessibility

No one really knows what *exactly* conceptual accessibility (CA) is. As a result, a variety of derived definitions that relate CA to "codeability", "imageability", "retrievability", etc. continue a circular definition practice. However, this concept is typically invoked in psycholinguistic studies in order to explain why some referents (or, more broadly, concepts) receive preferential treatment by the processor.

To arrive at the idea of CA, Bock and Warren (1985) used a sentence recall task. The materials were grouped into high-imageability and low-imageability groups based on Paivio's imageability norms (Paivio, 1969). Three types of structure were tested: (1) simple transitive structures (Active/Passive), (1) PO/DO dative structures (*The old hermit left the property to the university*), and phrasal conjuncts (*The lost hiker fought time and winter* vs. *The lost hiker fought winter and time*). Bock and Warren discovered that the structural variants that positioned the more accessible argument first were better recalled in the first two types of structures but not in the phrasal conjuncts. One important difference between the phrasal conjuncts and the other two structures is that the alternation of word order in phrasal conjuncts is not accompanied by changes in grammatical role assignment. The lack of CA effects in the phrasal conjuncts suggested that an increase in CA improves chances of the referent being assigned to a more prominent grammatical role in a sentence instead of provoking a simple word order positional effect.

The underlying idea behind the CA is that the referents that are "more thinkable" (Bock & Warren, 1985) tend to perform more prominent grammatical functions, and, therefore, appear in more prominent syntactic positions (usually at the beginning of a structure). Psychological explanations for CA are not as clear, however. The originators of the idea suggested that "Conceptual accessibility is the ease with which the mental representation of some potential referent can be activated in or retrieved from memory" (Bock & Warren, 1985, 50). A similar approach was taken in Sanford and Garrod (1981), who proposed that one important function of maintaining coherence in discourse is to constantly perform a successful search for discourse-relevant referents in the memory of the interlocutors. They called such referential situation a *scenario* and termed referents that form part of the current portion of the discourse, and, therefore, are actively maintained in the memory, more easily (or quickly) accessible than the referents that do not correspond to the current topic of discussion. The same idea is present, at least implicitly, in Levelt (1989), who related the production of referring expressions to the level of the accessibility in terms of the addressee's mental state.

In any case, more likely than not CA has to do with the language-related memorial activation status, so that the higher the activation of the referent in memory, the easier it is to retrieve. This status, however, directly depends on the allocation of limited attentional resources (see, for example, Kastner & Underlieder, 2000), so that the referent currently in focus does not need reactivation in the working memory, whereas the memorial retrieval involves more active deployment of attention. The linkage between the referent's CA and the focus of attention was advocated in a number of studies (Myachykov & Posner, 2005, for a recent review). Some recent data (Arnold & Griffin, 2007) support this claim. Using a story-telling paradigm Arnold and Griffin examined the process of choosing between pronouns and proper names in sentence production. The results revealed that even when a pronoun was not ambiguous in its relation to the referent's identity, the mere presence of another character decreased the chances of pronoun use and generally slowed down the access to the lexical form for the most prominent character. At the same time, the presence of another character in the preceding discourse generally reduced the chances of pronoun appearance in the following sentences. Based on these results, Arnold and Griffin suggested that the degree of CA varies as a factor of the attention allocated to each referent.

Another theoretical proposal about the nature of CA was made by Prat-Sala and Branigan (2000). They proposed a two-level understanding of CA claiming the existence of *inherent* and *derived* accessibility dimensions. The inherent accessibility is based upon the intrinsic properties of the concept, such as word frequency status, animacy, concreteness, and prototypicality. These are the features the concepts possess regardless of the interlocutors' intentions and the current discourse status of the corresponding referents. The derived accessibility is a temporary property of the concept that is dependent on the referent's current activation status in both linguistic and non-linguistic terms. The derived accessibility is driven by various means of priming. These two CA dimensions can overlap at any given time in discourse if the inherent prominence status is supported by the current, derived prominence status. On the other hand, the contributions of the inherent and derived accessibility forces can be contradictory if the prominence promoted by priming is not supported by the inherent prominence status of a referent. This idea is quite appealing as it introduces both the global and the local levels for the CA effects to appear. The model introduced in the last chapter of this thesis uses the bi-dimensional theory of CA quite extensively.

Evidently, making one referent more accessible than the other increases the chances of the former being mentioned early in a sentence (e.g., become the grammatical subject in active or passive sentences). In addition to imageability (Bock & Warren, 1985), different authors mention givenness (Arnold, et al., 2000; Bock, 1977), animacy (Altman & Kemper, 2006; Bock, Loebell, & Morrey 1992; Clark, 1966; Christianson & Ferreira, 2005; Ferreira, 1994; McDonald, Bock, & Kelly, 1993; Prat-Sala & Branigan, 2000; Sridhar, 1988), definiteness (Grieve & Wales, 1973), and prototypicality (Kelly, Bock, & Keil, 1986) among factors influencing the degree of referent's CA. The issue of imageability has been discussed in the paragraphs above. The importance of animacy, definiteness, and prototypicality as regulators of the referent's accessibility are not immediately relevant to the current research as these features were not independently manipulated in the reported experiments. The following two sections of this Chapter will discuss in detail the effects of givenness on the positioning of the referents in a sentence.

2.2. Givenness

2.2.1. Theoretical assumptions about givenness

Although it has long been noticed that the information flow in discourse can be divided into old or given and new elements, there is a number of theoretical approaches to the functional interpretation of givenness. In one sense, givenness represents the knowledge shared between the interlocutors. Therefore, the given information is the information that the speaker believes to be known by the listener. In contrast, the new information is the information the speaker is unfamiliar with (cf. Clark & Haviland, 1977; Halliday, 1967, Haviland & Clark, 1974).

Another view proposes that given/new distinction follows the referent's recoverability from the preceding context (Kuno, 1972). According to this view, if the information about the referent is recoverable from the preceding context, this referent is regarded as given. If such information is not recoverable, the referent represents a new discourse entity. McWhinney and Bates (1978) suggest their own interpretation of what influences the referent's newness relating this distinction to the amount of change the speaker tries to produce in the listener's mental state. In their view, a discourse entity is considered to be new when the speaker uses it to achieve a change in the information flow in the listener's working memory. By contrast, a discourse entity is considered to be given when no such change is attempted.

Finally, a quasi-psychological view on givenness suggests that given vs. new distinction correlates with the notion of cognitive activation of the concept. For example, Chafe (1976: 30) states that "Given (or old) information is that knowledge that the speaker assumes to be in the consciousness of the addressee at the time of the utterance. So-called new information is what the speaker assumes he is introducing into the addressee's consciousness by what he says." Implicitly, Chafe invokes the issue of salience and memorial activation in his interpretation of givenness. In order to explain what makes the referent given or new he suggests that the new information is "newly activated" at a given point in conversation, while the old information is the one that does not require such activation (Chafe, 1994, 72). Such activation status, among other things, depends on the speaker's perspective influenced by the salience of the processed material.

Givon (1992) proposed probably the most detailed picture of how interactions between the attentional properties of the discourse-related material and the activation of the referents in the speaker's memory can occur. Givon views grammar in general as a routinized processing system, where highly conventionalized structural cues trigger automatic cognitive processing responses. Using the idea that attentional operations and activation in working memory are strongly related processes (e.g., Erickson & Kintsch 1995, Kastner & Ungerlieder 2000), Givon argued that the mechanism behind the maintenance of the referential coherence can be described in terms of *mental processing instructions.* This mechanism operates on discourse nodes or files that are subject to constant attentional control and higher or lower activation in working memory. In this system, referential prominence correlates directly with the attentional activation.

According to Givon, main aspects of the discourse processing system are the following:

1. Attention is a limited resource; therefore, activation of a referent in the current discourse will depend on how much attention is currently being paid to it;
2. Processing of visual stimuli includes components of disengagement, move, and reengagement (Posner & Petersen, 1990). These components are rough correlates of (b) and (c) in Givon's system of four major operations in the attentional system manipulated by grammar of referential coherence:

(a) continue activation of the current open file;

(b) terminate activation of the current open file;

(c) activate a currently inactive file

(i) open-activate-a new file;

(ii) reopen-reactivate-an existing file (Givon 1992, 23).

3. Covert attention is influenced by grammar-guided discourse processing and involves phenomena like word order and syntactic structure of the clause;

4. The attentional system is subject to conscious high-priority assignments. This aspect of executive attention is recognized by Givon in grammar as a "gating" system (cf. "windowing" of attention in (Talmy, 2000)). Conscious attention in Givon's view assigns priority to incoming information" (Givon 1992, 42).

Regardless of their theoretical complexity and reliance on either a linguistic or a psychological tradition, one common element in all these proposals is the status of the given information as the one previously introduced in discourse, while the new information is considered novel to the interlocutors. I will follow this parsimonious interpretation throughout the rest of this thesis.

2.2.2. The control of ordering through givenness

2.2.2.1.Lexical priming and givenness

In order to explore how givenness affects the positioning of a referent in a sentence, some researchers employed lexical priming. This scenario typically involves a sentence recall or production task, in which participants are exposed to a word making one of the discourse referents primed for subsequent processing (e.g., Bates & Devescovi, 1989; Bock, 1977; Flores D'Arcais, 1975; Tannenbaum and Williams, 1968).

In one of the earliest studies, Tannenbaum and Williams (1968) used text vignettes repeatedly focusing one of the sentence constituents – subject or object – from sentence to sentence. These two priming conditions were contrasted with the control condition, in which neither constituent was primed. After reading the vignettes, participants were asked to produce novel active or passive voice sentences while inspecting the pictures of the events these sentences related to. Active voice sentences in the control condition were initiated on average faster than the passive voice ones, but the greatest reaction time benefit occurred when active voice sentences were produced in the subject-priming condition. Also, although passive voice sentences produced in the object-priming condition were still slightly slower than their active voice counterparts, this difference was greatly diminished.

In another study (Bates & Devescovi, 1989) participants were presented with a short film strip depicting a simple event (e.g., a hippo hitting a ladder), after which a participant was asked a question, which highlighted one referent in favour of the other (e.g., *Tell me about the hippo* or *Tell me about the ladder*). If the hippo was the lexically primed referent, the participants produced sentences like "*The hippo knocked over the ladder"* in 100% of cases. If the cued referent was the ladder, they tended to produce sentences like *"The ladder was knocked over by the hippo"* in about 70% of the cases. Evidently, interrogative lexical priming of either the agent or the patient of the event led to robust alternation of the grammatical voice in the participants' responses. A very similar approach was taken in another study (Bock, 1977). Participants first heard short sentences (e.g., *A psychologist cured a neurotic poodle*). Then participants had to answer a question with a preamble that cued one of the event's referents, for example, *the poodle*

(*The interior decorator was afraid she would have to get rid of her neurotic pet poodle because it was ruining the furniture, but she was able to keep it after all. What happened?*). As a result, participants tended to place the cued entity first in their answers to the questions, and, for the suggested example, produce sentences like *The neurotic poodle was cured by a psychologist*.

A variant of the sentence recall task proved to be efficient in another study investigating givenness effect (Bock & Irwin, 1980). However, the experimental materials in this study contrasted the primes that were lexically identical to the targets and those that were semantically but not lexically related. In experiment 1, participants heard lists of questions and answers about the event previously introduced by a single sentence (e.g., *The rancher had a stallion who kept running away*). Each question established one of the target referents as the given information and the other target as the new information (e.g., *What did the stallion do?*). The given noun then appeared in the answer list as either a lexically identical (*the horse*) or a semantically related (*the stallion*) referent. Also, the answers were designed as *appropriate* if they ordered given before new and *inappropriate* – if they placed new before given. After listening to the answer list, participants heard the question again and then wrote down the answer they recalled from the list. Sometimes the recalled answers were written down in their exact form, and sometimes the arguments in the answers were *shifted* in order to move the given referent in front of the new one. Bock and Irwin observed that shifting of the arguments occurred more often in the inappropriate answer condition, so that the given-new order was re-established in conflict with the order of the original target sentence. Also, such shifting occurred in both identical and related referent pairs, but this effect was stronger in the identical pairs. Hence, both referential and lexical accessibility provided the givenness effect in this study. One consequence of the Experiment 1 results that will become more important later is that while the lexical component of givenness probably stems from the levels of processing almost exclusively related to language, referential effects can be at least partially ascribed to the non-linguistic levels of processing. In Experiment 2, participants heard a short sentence introducing two entities (e.g., *The falling tree crushed the lumberjack*) and then were prompted to respond by the presentation of a single word that was identical to one of the referents in the previously encountered sentence (e.g., *tree* or *lumberjack*). Similarly to the previously discussed studies, such single word priming resulted in a tendency to recall the target sentence in a form that made the primed word the subject. Together, these two experiments provide support for both lexical and referential accessibility components in givenness.

The use of a virtually identical task in a study with a flexible word order language – Japanese (Ferreira & Yoshita, 2003) revealed very similar pattern of results: speakers tended to produce canonical and scrambled Japanese sentences so that the given referents were promoted to the frontal position in a sentence. Just like in Bock and Irwin (1980), this effect was more pronounced for lexically identical than for semantically-related arguments. One recent study used eye-tracking to explore the given/new contrast in another flexible word order language – Finnish (Kaiser & Trueswell, 2004). Experiment 1 in this study used a self-paced reading task together with manipulations of the context preceding the target sentence. Analysis of a variety of measurements revealed that difficulties typically associated with processing non-canonical sentences in Finnish (OVS, for example) are alleviated when the preceding discourse strongly supports patient-focused context. Experiment 2 used a Visual World task. Participants listened to the sentences while observing the visually presented events. Upon hearing OV-sentences, listeners made anticipatory eye movements toward a discourse-new referent in the scene; the reverse pattern was true for SV-sentences. The findings from Japanese and Finnish demonstrate that speakers of languages grammatically different from English are also sensitive to the given/new contrast.

Ferreira (1994) demonstrated that the choice between active and passive voice construction may not only depend on the difference between the new and the given status of the referents, but also on the type of verb used to describe a transitive event. In order to elicit active and passive voice sentences, Ferreira used a sentence completion task, in which participants read two nouns and a verb, and they had to produce a sentence using them. Sometimes the presented verb was "normal", thus inducing a regular sequentionality like *agent-theme* (e.g., *avoided*); sometimes the verb lead to a reverse interpretation of the event – *from theme to experiencer* (e.g., *challenged*). An additional manipulation was the animacy status of the presented nouns. The results of this study indicated that passive voice constructions were more likely to be used with "abnormal" verbs that promote the experiencer to a more prominent position. At the same time, the data revealed that the passives occurred more frequently when the referents differed in their animacy. Similarly to the previously discussed studies, Ferreira (1994) demonstrated that the use of syntactic alternatives (e.g., passivization of transitive sentences) promotes a more prominent referent to a more prominent syntactic slot in a sentence (e.g., Subject).

A combination of givenness and animacy effects was further explored in a cross-linguistic study using English and Spanish (Prat-Sala & Branigan, 2000). In Experiment

1, participants listened to a short story about two inanimate characters (e.g., *a swing and a scooter*) ending with the question *What happened?* The characters in the text were introduced as either given or new. After listening to the text, participants had to describe a picture of an interaction between the two characters (e.g., *a swing hitting a scooter*) using a single sentence. The results demonstrated that in both English and Spanish. Speakers tended to produce sentences that placed the given referent in a more prominent syntactic position. Experiment 2 used an animacy manipulation. The pictures were almost identical to the ones used in Experiment 1 except the patient in the picture was always animate. The preamble stories introduced either the animate or the inanimate referent as given. The givenness effect in this experiment overrode the effect of animacy although the tendency to use givenness as the syntactic cue was more pronounced for the animate referents. In relation to Prat-Sala and Branigan's interpretation of CA as a two-dimensional processing system with inherited and derived components, the effects of givenness have to belong with the derived accessibility whereas animacy represents the inherited accesibility.

Flores D'Arcais (1975) also used a single word priming paradigm in a study with Italian. Participants heard an Italian word that corresponded to an entity (say, *cat* or *dog*). Then, they were asked to describe a picture of a transitive interaction involving a primed entity and another entity (e.g., dog attacking cat). The results were virtually the same as those of Bock & Irwin (1980): If the primed referent was the cat, the participants tended to produce sentences like *"The cat is attacked by the dog"* in 67% of such cases. In a dog-primed condition, responses like *"The dog attacked the cat"* were observed in 77% of the collected utterances.

A less explicit approach was taken in Bock (1986a), which continued using single word priming tasks to explore the effects of semantic priming on the positioning of constituents in a sentence. She presented participants with a priming word such as *thunder* or *worship,* which preceded a picture of lightning striking a church. Importantly, none of these primes were lexically related to the target event, while both are semantically related to either lightning or church. Participants had to describe the target event extemporaneously. They were more likely to describe the scene with an active voice sentence (e.g., *Lightning is striking the church*) when primed by *thunder*. If they were primed by the word *worship*, participants tended to describe the picture with a passive voice sentences (e.g., *The church is being struck by lightning*). As with Bock and Irwin (1980), primes that were lexically unrelated but sharing semantic content with the targets promoted the activation of a primed referent to the earliest syntactic position in the sentence. Although it is not clear whether such semantic priming establishes its effects through givenness, such an explanation is quite possible. It is also unclear whether such associative priming works only through the starting point effect or somehow affects the assignment of the structural relations. For example, although the word *worship* is related to the word *church,* only the latter is unequivocally a noun. Therefore, it is not likely to affect grammatical role assignment.

Finally, a report that attempted to relate conceptual effects in sentence production to the issues of salience and perception comes from Osgood and Bock (1977). Their study used lexical sentential primes, in which various "salience" properties of the referents were systematically varied in order to elicit a number of English constructions: prepositional object dative constructions, voice-alternating constructions, genitive constructions, etc. Prior to the experiment, independent judges rated the experimental materials against a three-dimensional referent feature space. The first dimension – *naturalness* – referred to the tendency of referents to appear in the order mirroring the canonical event causality sequence (e.g., Agent-Action-Patient). The second dimension – *vividness* – reflected higher conspicuousness of some semantic features of a referent over others (e.g. the vampire vs. the man). These two dimensions were regarded as *inherent,* because the features related to them are naturally present in the referents themselves. The third dimension was the *motivation of the speaker.* This *attributed* feature of the salience (e.g. interest, concern, perspective) referred to the speaker's endogenous interest to order the referents according to some internally conceived mental plan. The analysis of the experimental data revealed that (1) speakers tended to use naturalness as the main determinant of the order of mentioning in the sentence (in other words, they relied heavily on the natural event causality and the canonical grammar of English), (2) Agents were more likely to be mentioned before Patients, and that (2) referents of a higher vividness status were more likely to be mentioned before the ones of a lower vividness status. Osgood and Bock contrasted their findings to the well-known tendency of the old discourse information to appear before the novel material (see below) suggesting that control of the ordering through givenness is not as powerful as the same process driven by the factors related to "naturalness" and "vividness". One problem with this interpretation is that the "perceptual" properties of the referents, like vividness were derived solely from the lexical ratings. Whether such vividness reflects a tendency for a preferential perceptual treatment of corresponding world referents is not at all clear.

2.2.2.2. Referential priming and givenness

The effects of givenness on the ordering of the constituents in a sentence were not only observed in studies using lexical priming. As the following paragraphs illustrate, tasks that employed non-lexical priming often lead to the same results: The referent made more accessible through the preview preceding the presentation of the target picture makes this referent more likely to be included into the sentential frame first. Although there are only a limited number of such reports, they are important for the purposes of this thesis as the conceptual manipulation used in Experiments 6-9 relies upon very similar set of principles.

One of the earliest report using referential priming paradigm (Prentice, 1967) was motivated by the earlier finding (Carroll, 1958) that the likelihood of using active voice versus passive voice sentences in description of transitive events varies 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*). Some of the characters were human beings, some – animals, and some – inanimate objects (e.g., *flower pot*). These pictures were paired with cue slides depicting one of the event's characters: the agent or the patient. Participants first viewed the cue picture and then the whole event, to which they provided spoken descriptions. The results of this simple experiment demonstrated that speakers were more likely to place the primed referent first in their descriptions of visually presented events. As far as animacy of referents is concerned, participants were more likely to passivize their sentences when both referents were humans and slightly less likely when one of the referents was an animal and the other human.

A very similar priming technique was employed by Olson and Filby (1972) in their study with adult speakers. Olson and Filby used a sentence verification task. In a typical variant of this task, a sentence is presented together with the picture, and participants have to verify whether the sentence felicitously describes the event portrayed in the picture. The pictures in Olson and Filby's task represented simple transitive interactions between the inanimate entities (e.g., *a car hitting a truck*). Prior to the picture presentation, a picture of either the truck or the car appeared on the screen. Exposure time for the picture was 650 msec., which is a substantial time to be able to recognize the referent and probably even access the lexical name of it. As a result of this manipulation, participants were faster to verify active voice sentences corresponding to the target event when the primed referent was the agent. A reverse pattern was true for the passive voice sentences: participants were faster to verify them as related to the target picture when the primed referent was the patient. In a question-answer version of their task (cf. Bates & Devescovi, 1982; Bock & Irwin, 1980), Olson and Filby presented participants with the same event pictures but manipulated their agent- or patient-based perspective by asking questions like *What about the truck?* The results once again were very similar: Participants tended to use active voice sentences to describe the target event when they were asked about the agent, and they were more likely to use the passive voice sentences when they were asked about the patient.

Turner and Rommetveit (1968) conducted similar research with children. They presented participants with the sentences and later asked them to recall these sentences. The materials were all active/passive voice sentences divided into four groups: (1) non-reversible actives, (2) reversible actives, (3) non-reversible passives, and (4) reversible passives (see examples in Table 1). One sentence of each type was randomly paired with a picture of one of the following: The agent, the patient, the whole described event, or the blank. Presentation of one of the referents made that latter referent more prominent (or given) compared to its counterpart. The presentation time varied as the experimenter presented the pictures manually; however, it was always long enough for a participant to identify the referent. The pictures were shown to the participants both at the time of the sentence storage and recall. Turner and Rommetveit then analyzed the recalled sentences in terms of the correct vs. incorrect responses and in terms of the syntactic transformations that occurred as a factor of which referent was cued. The results demonstrated that (1) active voice sentences were better remembered than the passive voice ones, (2) the active voice sentences were more likely to be recalled correctly if the primed referent was the agent, while the passive voice sentences were better remembered if the primed referent was the patient, (3) the non-reversible sentences were remembered better than the reversible ones, (4) sentence reversibility interacted with the picture type so that the non-reversible pictures were remembered better than the reversible ones when the primed referent was the agent, and (5) the passive voice sentences were more likely to be transformed than the active ones except for the cases when the primed element was the patient. Together these results demonstrated that (1) processing passive voice sentences (at least for children) is more difficult than processing active voice sentences and that (2) making one of the referents more prominent (through givenness enhancement) facilitates the processing of the syntactic structure that uses this referent as its subject.

While the three experiments discussed thus far involved transitive sentence processing, Clark and Chase (1972) used a referential priming paradigm to investigate preferential ordering of the constituents in the locative sentences. In Experiment 3 of their study, Clark and Chase used a variant of the sentence verification task, in which participants were to compare locative sentences, like *Star is above plus* to pictures and decide whether the sentence and the picture were felicitously related. At the same time, they were instructed to direct their attention to either the top or the bottom of display. This attentional manipulation resulted in faster and more accurate interpretation of the sentences that started with the referent in the cued location (cf. Forrest, 1996).

One problem with interpreting the results of studies using referential priming paradigm was already known to Prentice (1968). Speaking about the observed tendency to alternate word order as a function of the primed referent, she notes:

> "Whether the change occurs first at a verbal level, in tendency to name, or at perceptual level, in interpretation of the event, is the question at hand. Does the subject sometimes name the cue element before he "knows" which slide he will describe? Alternatively, does the subject perceive an event differentially as a function of cue?"

Prentice herself favoured the idea that the cue to one of the referents triggered changes in the deep syntactic structure not in the conceptual analysis of the event. This interpretation is later echoed in Tomlin (1995), who also supported the direct link from perception to the assignment of the syntactic positions in a sentence. However, to tease apart the perceptual and the conceptual effects using a study, in which visual and referential priming are confounded is impossible simply because both properties are present in the same cue (e.g. Prentice, 1968) or because both the target and the cue are presented together (Tomlin, 1995). As a result, the experimenter's interpretation is left with the situation when she cannot (or should not) exclusively favour either perceptual or conceptual priming explanation of the word order alternation. The question whether the referential priming effects on the choice of word order are due to the elevated perceptual salience of the previewed referent, the establishment of the given/new contrast or a combination of both effects remains open. Chapter 6 will help address the issue of perceptual versus referential priming of the word order by comparing the two studies (Experiments 6 and7) that use identical cueing parameters (location and duration) but employ different priming paradigms.

2.3. Conclusions. Chapter 2

Chapter 2 provided a comprehensive discussion of the studies that demonstrate how increasing conceptual accessibility through lexical and referential priming elevates the chances of a referent assuming a prominent position in a sentence. Many reports relied on the *starting point* hypothesis – the idea that the fluctuations of the CA status correlate with the ability of the referent to appear at the beginning of a sentence. In many cases such preferential positioning does not require a change of the grammatical structure that goes with it (e.g., locative sentences). In other cases (e.g., voice alternations), the assignment of the referent as the sentential starting point entails an immediate and possibly costly choice between the available structural options. The difference between the positional and the structural effects that can both potentially result from the CA manipulations is very important. While the former has to do with incremental word order assignment, the latter may rely on a certain degree of structural preplanning. The experiments discussed in Chapters 6 and 7 attempt to cast some light on this issue.

Although a battery of studies supported the given-before-new account of word order priming some issues remain unclear. For example, it is uncertain whether exposing the primed referent for duration of time insufficient for its recognition will result in much weaker word order alternation effect? Hypothetically, the result of such "implicit referential" priming might be reduced to perceptual factors: If speakers do not have enough time to process the identity of the referent, the cue becomes functionally indistinguishable from a strictly perceptual one. At the same time, research in rapid object and scene recognition demonstrated that people can be quite good at recognizing the events and their referents even at very short presentation times. Although psycholinguistic evidence for people's ability to name things under implicit presentation conditions is limited, one study (Dobel & Gumnior, 2004) revealed that viewers were able to recognize and name objects at 75% accuracy starting with 200 msec presentation window, and properly identify and name "who-did-what-to-who" with over 50% accuracy at 150 msec stimulus duration. Together, these reports suggest that referential priming at short presentation times may still add a conceptual boost effect to the main of the cue location.

Second, it is unclear how perceptual and referential priming effects interact during sentence processing. From methodological point of view, any direct comparison of the existing literature is impossible because the parameters of the experimental protocols are too different. From theoretical perspective, it is also difficult to reconcile the attended-first with the given-before-new hypothesis. Bock, Irwin, & Davidson (2004) provide a comprehensive account of this theoretical controversy. They justly note, that "the focused first" and "the old first" proposals are contradictory because the information that attracts the focus of attention is typically the new elements of the scene, whereas givenness promotes the already established background. The lexical-semantic factors (e.g. old-before-new) and the perceptual factors (e.g. focused first) should, therefore, produce competing effects. The second problem is what Bock and colleagues call the *levels problem.* Even if one were to assume that perceptual factors in fact predict the assignment of the syntactic roles, it is unclear what level of representation –perceptual, conceptual, or linguistic – these effects are particularly active at. Finally, there is a *language variation problem.* The latter derives from the fact that different languages impose different grammatical restrictions on what can be placed first. The grammatical properties of a given language can be understood as a set of automatic constraints (e.g., canonical word order), while violations of such constraints represent processes under voluntary control of the speaker (e.g. regular passivization). Then, the choice of a particular syntactic structure in a particular discourse situation may reflect a blend of automatic and controlled processes. The exact instantiation of such a blend may differ from language to language. A good candidate for the fluent regulator of the interplay between automatic and controlled processes is the language syntax (Bock, 1982). As such a regulator, the language syntax is organized around a set of automated defaults and a set of alternative options that require deeper processing with the allocation of more cognitive resources. Experiments 6 and 7 explicitly address these issues by manipulating both (1) the explicit/implicit property of the cue and (2) the perceptual/referential nature of the cue and by (3) using languages other than English.

So, it seems likely that cueing speakers' attention toward one of the event's referents improves chances of the latter becoming the sentential stating point, therefore biasing the organization of the resulting sentence. At the same time, exposing participants to one of the event's referent's name or picture promotes the use of the syntactic structure that includes the primed referent as its subject. On the other hand, speakers across languages also have to rely upon the syntactic environment of the constantly changing discourse in addition to tracking the perceptual and the conceptual properties of the entities and the events they talk about. The next chapter discusses how the previously encountered syntax influences the structural choices made by the interlocutors in the unfolding discourse. Such phenomenon, known as syntactic or structural priming (Bock, 1986b) enjoyed both the theoretical and the experimental dedication in the psycholinguistic community.

Chapter 3. Syntactic effects in sentence processing

The previous Chapter helped to establish that processing a semantically or lexically related word or preview of a picture representing a referent can facilitate the subsequent placement of this referent in an upcoming sentence. This effect, typically referred to as *priming* effectively enhances the concept's activation and results in its preferential treatment in subsequent processing. In word-to-word priming, activation of a concept from the same semantic field (e.g., *worship*) successfully activates a wide range of related concepts from both the same and different word classes (e.g., *church*). In word-to-structure priming, the same activation logic explains why being primed by the word *worship,* speakers tend to describe a picture of a lightning striking church by using the structure that uses *church* as its subject (Bock, 1986a). However, despite some existing evidence (e.g., Bock & Irwin, 1980), it is still unclear whether such priming effect simply leads to the positional preference to use the primed word or concept before the other words or concepts or it influences the grammatical role assignment and, through such assignment, the syntactic composition of a sentence.

It would of course be logical to suspect that not only conceptual processing relies on the existing preactivation of the referents and words in discourse but both parts of the syntactic structure and the whole sentence syntactic frames might be able to influence the likelihood of their re-use in the subsequent discourse. It has been found to be true. Processing a certain syntactic structure within a sentence affects the ease of processing of the same or related structure in the subsequently comprehended or produced discourse. Such effect is known as *structural (*or *syntactic) priming* (Bock, 1986b).

Recent reviews provide comprehensive account of the abundant literature on the matter (Branigan, 2007; Ferreira & Bock, 2006; McLean, Pickering, & Branigan, 2004; Pickering & Branigan, 1999). In a very short summary, structural priming was observed in speaking (e.g., Bock, 1986) and writing (Branigan, Pickering, & Cleland, 2000; Hartsuiker & Westenberg, 2000). It is persistent in language production (e.g., Bock and Loebel, 1990), comprehension (Branigan, Pickering, & Cleland, 2000; Branigan, Pickering, McLean, 2005; Scheepers & Crocker, 2004), between these two processing modalities (Bock, et al., 2007), and in dialogue (Branigan, Pickering, & Cleland, 2000). It has been confirmed by using experimental (e.g., Bock, 1986), corpus-based (Gries, 2005; Szmrecsanyi, 2005), internet-based (Corley & Scheepers, 2002), and naturalistic studies (Weiner & Labov, 1983). The effect of structural priming was observed using different kinds of structures (Cleland & Pickering, 2003; Ferreira, 2003), in different languages (Hartsuiker & Kolk, 1998; Scheepers, 2003; Yamashita & Chang, 2006), and between two languages in a bilingual mind (Desmet & Declercq, 2006; Hartsuiker, Pickering, & Veltkamp, 2004; Loebel & Bock, 2003). Structural priming affects sentence processing both in adults' (e.g., Bock, 1986) and in children's (Huttenlocher, Vasilyeva, & Shimpi, 2004; Savage, et al., 2003; 2006) speech. Finally, structural priming seems to persist over considerable portions of discourse (Bock & Griffin, 2000). As structural priming is characterized by such a ubiquitous presence, several theoretical accounts, each trying to explain what structural priming is and what it is for, have been proposed. While some view structural priming as a short-term memory residual activation phenomenon (Pickering & Branigan, 1998; Pickering, et al., 2000), others propose that it in fact reflects fundamental processes of language acquisition through implicit learning (Chang, et al., 2000) or a vehicle of automatic coordination between interlocutors (Pickering & Garrod, 2004). I will briefly discuss the essence of these accounts and the existing experimental evidence supporting or contradicting them.

3.1. Autonomous structural priming

Research in the structural priming was motivated by the earlier work that demonstrated that the segments of the syntactic structure are actively repeated by the speaker from sentence to sentence (e.g., Schenkein, 1980). For example, Levelt and Kelter (1982) investigated syntactic repetition in discourse using question-answer task. They found that the question's surface form influences the form of the subsequent answer. For example, if participants were asked a question like *At what time do you close?* They were more likely to provide an answer like *At five o'clock.* Similarly, a non-prepositional form of the same question led to a non-prepositional answer. Their suggested interpretation of this effect was that such repetitive use saves interlocutors' cognitive resources by making speech processing more economical. Instead of constant consultations with the long-term grammatical knowledge, communicators, therefore, simply recycle the syntactic material currently activated in unfolding discourse. It is important, however, to note that Levelt and Kelter believed that the repetition advantage is limited to the tendency to repeat words (prepositions in their study) and not the structural frames.

Another study that prompted a similar theoretical account was conducted by Weiner and Labov (1983). Using sociolinguistic interviews, Weiner and Labov found among other things that an important factor influencing the speaker's choice between active and passive voice sentences was the presence of a sentence in the same syntactic form in the preceding 5 clauses. They interpreted such speakers' dependence on the previously used syntax as a result of the "mechanical" syntactic effect, which helped speakers maintain automaticity in spontaneous speech.

Although both aforementioned reports hinted at structural repetition component present in their results, arguably the first study that systematically addressed the issue in controlled experimental settings was Bock (1986b). In her study, Bock used a picture description task. Prior to describing, for example, a pictures of a transitive event (e.g., *lightning striking a church*), participants produced a priming sentence semantically unrelated to the target picture but presented in either the active or the passive voice form (e.g., *The referee was punched by one of the fans*). As a result of this priming manipulation, participants were more likely to repeat the grammatical voice from the priming sentence in the sentence describing the target event. Experiments 2 and 3 of the same study further tested whether the syntactic repetition effect was due to any overlap in the conceptual structure between the prime and the target. Both experiments manipulated the humanness of the agents between the priming sentences and the target pictures. Such conceptual overlap did not interfere with the priming effect. Based on these results, Bock concluded that structural repetition is a relatively encapsulated syntactic process and interpreted her results as consistent with the idea that syntactic processing constitutes an independent operational system within the linguistic code.

Some later studies supported the tendency to reproduce the syntactic structure regardless of the differences in the conceptual content. One study (Bock, 1989) addressed the issue of the closed-class words as possible enhancers of the Structural priming effect. If, as many believe (e.g., Garret, 1982), closed-class words form an indispensable part of the syntactic structure, manipulating this parameter might disrupt the priming effect. The same picture description paradigm as in Bock (1986b) was used. Participants had to describe pictures after being primed by a sentence that contained either closed-class or open-class words. The results revealed that there was no tendency for the priming effect to increase as a factor of sharing closed-class words between primes and targets. Another study (Bock & Loebell, 1990) supported the "modular" account of the Structural priming phenomenon by testing whether sharing event structure between primes and targets interacts with the priming effect. Once again, a picture description task was used together with production-to-production priming protocol. The priming materials were the sentences of either prepositional dative (e.g., *The wealthy widow gave an old Mercedes to the church*) or prepositional locative (e.g., *The wealthy widow drove an old Mercedes to the church*) construction. The target pictures depicted dative events referentially distinct from the primes with three participants: the agent, the theme, and the beneficiary. Therefore, in half of the trials the event structure was shared between the primes and the targets; in the other half only syntactic structure was shared. The results of Experiment 1 clearly demonstrated that conceptual overlap between primes and targets was not more likely to produce the priming effect than just syntactic overlap. Experiment 2 supported this result by testing past tense passive and past tense locative constructions (e.g., *The 747 was alerted by the airport's control tower* vs. *The 747 was landing by the airport's control tower*).

Similar findings supporting independence of structural priming from lexical and conceptual structure were observed in a recent study on the sentences using optional complementizer *that* (Ferreira, 2003) and a study on relative clause attachment

(Scheepers, 2003). For example, the latter study established that language producers tend to repeat the relative clause attachment patterns driven solely by the structural properties of the priming sentences but not due to the repetition of the referential scheme. Also, one study using Wernicke's and Broca's aphasics (Faroqi-Shah & Thompson, 2003) showed that both types of aphasics could not produce passive sentence well even when lexical cues, such as relevant nouns and uninflected verbs were provided. On the other hand, when the verb cues were given to the aphasic speakers with auxiliaries and tense morphemes, production of passive sentences reliably improved. This result supports that syntactic planning is a somewhat independent component of sentence processing.

3.2. Lexically mediated structural priming

The reports discussed so far suggested that structural priming relies on an independent syntactic stratum that is being shared between the priming and the target sentences. The effects of structural priming seem, therefore, to be quite autonomous from the referential and the lexical environment, in which sentence processing occurs. However, a number of other studies provided evidence for a contrasting account – the one suggesting that the structural priming effect may, at least partially, rely upon the repetition of the lexical-semantic information from sentence to sentence.

For example, Pickering and Branigan (1998) demonstrated that the magnitude of the structural priming effect can vary dependent on the repetitive use of the same lexical material from the priming to the target sentence (e.g., the verb). In a series of experiments, Pickering and Branigan used sentence completion paradigm where participants first had to complete priming sentences with the initial fragment given to them (e.g., *The racing driver SHOWED the torn overall…*or *The racing driver GAVE the torn overall…*), and then they also had to complete target sentences like *The patient SHOWED….* Sometimes the verbs used in primes and targets were the same, and sometimes they differed. Also, various parameters of the verbs used in primes and targets were manipulated including their tense, aspect, and number. The results revealed that priming occurred both when the primes and the targets used the same verbs and when the verbs were different. Also, the magnitude of the priming effect did not depend on whether the verbs in primes and targets shared their tense, aspect, and number parameters. However, a reliably stronger priming effect was observed when the verbs were repeated from sentence to sentence. Hence, repetition of lexical information from sentence to sentence may result in a *lexical boost* effects that successfully modulates the strength of structural priming.

The verb-related lexical boost effect has later been replicated in a study using confederate priming paradigm (Branigan, Pickering, & Cleland, 2000). Confederate priming simulates a dialogue situation in experimental settings so that one of the dialogue parties (the experimenter's confederate) uses scripted text, whereas the other participants remains naïve simply producing sentences in response to the confederate's ones. Participants in the current study used a set of pictures portraying ditransitive interactions (e.g., *A cowboy offering a robber a banana*). The confederate described her picture using a provided sentence varying PO/DO dative constructions (*A cowboy is offering a banana to a robber/A cowboy is offering a robber a banana*), and the naïve participants had to describe a card with another ditransitive event. The cards in the naïve participants set were ordered so that sometime the verb necessary to describe the event was the same as the one just used by the confederate and sometimes these verbs differed. Hence, a manipulation similar to that of Pickering and Branigan (1998) was used. The results of this study confirmed both the Structural priming and the lexical boost effect in a simulated dialogue environment. These results are important for two reasons. They suggest that (1) structural priming is quite ubiquitous as it persists in both monologue and dialogue, and (2) they provide further support for the existence of verb-related lexical boost effect.

Corley & Scheepers (2002) further validated the existence of verb-modulated Structural priming by using the set of materials from Pickering & Branigan (1998) in an internet-based study. Participants had to complete PO/DO dative sentences, in which with and without verb overlap between primes and targets. The study replicated the original results for both the structural priming and the lexical boost effects using World-Wide Web instead of thoroughly controlled experimental setting. In fact, the lexical boost effect from verb repetition was twice as high as in Pickering and Branigan (1998). Also, the name onset latency analysis demonstrated that people are not only susceptible to structural priming categorically, by means of repeating the sentential syntax; they are also faster to provide linguistic responses when producing structurally primed sentences. A similar reaction time advantage to produce syntactically-primed sentences was observed in another study (Smith & Wheeldon, 2001) by using both sentence and picture priming protocols. Interestingly, the same study provided evidence for the idea that the structural priming effect can be achieved by the presentation of non-linguistic materials – pictures of semantically different but structurally similar events. However, the observed priming effect seems to result from purely syntactic correspondence between primes and targets and not the similarities in conceptualizations of the described events. Also, the last experiment in a series demonstrated that the priming effect is localized to the generation of the first phrase in a sentence prior to speech onset.

The verb dependence of PO/DO priming seems to be particularly strong in sentence comprehension. One recent study (Arai, van Gompel, & Scheepers, 2007) used a visual-world task to demonstrate that verb-repetition between primes and targets is a necessary prerequisite for PO/DO priming in sentence comprehension. Following the idea advocated in earlier studies on verb repetition and structural priming, Arai and colleagues designed two eye-tracking experiments, in which participants had to listen to PO/DO sentence while looking at the arrays of pictures containing the referents of the comprehended sentences interspersed with unrelated distracters. The duration and the proportion of looks to the beneficiary and the theme of the event were clearly influenced by the PO/DO primes. In other words, participants were more likely to preferentially interrogate a ditransitive event's theme if the preceding prime contained a PO rather than DO construction. Most importantly, this was only true when the verbs were repeated between primes and targets. When the verbs were different, no effect of structural priming was observed.

Finally, one more study (Melinger & Dobel, 2005) managed to achieve the structural priming effect by using single verb presentation. German and Dutch participants read single verbs that were restricted to either PO or DO construction use. A verb presentation was followed by the display of an event that could be described by using either of the two constructions. Such single verb exposure successfully constrained speakers’ tendency to produce one of the two structures, which further supports lexically-motivated explanations of structural priming

However, the lexical boost effect is not limited to verb repetition. Cleland & Pickering (2003), for example, found robust lexical boost effects in by repeating the structure and the lexical-semantic content of Subject noun phrases. Speakers in this study were more likely to repeat the syntactic structure from priming to target sentences if both sentences contained a pre-nominal adjective (*the red square*) Subject NP than when the Subject NP was a post-nominal relative clause (*the square that’s red*). Also, they showed that repeating the head noun (*square*) between prime and target further enhanced the priming effects. Finally, the same tendency was observed when the head nouns in primes and targets were lexically different but semantically related (*goat and sheep*). The latter finding is very important because it provides evidence for the idea that not only lexical overlap but also sharing conceptual information between primes and targets can modulate the magnitude of the priming effect.

The reports discussed in this paragraph provide cumulative evidence for the idea that structural priming may not only depend on the repetition of the syntactic frame of sentences but also on the amount of lexical-semantic overlap between primes and targets. The next section of the Chapter discusses how conceptual similarities between sentences can boost structural priming effect.

3.3. Conceptually mediated structural priming

One of the first efforts to explore the issue of how conceptual overlap between the sentences may impact the effect of structural priming was undertaken by Bock, Loebell, and Morey (1992). In this production study, Bock and colleagues systematically varied the animacy of the sentential subjects and objects while priming the production of active/passive voice transitive sentences. Participants repeated sentences read to them by the experimenter and then described pictures of transitive events. In half of the experimental trials, animate subjects were paired with inanimate objects; in the other half a reverse pattern was used (e.g., *The boat carried five people* vs. *The boat was carried by five people*). The results revealed priming effects both when the animacy of subjects or objects was repeated and when such correspondence was not maintained. In other words, speakers preferred to bind semantically similar referents from primes to targets, so that the priming effect was directly affected by the referents animacy. One of the caveats in interpreting the results of Bock, Loebell, and Morey (1992) is that it is unclear whether the animacy effect in their study is limited to positional preference to assign animate entities to the pre-verbal position or that the animate status of a referent directly affects the grammatical function assignment immediately related to the Structural priming effect.

A more recent study (Griffin & Weinstein-Tull, 2003) used a sentence recall paradigm (Potter & Lombardi, 1990) to investigate the tendency of speakers to paraphrase the finite complements of object-raising verbs as infinitive complements. The following types of priming constructions were used: object- or subject-infinitive sentences (e.g., *Mr. Forbes ordered his servant to be faster/Jenny actually intended to be a runner in the race*), object- or subject-raising sentences (e.g., *The cook preferred sauces to be spicier/Walter finally started to be kind to his mother*), and intransitive sentences (e.g., *Roy's grade point average slowly improved*). The targets were object-raising sentences with finite complements (e.g., *Allison wished that the bad news was a* *mistake*). Participants read a target sentence followed by a prime, and then they were prompted to recall the target sentence by reading, for example *Allison wished…* Griffin and Weinstein-Tull analyzed the probability to paraphrase the target sentences as a function of the priming construction type. The results demonstrated that the structure that primed the most paraphrases was the object-raising construction, while the fewest paraphrases resulted in the intransitive priming condition. Object-raising sentences have the same word order as, for example, object-infinitive sentences; what makes these structures different is an additional conceptual role in the latter construction. Therefore, a difference in the conceptual structure of an event within sentences that otherwise use the same word order can influence grammaticalization of upcoming sentences and modulate the magnitude of the structural priming effect. However, one recent study using priming from PO/DO sentences with a subordinate clause preceding main clause to sentences containing only main clause (Branigan, et al., 2006, Experiment 3) revealed a reliable priming effect between the sentences that differed in both semantic and structural complexity.

The reports discussed thus far provide mixed evidence about the relations between seemingly autonomous syntax and the lexical-semantic composition of the sentences as revealed by the means of Structural priming. On one hand, some studies supported the existence of an independent syntactic component that is unaffected by the amount of conceptual similarities between primes and targets; on the other, there are strong reasons to believe that repeating of the lexical or semantic environment from one sentence to another interacts with the strength of the structural priming effect. Finally, subtle differences in conceptual structures of otherwise syntactically similar sentences also affect the amount of structural priming propagated from sentence to sentence. Such difference in empirical findings prompted the emergence of considerably different theoretical explanations of what structural priming is for.

3.4. Structural priming. Implications for theory

We have so far established that structural priming seems to be a powerful force influencing both the production and the comprehension of syntax in human discourse. However, the ubiquitous nature and the persistence of this phenomenon beg a number of theoretical questions. The first question is more computational: at what level or stage of sentence planning are the effects of structural priming the most active? Another question is more psychological: what functions does structural priming perform in language development and adult linguistic processing?

3.4.1. Global and local structural priming

In order to answer the first of the two questions, one needs to assume the following: It is very likely that in order for structural priming to exist, a speaker has to rely on temporary or long-term routinization of some structure-building procedures. As Branigan and colleagues (Branigan, et al., 2006) justly note, such procedures differ from language to language, and they can be realized either at a local level (priming of the constituent parts of the structure that constrain their surroundings) or a global level (priming the structure as a whole). The local account suggests that the structural priming procedure creates preferential ordering within the locally built verb, noun phrase, prepositional phrase, etc. (e.g., Hartsuiker & Westenberg, 1999; Kempen & Hoenkamp, 1987; Pickering &

Branigan, 1998) According to the global account, the speaker constructs a structure that contains both hierarchical relations between the constituents and the linearization mechanism for the incremental assembly of the structure (e.g., Chomsky, 1981; Scheepers, 2003; Smith & Wheeldon, 1999). These two accounts are not mutually exclusive, but the specific predictions for particular structure priming may differ.

To provide support for the locally bound structural priming, Branigan, et al. (2006) conducted a set of syntactic experiments investigating details of PO/DO structure priming (arguably, the most commonly primed structure in psycholinguistic studies). However, while most of the previous studies used the materials, in which primes and targets shared global structure, Branigan and colleagues manipulated the extent of global structural overlap between primes and targets. They used a sentence completion paradigm similar to Pickering & Branigan (1998): Participants read and completed prime and target sentences. In two experiments they made use of an additional clause that was present in the priming sentences but not in the targets. For example, the priming materials in Experiment 3 contained an additional subordinate clause:

1. PO Prime: *As Anne claimed, the racing driver showed the torn overall…*
2. DO Prime: *As Anne claimed, the racing driver showed the helpful mechanic…*
3. Target: *The patient showed…*

The same type of manipulation was applied in experiments with an additional adverbial clause. Another set of experiments made use of priming of the local parts of the structure (e.g., main clause to main clause or subordinate clause to subordinate clause) or from one part of the structure to another (e.g., form main to subordinate clause, or vice versa). Finally, a cumulative approach was taken in the last two studies, in which priming of main and subordinate clauses together was extended to just the main clause or just the subordinate clause. The results of all 8 experiments converged on the same conclusion: while priming can be affected by the repetition of the global structure in terms of the repeated clause type, most of the priming effects reside in the tendency to repeat the structural properties locally, as confirmed by the experiments, in which the global structure of the primes and the targets differed.

A somewhat similar idea is advocated in a number of studies by Hartsuiker and colleagues (Hartsuiker, Kolk, & Huiskamp, 1999; Hartsuiker & Westenberg, 2000). The former of these two studies used the grammatical option of partial scrambling of the constituents available to Dutch speakers. For example, it is possible for a speaker of Dutch to say both *On the table is a ball* and *A ball is on the table.* Both these sentences retain the same global structure, but they differ in word order linearity. Participants read similar sentences as primes, and then described unrelated target pictures. As a result, speakers were more likely to repeat the word order from primes to targets regardless of the global structural overlap. The authors take this result as supporting the idea of *linearization* process involved in structural priming. In its simple version, linearization is the mechanism responsible for incremental sentence formulation, so that the speaker does not wait for the complete sentence encoding processing chunks of the formulated sentence as they come online. A similar theoretical approach is taken, for example, in Vigliocco & Nicol (1998).

Theoretical ideas expressed in the studies discussed above are quite appealing as they help explain incremental priming of structures like PO/DO datives or active/passive transitives. An important caveat, however, is that the stable local priming may be limited just to these and similar structures, the formulation of which is strictly bound by restrictions on the conceptual and lexical surrounding, as indicated by the fact that certain verbs (e.g., *donate*) do not even permit PO/DO alternation. The same is true about choosing active versus passive transitive frame. On one hand, the choice between these two syntactic forms is strictly constrained by the grammatical subject assignment to either the agent or the patient of the event; on the other, the form of the verb phrase (e.g., *hit* vs. *is hit by*) further limits the flexibility of the structural choice. In other words, the "structural" in the local priming proposals may in fact be the result of "lexical" foundation for PO/DO and active/passive sentence priming.

In order to provide an independent test of the existence of locally-bound structural priming, one needs to test a structure, which does not rely, in any part, on the lexical constraints on its formation. Such was done in Scheepers (2003). Scheepers conducted a set of German sentence completion studies with the sentences containing global structural ambiguities. A sentence like *The assistant announced the score of the candidate that was unexpectedly poor* contains the relative clause (RC) *that was unexpectedly poor,* which can modify both *candidate* and *score.* The results of this study revealed that speakers of German were more likely to produce a high-attached RC sentence when primed by another sentence with the same pattern of modification. Of course, both types of attachments (high and low) use the same type of local structure; what differs is the global set-up. Based on his results, Scheepers concluded that his results could not be felicitously explained by the increase of the activation at the local level. Instead, the processing system seems to refer to the global plane when computing the locally different structures.

In order to answer the second of the question posed at the beginning of this Chapter, one has to enquire what processing purpose structural priming serves. A number of existing models attempt to answer this question.

3.4.2. Models of structural priming

3.4.2.1. The residual activation account

The information preserved in structural priming has inspired detailed models of grammar representation within the production of lexicon, most explicitly so in Pickering & Branigan (1998). Taking the spreading activation account of Roelofs (1992, 1997) as a starting point, they differentiated between three types of information associated with a verb's *lemma*, namely (a) nodes representing syntactic category (distinguishing verbs from nouns, adjectives etc.), (b) nodes representing syntactic features such as tense, aspect, or number, and most importantly, (c) nodes that represent the verb's combinatorial properties (roughly, the syntactic environments in which the verb can occur). Take for example a set of sentences in (1-4):

1. The rock star sold the guitar to the agent.
2. The shopkeeper handed the groceries to the customer.
3. The rock star sold the agent the guitar.
4. The shopkeeper handed the customer the groceries.

In Pickering in Branigan's theory, the lemma <SELL>, for instance, would link to (at least) two different combinatorial nodes, representing the PO (as in 1, 2) and DO (3, 4) constructions that the verb can legally combine with. Using standard assumptions about decaying activation, Pickering & Branigan explain the priming of a sentence like (2) after

(1) by suggesting that, once the PO node has been activated in the prime (1), it retains some residual activation and thus reaches threshold more easily when making the target utterance (2). Importantly, they further assume that not only individual nodes, but also the links between them may retain residual activation. Hence, they predict that structural priming should become more pronounced when prime and target utterances employ the same verb (e.g., if the targets in (2, 4) also employed the verb *sold* rather than *handed*) because in this case, combined activation from the combinatorial node and the link between lemma and combinatorial node would result in a stronger pre-activation of the relevant structure than if only the combinatorial node were pre-activated.

Some recent research using the implicit learning paradigm (Scheepers & Myachykov, 2006) provided evidence for the reliance of the structural priming mechanism on the gradual decay of the activation. Scheepers and Myachykov showed English speaking participants both grammatical and ungrammatical Russian sentences and asked them to judge their grammaticality. The syntactic structure of some of the trial pairs were repeated to test for effects of structural priming. The results revealed a strong priming effect only locally, from trial to trial. However, the over-time implicit learning occurred independent of the locally established priming effect.

It is unclear at this point whether the lexical boost effect is localized to the information contained in the head NPs and the sentential verbs. Would, for example, an overlap within the constituents playing other grammatical roles lead to a similar boost effect? At the same time, it is uncertain whether the lexical boost effect is a strictly positional or an additive effect. In other words, would it matter for the magnitude of structural priming effect how many constituents in a sentence experience repetition?

Finally, from a strictly lexical point of view, and important question is whether occurrence of the same lexical material in different grammatical roles would reduce or increase the likelihood of structural priming. These and similar questions are the matter of some currently conducted research (Scheepers & Myachykov, in progress).

3.4.2.2. The Dual-Path account

Another model that predicts consistent interaction between lexical-semantic and syntactic levels of processing during language production and comprehension is dual-pathway architecture proposed by Chang (2002). This model is favourably different from other connectionist analogues as it (1) proposes a mechanism for symbolic generalization and (2) places constraints on how sequential information can interact with lexical semantics, effectively creating two pathways in the architecture. The first pathway is represented as *message-lexical system.* As such, it organizes passing of (1) perceptual/thematic and (2) semantic information from message level to the lexicon. The second pathway is responsible for *sequencing* in a sentence the information available from message-lexical system. These two layers of processing work in an interactive fashion. This permits mutual effects from one pathway to the other; therefore preferential coordination of information at one level should lead to corresponding coordination at the other level. Such an interaction between levels of processing is another good candidate for explaining why syntactic priming is boosted by lexical overlaps.

Would the magnitude of a structural priming effect depend on what type of lexical information creates an overlap? Chang's model does not suggest clear-cut predictions that would address this question. It is sufficient that, according to the dual-path model, lexical-semantic priming of components should increase priming of relevant structures during their sequencing in a sentence. It seems plausible, however, to assume that for the architecture described in Chang (2002) it would not make a difference whether, say, a head noun phrase or a verb becomes more accessible at lexical-semantic level; what is important that any lexical overlap should lead to a corresponding increase in the magnitude of syntactic priming.

3.4.2.3. The implicit learning account

The implicit learning model provides a theoretical account quite different from the residual activation proposal. It views structural priming as a vehicle of people's acquisition of complex grammatical knowledge (Chang, et al., 2000; Chang, Dell, & Bock, 2006). Implicit learning typically refers to the fact that parts of knowledge are acquired and adjusted based on human experience – a trial-and-error mechanism that allows knowledge to consolidate outside of explicit processing. As far as structural priming is concerned, it proceeds from the fact that speakers are typically not aware of re-using structure. The conditions under which structural priming occurs provide, in such view, important insights into the mechanisms of *grammatical encoding* – the process of selecting the lexical elements and assembling the syntactic framework for a message to be conveyed (Bock & Levelt, 1994). Over a large number of instances of use, language users may, therefore, attune the syntactic assemblies of their language to the details of the world they speak about. Once they have done it once, they may rely upon this correspondence in an automatic manner when they have to extend the existing grammatical knowledge to new situations of use. Structural priming provides one of the strongest pieces of evidence for the psychological reality of abstract syntactic representations at least partially independent from the details of the event's relational semantics. The proposed mechanism responsible for relating the open-ended syntactic inventory to the particular details of the described events relies upon the existence of additional structures in the sentence processing system. The information computed at the message-relational level, for example the event's referential scheme (who does what to who) structures speakers apprehension of the event. This representational scheme is than mapped onto the syntactic option existing in a given language. Such mapping is on one hand semantically structurally broad; on the other, it may be structurally constrained. The repetitive use of certain conception-syntax correspondences allows for experience-based tuning of the system (Ferreira & Bock, 2006).

In order for implicit learning to take place, the lexical boost effects have to result from memory-based principles that are independent of syntactic priming. Chang, Dell, and Bock (2006), for example, stipulate that lexical boost effects may be due to explicit memory for the wording of the prime: when the target is being formulated, repeated content words serve as cues to the memory of the prime, which biases the speaker to repeat the prime's structure. Again, this predicts that lexical boost effects should not be specific to lexical head repetition, and also, that the more content words are shared between primes and targets, the more priming should be observed.

3.4.2.4. The interactive alignment account

One more theoretical proposal that explicitly addresses the interplay between lexical and syntactic information during language processing is Pickering and Garrod's (2004) *interactive alignment model.* This model suggests a highly mechanistic dialogue interface between interlocutors, according to which both speaker's and comprehender's mental representations of the discourse situation are simultaneously adjusted at multiple discourse levels: conceptual, lexical, syntactic, etc. Because correspondence between different levels is established simultaneously, alignment at one level promotes alignment at all other levels. At the same time, misalignment at one level leads to a similar effect at other levels. The major mechanism that helps establish such alignment is imitation of each other's choices during linguistic communication. In addition, Pickering and Garrod claim that production and comprehension of speech largely rely on the same set of representations. Therefore, the probability of using a lexical item or a syntactic structure during production of speech increases the probability of recognizing the same item or structure during speech comprehension. Consistent simultaneous alignment at different levels results in creation of *routines*, which promote the ease of linguistic communication.

Although, interactive alignment theory does not suggest a general mechanism behind the establishment of the dialogue routines, structural priming is cited among possible candidates. In relation to Pickering and Garrod's model, experimental results from structural priming studies suggest that at least within the syntactic level of processing interlocutors actively align their linguistic representations. However, if the mechanism of priming is not encapsulated with one level of processing, it is natural to assume that priming at lexical level should lead to increased priming at syntactic level and vice-versa. What remains largely unknown is whether (1) the strength of lexical overlap effect is similar regardless of which type of lexical items is being primed – a verb or a noun phrase and (2) whether this effect is additive in nature, so that the more items are consecutively primed at lexical level, the stronger the priming at syntactic level.

3.5. Conclusions. Chapter 3

Chapter 3 discussed the empirical evidence and the theoretical assumptions about the phenomenon of structural priming. While some theories propose the existence of an autonomous component that is responsible for the existence and the longevity of structural priming, other accounts suggest a conceptually and lexically-mediated structural priming architecture. These two accounts are not mutually exclusive, and, while some syntactic structures may rely upon processing of the globally organized frames, other configurations may be more susceptible to the conceptual map of the described scenes and/or the lexical constraints imposed onto the organization of the sentence's syntax. Also, the proposed theoretical explanations of the functional purposes served by structural priming also differ in their account of what structural priming is for. On one hand, it is possible that structural priming serves fundamental purposes of the acquisition and consolidation of the grammatical knowledge; on the other, it may be limited to the residual activation in the working memory. The existing literature provides mixed evidence for both proposals, which further supports the idea of complimentary nature of the two theories. What will be of a concern during the discussion of the current experimental data, is the degree, to which perceptual, conceptual, and syntactic effects can interact in the production environment where all three informational components are available to the speaker. Discussion of the data from Experiments 8 and 9 directly addresses this issue.

Chapter 4. Language production models

To understand the processes that underlie the speakers' choices regarding the structure of the sentence they utter, some considerations about the existing theory of language production must first be made. In the past 30 or so years, a variety of processing models have been developed that, with varying degree of success, attempt to computationally represent what happens in the human mind when it processes sentences. I will only discuss a very small number of such models, and I will discuss them very briefly. Good recent reviews can be found, for example, in Alario, et al. (2006) or Vigliocco & Hartsuiker (2002). Also, at least one of the proposals I discuss here is not a model of sentence processing *per se*; it rather provides a framework for modelling the language faculty as a whole (Jackendoff, 2002). But at the end of the day, linguistic models inform us about necessary constraints on the organization of the language architecture while the job of psycholinguists is to assess how these constraints are realized by the processor in real-time communication.

4.1. Modularity in sentence processing models

Vigliocco and Hartsuiker (2002) justly suggest that existing sentence processing models can be roughly divided into *modular* or *minimalist* (e.g., Bock & Levelt, 1994; Ferreira & Clifton, 1986; Fodor, 1983) and *interactive* or *maximalist* (e.g., Bates & McWhinney, 1982; Dell, 1986; Jackendoff, 2002; Roelofs, 1992; Kempen & Vosse, 1989). In short, modular accounts propose a certain degree of information encapsulation, that is, they postulate that processing at each level within the system is a somewhat autonomous set of operations *"invisible"* (Jackendoff, 2002) to the other processing modules. One consequence of the module encapsulation is that operations at any subsequent level cannot start until processing of the relevant information at the preceding levels has been finished. Also, the flow of information from level to level in modular models is typically unidirectional – from conceptualization to articulation, and, consequentially, feedback from the higher levels of processing to the lower is highly restricted. These two fundamental properties lead to the following logic: Sentence generation starts with the analysys of the preceptual properties of the elements that compose the cognized event; once this process is finished, the conceptual apprehension of the event occurs proceeding from the asymmetries inherited from the preceeptual analysis; the product of this operation – *the message* – is fed forward to the module that is responsible for the selection of the lexical material for the building blocks of the upcoming sentence; finally, these blocks are assembled into a string of sentence constituents, phonological forms are retrieved, and the sentence is articulated.

The motivation for such a division initially came from analyses of speech errors in natural speech. It was observed that such errors typically occur within the same level of processing (Garrett, 1975). For example, lexical exchange errors tend to occur within the same word class, so that nouns exchange with nouns, verbs with verbs, etc. This leads to the logical conclusion that the lexical exchange errors reflect the processing specifics that are limited to the selection of the sentence's noun phrases, verbs, etc. Sound exchanges, on the other hand, often occur between words of different classes rather than between words of the same class. Hence, distribution of phonological erros is taken as word-class insensitive and evident about the fact that accessing phonological forms occurs after the word forms themselves were selected. The ultimate idea, therefore, is that sentence formulation proceeds by passing of information along the assembly line from one station to another until the final product has been put out.

4.1.1. Blueprint for the speaker

A good example of a modular approach to sentence production is Levelt's *blueprint for the speaker* (Levelt 1989; Bock & Levelt, 1994). Reflecting the abovementioned production principles, Bock and Levelt's model includes three levels or stages of sentence generation labeled as (1) MESSAGE component, (2) GRAMMATICAL component, and (3) PHONOLOGICAL component, or ARTICULATOR (see Figure 3).

Each of the processing stages receives input from a level preceding it. To begin producing an utterance, a communicative intention is created. This intention is called a message (cf. Garrett, 1975). In brief, at the message stage, the non-verbal information is processed and organized into a conceptual scheme of the event before any linguistic processing can occur. Although the authors of this and a number of derivative models do not explicitly state that, it is at this level that the perceptual effects are supposedly the most active biasing conceptualization of the event according to the event's salience map. The message captures features of the speaker's intended meaning and provides the raw material for the grammatical encoding. The grammatical component is divided into two sub-stages: Functional processing level and Positional processing level. The functional level is responsible for the selection of word *lemmas.* Lemmas are thought to be "amalgams" of an individual lexical concept's properties including its semantic representation and its morpho-syntactic features. However, they are not the lexical forms yet. Also at this level, grammatical functions, like Subject and Object are assigned. The lemma and the grammatical function information is fed to the Positional sub-component. It is here that representations of words are sequentially inserted into a sentence frame that later becomes fixed as the order of the elements in an utterance. This ordering may not be imposed during functional processing. Evidence for this comes from different types of errors. For example, when sounds are exchanged, they originate in the same phrase 87% of the time as opposed to whole word exchanges that occurre within the same phrases only 19% of the time (Garrett, 1980). Finally, at the Phonological level, the phonological forms for the words are retrieved and an overt utterance is produced.

A somewhat similar account was proposed by Tomlin (1997). Tomlin used a similar set of processing stages but with different labelling. Tomlin's labelling reflects the fact that the model is specifically adjusted to the generation of sentences driven by the perceptual properties of the cognized scene (See Figure 4). Due to the same intent, Tomlin provided a detailed account of how the scene is interrogated rather than how the utterance itself is articulated. Therefore, the model's focus is on interfacing non-linguistic and linguistic processes. Tomlin refers to this process as *mapping* of the perceptual plane onto the linguistic ones. According to his model, the perceived event is conceptualized as set of three features: (1) *Field* – conceptual area in which event occurs (in other words, the production environment), (2) *Parameters* – stable conceptual units in the field: Objects, object-complexes, or abstract concepts, and (3) *Action* – a state of relationship between the parameters or a change of such state. The parameters of the scene are mapped via attention detection filter that assigns higher priority to the focused elements onto the reference frame and the grammatical functions within the produced sentence, such as Subject, Verb, and Object.

Both models provide viable explanations of how sentence generation may proceed, with the latter one specifically geared toward explaining the relationship between the perceptual and the grammatical processes. However, many questions remain unanswered. First, it is still relatively unclear which processes are localized to which levels. For example, if one were to be absolutely precise, one might need to assume yet wider division of labour with perceptual operations assigned a separate processing node preceding conceptual analysis. Although perceptual properties of the cognized objects are typically extracted in parallel with the conceptual information, it does not always have to be the case. A situation is quite possible when a visual trace or a perceptual cue, conceptually uninformative themselves, provide important visual indexing within the field that facilitates later processing of the elements that occur in the marked locations (e.g., Altman & Kamide, 2004; Spivey, Richardson, & Fitneva, 2004). Such facilitation, when it occurs, has nothing to do with conceptual processing and it effectively precedes it. Second, although it is postulated that the word order assembly is the prerogative of the Positional component, often lexical information provides powerful constraints on what function the word can perform in a sentence and where in a sentence this word can appear. Hence, a situation is possible when procesing at assembly level is already restricted when lemmas are retrieved. Third, it is unclear whether processing at each level is in fact encapsulated, so that the efficiency and automaticity of the production is achieved through the completeness of the information passed on to the other levels. Alternative reasoning (see below) is that processing efficiency can be achieved by spreading activation across levels. If, on the other hand, interactions between levels do occur, it is uncertain whether they occur only within the neighboring levels or are they possible between levels that are not adjacent to each other (e.g., Message and Grammatical encoding)? Finally, it is questionable that processing is in fact unidirectional. More often than not, sentence production relies on availability of previously processed information, both linguistic and non-linguistic. In this situation, it is quite plausible that priming of material at higher levels would lead to virtual absence of preplanning at levels preceding them during generation of subsequent sentences.

4.2. Connectionism in sentence processing models

Another type of language processing model was motivated by the development of the connectionist theory (McClelland & Rummelhart, 1981). It is important to note that connectionist models do not come as a theoretical opposition *proper* to either modular or interactive proposals. In fact, both modular and interactive architectures can be and have been modelled using a connectionist approach. What makes connectionism theoretically unique is its treatment of how processing happens. Generally, connectionism abandoned understanding of mental processes as operations on certain types of *representations*. Instead, it assumes that any task including linguistic ones can be successfully represented in a network model whose performance relies on a complex set of connections between neuron-like processing units. Dynamic changes in the *weights* of the connections (or the connection's strength) are what effectively constitute processing in a connectionist model.

So, a typical neural network consists of layers of units that are connected to each by a pattern of connections. There are usually three types of units in a neural network: input units, hidden units, and output units. As Garson notes, "If a neural net were to model the whole human nervous system, the input units would be analogous to the sensory neurons, the output units to the motor neurons, and the hidden units to all other neurons" (Garson, 2007). To implement such networks, connectionists develop computational models that are first trained on a limited number of input examples. Such gradual learning is based on continuous changing of the *weights* of the connections between the networks' units. These changes determine activation patterns and the model's behaviour. The logic of the weight change is as follows. First, a training set of inputs and their desired outputs is selected. Examples in the training set are assigned initial weights that are determined, for example theoretically or based on observed frequencies. Before training occurs, the net's weights are assigned as random values. Then, the net is exposed to the training set. The outputs of the net's performance are compared to the desired output for the training set member. The weights in the net are then adjusted toward the direction of the desired input. Upon repeating the training process a number of times, the net is able to produce the output reliably similar to the desired one for each member of the training set. Hopefully, the model can then generalize the acquired rules to novel materials. Although connectionist models differ in the architectures they are based on, this principle of organization remains the same.

With regard to language processing, connectionism came as an opposition to the idea that the acquisition of the language ability has to rely upon an innate mechanism that enables a child to extract the language lexicon and grammar in a situation of limited input. Instead of postulating a certain set of representations that a language learner has to rely upon during language development and later use in adult linguistic performance, connectionist theories attempt to explain behavioral phenomena, including language, in terms of networks of simple, neuronal processing units (e.g., Chang, 2002; Cottrell, 1989; Dell, Chang, & Griffin, 1999; Elman, 1990; 1991; McClelland & Rummelhart, 1981; Rohde & Plaut, 2003; Seidenberg, 1997). Some connectionist models of speech production are restricted to learning units at word level (e.g., Dell, 1986; Harley, 1993). This, of course, limits the inferences to the levels of lexical retrieval and phonological encoding and does not permit modelling of the full-fledged sentence production. Other models (e.g., Dell, Chang, & Griffin, 1999; Gasser, 1988; Kalita &Shastri, 1994; Rohde, 2002; Ward, 1991) permit generation of simple sentences including active/passive voice alternation.

For example, Dell, Chang, and Griffin' s *prod-SNP* (1999) tried to adapt Elman's Simple Recurrent Network model (SRN) (Elman, 1990) in order to model how the mechanism responsible for the structural priming (Bock, 1986) can also explain general sentence production. The model mapped the sentence propositional content with the help of a slot-based approach so that each subsequent word was inserted sequentially – based on the properties of the preceding constituent. The Dell, et al's model learned representation of the grammatical voice alternation in English by assigning the higher weight to either agent or the patient role of the constituent in the head NP. Also, it was able to learn two types of the English dative construction: The prepositional object and the double object. The model was quite successful in mimicking the structural priming data typically observed in experimental studies with human subjects.

Another example of connectionist architecture capable of modeling production (and comprehension) of relatively complex syntactic strings is Rohde's (2002) CSCP – the Connectionist Sentence Comprehension and Production model. CSCP uses a large-scale SRN representation to be able to model both comprehension and production of relatively complex grammatical phenomena in English including voice alternation, multi-clausal sentences, etc. Comprehension and production mechanisms in this model are integrated and for the most part, rely on the same set of processes in such a way that the production mode is acquired by making predictions from comprehending sentences. The model uses the *semantic system* component that maps propositional content into a single static representation. In the comprehension mode, the model receives a sequence of words with their phonological properties and represents the sentence meaning so that it could be decoded by the semantic component. When producing a sentence, the model predicts the sentential starting point by associating the message layer with the correct word meaning. The model then feeds the most strongly predicted word back into the comprehension input and continues with the production of a new word. Hence, this sequence repeats itself feeding information via both production and comprehension routs. A variety of training tests proved that CSCP is capable of representing such language phenomena as frequency effects, structural priming, and number agreement.

Arguably, one of the most influential connectionist models of sentence production is *dual-path model* (Chang, 2002). In his model, Chang proposed two pathways for sentence generation: one for mapping entities' semantics onto the word forms and the other for mapping the entities onto sentence positions. Chang labeled the first pathway *the message-lexical system.* The connection weights in the message-lexical pathway relayed representations in two layers: *where* units, responsible for representation of the thematic information and *what* units, responsible for representation of word meanings independent of the event roles. The second pathway – *the sequencing system* – used a variant of SRN to be able to sequence each unit in an upcoming sentence. Interestingly, dual-path architecture was successful in both acquisition of syntax and modeling double dissociation in aphasia. What makes dual-path model theoretically relevant for the research reported below is that Chang goes beyond a somewhat simplified level of reasoning in his attempt to model processing at the message level specifically suggesting that perceptual biases on sentence formulation can be fed forward in parallel with processing at linguistic levels. Theorizing on how exactly attentional control can propagate biases in his model Chang uses evidence from Griffin and Bock (2000) that I have already discussed. More specifically, he proposes that the tight linkage between attention to referents' spatial locations and the thematic organization of sentences is related to the activation of *where* units in the dual-path model. According to this view, online structural decisions and eye movements as correlates of attentional foci can be synchronized in order to properly map the thematic scheme of the event onto the structural frame of the sentence.

4.3. Interactivity in sentence production models

Many theories of language processing including a number of connectionist architectures mentioned above rely upon a set of interactions between units representing general cognitive processing and those limited to linguistic operations. Interactive theories of sentence processing suggest a view theoretically very different from modular accounts. In order to represent interactive properties in the system, one has to allow a certain degree of simultaneous processing at a number of levels and bidirectional flow of information from one level to another. Interestingly, most current theories of sentence comprehension assume exactly this type of processing (e.g., Knoeferle & Crocker, 2006; Tabor & Tanenhaus, 199; Tanenhaus & Trueswell, 1995). Some theorists even propose the existence of a more-or-less single system that emulates processing in both interaction channels (Pickering & Garrod, 2006; Rohde, 2002). According to interactive comprehension accounts, sentence comprehension is a very opportunistic process when information becomes available to the comprehender at different levels at the same time, and the comprehender's task is to actively update her understanding of the discourse based on the various types of information almost simultaneously. However, much of psycholinguistic research heavily relies on modular accounts. That said, there are a number of sentence production models that rely upon a considerable degree of parallel activations and interactions between levels of processing. According to many interactive accounts, generation of sentences proceeds by virtue of *cascading* or *spreading of activation* (e.g., Dell, 1986; Roelofs, 1992), so that processing at one level does not have to be complete before processing at the subsequent level starts. On the other hand, many models assume that the production system's performance is governed by a system of constraints that are accommodated within the language grammar (e.g., Prince & Smolensky, 1993; Stemberger, 1985; Frazier, 1995; Tabossi, et al., 1994). In short, such constraints tell the processing system what it can and cannot do according to the grammatical rules of the language.

Parallel simultaneous contributions from perception, local semantics, and general linguistic knowledge are typically described by invoking metaphors, such as "workspace" (Ferreira, 2005) or "blackboard" (Jackendoff, 2002; van der Velde & de Kamps, 2006). Like connectionism, parallelism has its prerequisites in cognitive science. A good example of a general-purpose parallel architecture model is found in Dehaene, Kerszberg, & Changeux (1998), Dehaene and Naccache (2001), and Dehaene and Changeux (2004) (See Figure 5).

Two important features organize Dehaene's global workspace model: Dehaene's model avoids strict directionality of input-output relations between the processing stages; and permits wide parallel interfaces which feed signals from multiple sources of information to the processor. Each signal operates as a possible "prominence candidate" attempting to bias the decision of the processor and, therefore, the behavioral output. Executive processing occurs within a workspace that receives simultaneous input signals from different sources of information (in parallel). For example, explaining a model largely similar to Jackendoff's, Dehaene suggests that processes within workspace architecture do not "obey a principle of local, encapsulated connectivity but rather break the modularity of cortex by allowing many different processors to exchange information in a global and flexible manner. Information which is encoded in workspace is quickly made available to other brain systems for overt behavioral report." (Dehaene & Changeaux, 2004, p. 1147). This type of architecture allows modelling of real-time language processing as a set of constraint-controlled cascading parallel interactions between different types of information, each of which tries to bias the course of preparation, production, and comprehension of speech.

In a similar fashion, Jackendoff (2002) invokes Parallel Architecture as a model of the language faculty that helps to represent the interactive properties within the linguistic system and analyze how units at lexical and grammatical levels are interrelated with the specifics of the conceptual and perceptual planes. Jackendoff can be congratulated with a successful attempt to reconcile "processes vs representations" debate. He hypothesizes that while elements within the same level are probably organized in a more or less representational form (some domain-specific, some not), interfaces between the levels can rely upon the same set of distributed processing parameters. This assumption is important because it allows avoiding seemingly inevitable theoretical controversies and attempting to accommodate both modularity within the processing levels and distributed properties within the interfaces between the perceptual, the conceptual, and the linguistic planes.

4.4. The interface model: A unified approach

A detailed analysis of the proposed interface model can be found in (Myachykov, Posner, & Tomlin, in press). Here, I will only describe the basic parameters of the proposed architecture in order to later accommodate the experimental results within the model.

Leaving connectionism aside, I am primarily interested in how representation of information within processing modules and parallel processing at multiple levels can be achieved within the same architecture. Such an eclectic approach is of course not novel at all. For example, Levelt stressed in the opening pages of his seminal book *Speaking: From Intention to Articulation* (1989) that the modular approach does not advocate total encapsulation of processing within levels; rather, it suggests that processing at respective levels is easier to study experimentally with the modular approach in mind. Similarly, Jackendoff (2002) leaves space for separate sets of representations within modules and persistent interactions both within language layers and the interactions between language and other cognitive domains.

Here, I will combine the parallel processing approach with Levelt and Bock's feed-forward formulator to the analysis of the chronometric and functional properties of the interface between the grammatical system of a particular language, relative accessibility of lexical/semantic information about the discourse referents, and the distribution of attention in the described scene. Figure 6 illustrates parameters of the resulting model. The proposed model is organized around a number of principles. These principles help generate specific hypotheses about the model's behavior in different perceptual, conceptual, and syntactic environments and with different language systems. The processing level (workspace) includes traditional production stages – message, lemma, and assembly. Operations at the processing level occur in parallel cascade fashion. Therefore, processes within neighboring nodes can temporarily overlap.

Processing levels receive inputs from two input levels – global and local (cf. Prat-Sala & Branigan, 2000). Informational inputs at both levels are *roughly* divided into four types – perceptual, conceptual, lexical, and syntactic. Inputs from the global level operate a set of permanent *constraining parameters.* Inputs from the local level comprise a set of *priming parameters.* Information can be primed at each corresponding level. Successfully primed information is propagated to the level immediately following the one it was primed at. Therefore, processing at lemma level can depend on both priming of lexical information and the priming effects carried over from the message level. Constraints are established at each corresponding level, and they can control information processing at the level preceding it. For example, constraints at assembly level (e.g., canonical SVO) can partially control the input from the lemma level (e.g., prefer to enter the noun in the nominative case as the sentential starting point). Therefore constraints establish control locally and can propagate it from the higher levels of processing to the lower, while priming operates locally and can be promoted from the lower to the higher levels.

Global constraints reflect the speaker's processing preferences stored in long-term memory. They are higher probability tendencies, for example, to visually interrogate agents before patients or to use higher frequency lexical items in case of synonymous choices. A part of the constraining system is the speaker's grammatical knowledge, which can be understood as a set of (1) rules, (2) affordances, and (3) preferences. Rules include all possible well-formed structural options existing in the language grammar. Affordances specify what grammatical structures are felicitously applicable to a specific situation. Preferences are higher frequency structures with a large probability of use regardless of the distribution of priming parameters in the described array. In other words, preferences form a subset of the language grammatical defaults. An example of a grammatical constraint is a speaker's tendency to utilize canonical grammar in a wide range of situations.

While perceptual and conceptual preferences might be very similar across speakers of different languages, the power and the nature of linguistic constraints (lexical and syntactic) may differ across grammatical systems. In other words, some morphological and syntactic properties of one language may be more constraining than of another, at least with regard to visually-mediated sentence production. For example, an additional regulator for lemma selection and the assignment of word order, case marking, might lead to higher dependence of the speakers of case marking languages to use canonical structures. One, for example, might expect that there will be stronger priming effects across all three dimensions (perceptual, semantic, and syntactic) in English than

Russian because the promotion of a referent or a structure in Russian is governed by an additional constraint – case marking, which is not active in English grammar.

Priming parameters are available to the interlocutors locally: They originate from the details of the communicative situation. Such parameters can be understood as oscillations in referents' accessibility as a result of the changes in their relative prominence. An increase in priming parameters will lead to promotion of a referent within a respective processing node, i.e. a visually focused object (in case of visually-situated speech), a lexically primed name referring to an object or a concept, or syntactically primed structure for the description of the event. In the absence of specific priming parameters or in the situation when primes promote same items and structures as constraints, the processor's default is to use constraints. This should lead to the ceiling effect in the behavior of the model. In this case, processing is very shallow, and it utilizes little cognitive resources. For example, in the absence of competing cues, speakers of SVO languages will tend to fixate agents first, mapping their canonical SVO sentences onto the preferred causality map in virtually 100% of cases.

Both constraining and priming inputs carry information only to the nodes they are connected to. For example, perceptual effects cannot influence processing at assembly level directly; they have to exert their influence during the initial analysis of the visual information. However, perceptual biases can interact with the conceptual details of the scene. If interaction is powerful enough, its effect will be carried to the choice of lemmas and potentially influence the referent's preferential positioning in a sentence. Operations within "language-related" nodes (lemma and assembly) are more capable of influencing linguistic choices because processing within these nodes occurs closer to the output.

Therefore, an increase in the semantic or syntactic complexity will decrease the likelihood of perceptual effects biasing the output sentence. One possible consequence is that perceptual priming will be more likely to affect the organization of transitive than ditransitive sentences.

The interaction between Input Levels occurs within the Processing Level. It is governed by the following principles. *Competition*: Priming parameters will sometimes concur, and they will sometimes compete with the constraining parameters. Due to the higher probability of general use and, their automaticity and ability to minimize the cognitive load during sentence processing, constraining parameters are stronger than priming parameters by default. To violate a constraint, the competing priming input should be relatively stronger. *Weight*: The larger the weight of an input, the greater the chances its effect will survive competition with other competing or constraining inputs. Because of that, endogenous restrictions on attentional focus may increase the relative weight of the perceptual priming parameter and help promote its bias to the lemma level, leading to higher probability of inclusion of the corresponding name at the frontal position of an upcoming sentence. *Connection*: The closer the input connection to the output, the greater the chances that its influences will be realized in the output. For example, perceptual priming will bias the resulting syntactic structure to a lesser extent than syntactic priming because the effects of the former are realized only within the message level. As a result, perceptual priming effects during sentence production should be relatively weak compared to, say, lexical or syntactic effects. *Interaction 1*: Only the processing nodes adjacent to each other will interact. For example, because there is no direct connection between the message and the assembly nodes, priming at message level will have to be supported by similar priming at lemma level for the effect of the former effect to be promoted to the assembly level. *Interaction 2*: Only the input nodes adjacent to each other will interact. As a result, one should expect interactions between perceptual and conceptual, but not between perceptual and syntactic priming components in the case of multiple sources of priming.

4.4.1. Motivations for the measurement set

The proposed model helps to justify the measurements I will use in 8 out of 9 studies reported in this thesis. Four separate data sets will be reported in each experiment: (1) probability of alternating the produced word order as a factor of the experimental manipulations, (2) onset of the initial fixations to the event's interest areas – referents and action, where possible, (3) eye-voice spans for each of the sentence constituents – also where possible, and (4) the timeline of sentence production reflected in each constituent name onset latency.

The first measurement – word order choice – represents the structural organization of the final product – a full-fledged sentence. The other three measurements – the first fixations, the eye-voice spans, and the name onset latencies – provide sequential windowing into (1) perceptual, (2) conceptual, and (3) linguistic processes during formulation of sentences (see Figure 7). Because these three successive stages overlap, the specified measurements also extend the boundaries of the processing stages.

The first fixations to the information-laden elements of the scene provide the earliest window into the initial uptake of information before any "serious" conceptual processing can begin. The distribution and the timeline of these early gazes reflect, in the first place, the bottom-up perceptual processes associated with the uptake of the visual information. As such they typically reflect allocation of attentional resources within the cognized scene. When an experimenter uses manipulations of attentional focus by means of visual cueing, first fixations to the cued referent obviously inform her about the efficiency of the cueing procedure. More importantly, the initial looks to the referents, especially the looks to the non-cued referents, may be informative with regard to what I will refer to as *rapid (event) apprehension* (cf. Griffin & Bock, 2000). Rapid apprehension provides the analysis of the most basic, raw conceptual information the scene contains. This process is likely to be non-linguistic for the most part, and it does not have to be complete before the next stage processing starts when, for example, the perceptual information promotes highly predictable conceptual and linguistic structures. Rapid apprehension logically provides the best opportunity for the factors related to perceptual salience and event causality to have their greatest chance at offering their biases for subsequent sentence production because, if not confirmed by other factors, these effects have to compete with various linguistic biases.

The next measurement – eye-voice span (EVS) – reflects operations that bridge non-linguistic and linguistic processing. This measurement is motivated by the fact that, after apprehending the event, speakers tend to look at referents shortly before naming them. What eye-voice span measures is the time between the onsets of these gazes and the onsets of the corresponding names. Of course, a situation is possible when the first fixation to a referent is the last one relative to the corresponding name onset. This explains *potential* overlap between the two measurements. However, such situation is not inevitable: as I will demonstrate, in regular sentence production, rapid apprehension is completed before the looks preceding sentence formulation occur (cf. Griffin & Bock, 2000). Therefore, it is the processing stretch from rapid apprehension to sentence formulation that the EVS generally measures.

One important issue related to the EVS measurement in visual world production research is existing ambiguities with regard to what exactly EVS reflects in terms of information processing. The first study that applied EVS in picture description tasks (Griffin & Bock, 2000) reported EVSs of around 900 msec. for both Subjects and Objects of English transitive sentences. Because such magnitude was comparable to the measurements of lexical access during single referent naming, the conclusion was that EVS probably reflect the same process during the formulation of sentences. A recent study (Bock, et al., 2003) reported much shorter eye-voice spans for narratives describing time displays (circa 350 msec). Reporting time is probably structurally easier than producing full-fledged sentences, but the same cannot be said about the associated lexical retrieval. There is no good reason to suggest that the retrieval of names necessary for reporting time is easier than the same process during ditransitive sentence production. Therefore, eye-voice spans have to represent something else in addition to reflecting lexical access during incremental production of sentences.

Based on their chronometric span and on what is known about the fluctuations of the EVS values from visual world studies I suspect that they have a more complex nature. I propose that they represent the relay of the rapid apprehension product to the linguistic modules. This process begins with finalizing conceptual analysis and extends through lemma selection to phonological encoding. As such, it has to include lexical access and mapping of the referential scheme onto the grammatical plan of the sentence. Facilitation of processing within one or both of these subcomponents might lead to deflation of the EVS magnitude. Also, because eye-voice spans reflect conceptual analysis, facilitation of the latter by manipulating the perceptual or conceptual environment might lead to the same effect. All together, it is *the processing difficulty* associated with both conceptual and linguistic analysis of the scene referents that eye-voice span reflects. If this is true, one would expect the magnitude of EVSs to change as a function of (1) conceptual complexity (2) lexical complexity, and (3) syntactic complexity involved in the promotion of a referent to its sentential role.

Finally, latencies to produce sentence constituents chronometrically include both aforementioned measurements and reflect (1) completion of all the cognitive operations on building blocks in a sentence and (2) the general ease of structure building. It is important to separate the EVS measurement from the name onset latencies because the latter confound all three stages of sentence formulation, while the EVS permits separation of rapid apprehension from the rest of processing. Hence, I will report the collected data starting from non-linguistic uptake of information through conceptual processing to language related operations.

Chapter 5. Perceptual priming and sentences production in different languages

The series of experiments discussed in Chapter 5 further test the theoretical proposal discussed in Chapter 2, that directing attention to one of the event's referents increases the conceptual accessibility of the latter and its chances of claiming the prominent syntactic slot in the upcoming sentence. A number of theories (Bates & McWhinney, 1982; Kempen & Vosse, 1989; McWhinney, 1975; Tomlin, 1995; Chang, 2002; Dell, Chang, & Griffin, 1999) predict such preferential placement through the tendency to construct sentences incrementally, starting from the most conspicuous or important element. The rest of the syntactic structure then inevitably becomes biased due to the assignment of the starting point of a sentence to a particular referent. So, for example, in English the passive frame follows if the starting point of a sentence, and therefore the grammatical subject slot is assigned to the event's patient.

All experiments except one (Experiment 5) analyze transitive sentence production across languages with different degrees of word order flexibility. Experiment 1 uses the Fish Film paradigm (Tomlin, 1995) in Russian – a language that allows a very flexible ordering of the constituents within a sentence. Experiments 2 and 3 use eye-tracking methodology to address the issue of regular transitive sentence production in a no-manipulation setting in both English and Russian. Experiments 4 and 5 apply implicit attentional cueing in and eye-tracking methodology in the analysis of both transitive and ditransitive sentence production in another flexible word order language – Finnish.

5.1. Experiment 1. Perceptually-primed sentence production of Russian transitive sentences

5.1.1. Some facts about Russian word order

Experiment 1 uses Fish Film paradigm (Tomlin, 1995) in order to analyze the structural and the dynamic properties of visually cued sentence production in Russian. Two important comparisons between English and Russian grammars will help us generate specific hypotheses for the study.

First, NP constituents in an English sentence do not assume specific case markers in either subject or object position while Russian ones almost always do. The lack of case marking makes it possible for English speakers to maintain early commitments to the sentential starting point and decide which structure they want to produce later when it becomes clear what thematic role the entity in focus performs (cf. Bock, 1986a). This strategy is not available for a Russian speaker because it is hardly possible to separate lexical access from sequencing constituents in a Russian sentence. The case marker, a part of the lexical form, carries information about the thematic role the word can assume in a sentence therefore biasing sentence production toward a particular order. Rather than word order, it is the case system and, therefore, the endings of the words that tell you which element of the Russian sentence is its subject, or its object.

Second, the choice of word order in English is limited to the alternation between active and passive voice. The grammatical case system of Russian creates wider opportunities for syntactic scrambling. Because case marking indicates the syntactic roles, all possible combinations of Subject, Verb, and Object are possible without alternation of the sentence meaning. Apart from scrambling, Russian makes use of

Passive Voice construction with the auxiliary verb *byt' (to be)*, but this construction is very infrequent in regular speech. Finally, Russian has a higher-frequency canonical order – active voice SVO. This pattern is usually treated as "objective" or "neutral", while alternative word orders encode focalization or topicalization functions (Kovtunova, 1976; Yokoyama, 1986; Krylova & Khavronina, 1988; Mathesius, 1947; Vinogradov & Istrina, 1960;). Because the Russian language makes active use of case marking, the grammatical function of Russian word order is restricted to a small number of specific cases. For example, sentences like *"Mat' ljubit doch'"* – "Mother (Nom/Acc) loves daughter (Nom/Acc)" (Jakobson, 1936) are ambiguous because the noun forms in this sentence do not differ in their Nominative and Accusative form. When presented in isolation, such sentences are consistently interpreted as SVO (King, 1995).

The communicative function of Russian word order is traditionally believed to depend upon the relationship between theme and rheme in the sentence. According to this view, old or given information always precedes the new information in a spoken sentence. Sgall, et al. (1986) defines this phenomenon as a reflection of *Communicative Dynamism* (CD). This view ascribes constituents with more CD to the newest informational units. These constituents should appear after the ones bearing older information in Russian utterance. Sometimes thematic prominence is viewed as a more diversified system, which makes use of more than two possible categories.

The traditional view of Russian word order as a reflection of thematic importance of the constituents assumes that in naturally occurring and stylistically neutral speech at least SVO and OVS should be readily available to the speakers provided the thematic relations between the referents warrants one of these two patterns.

5.1.2. Experimental hypotheses

How can the grammatical differences between English and Russian inform us about possible behaviour of Russian participants in the Fish Film task?

1. One possible hypothesis is that (1) because in the Fish Film the discourse preceding the target event actively manipulates the accessibility of both referents with the visual cue making one of the fish the given information and (2) because this type of mapping is supposed to be the driving force of Russian non-canonical orders, word order scrambling should be the preferred device to map the cueing scheme of the Fish Film onto the target sentences. In other words, word order scrambling should do in Russian what the voice system does in English. If this is the case, one should observe agent-initial constructions in the agent-cued condition (e.g. SVO, SOV), while the patient-initial constructions should dominate in the patient-cued condition (e.g. OVS or Passive Voice).

2. However, the existence of case marking in Russian might create a problem for an incremental processor. Because the Fish Film instruction elicits descriptions of everything that happens before the target event, a speaker of Russian is likely to initially refer to the cued fish with the Nominative form (*ryb-a*) hoping to use it as the agent in the description of the target event. If this expectation is not confirmed and the cued fish appears to be the patient, a repair toward Accusative *ryb-u* is necessary so that an OVS sentence could be produced. This is difficult due to (1) the time pressure of the task and (2) because the participants are instructed to describe the target event extemporaneously, as they view it. Another option is to maintain the commitment to the Nominative word form but change the perspective from the cued patient to the non-cued agent and assign the syntactic positions post-hoc based on a new apprehension of the event. In either case,

the necessity to use case-marked forms throughout the experiment should make consistent alternation of word order problematic. In this case, Russian speakers may have to rely upon the use of the canonical word order in both cueing conditions. This should make them switch attention from the cued fish to the non-cued one in order to arrive at SVO, which should result in a reaction time penalty and slow down sentence production in the patient-cued condition.

Finally, one could expect Russian speakers to use Passive Voice sentences. Passive Voice would make it possible for Russian participants to behave like their English counterparts and maintain the commitment to the Nominative word form regardless of the cue location. I am not aware of any study that attempted to elicit Russian passive voice using a psycholinguistic task. However, a study using Serbian – a language grammatically similar to Russian – (Gennari, et al., 2005) found it extremely difficult. Based on this knowledge, I did not expect to observe many passive voice constructions in my Russian sample.

Also, wide range of syntactic choice may both complicate and facilitate sentence production. If some early syntactic planning occurs in parallel with conceptual analysis, a multiplicity of choice should complicate production because the speakers would need to entertain a number of candidate structures for the upcoming sentence. This should increase the cognitive load associated with linguistic processing and complicate the establishment of a strong coupling between the attentional focus and the positioning of the sentence constituents. However, there is some evidence that a wider choice of structural options can instead make production of sentences easier (Ferreira, 1996). In his experiment, Ferreira asked English participants to create sentences using the words that either permit or restrict the use of multiple structural choices. It turned out that the sentences that had structural alternatives were constructed faster, which suggested that availability of choice facilitates the production of sentences. This result is usually taken as supportive of strictly incremental models of sentence production with minimal syntactic pre-planning in mind. From this perspective, the performance of Russian participants on Fish Film task should be at least as good as that of their English counterparts.

Therefore, two working hypotheses for this study are possible: (1) a wider inventory of structural choices in Russian should facilitate for syntactic scrambling prompting speakers to use patient-initial constructions when the event patient is cued while the agent-cued trials should be predominantly described by agent-initial constructions; (2) the necessity to commit to a particular case form will make the establishment of a consistent mapping between attentional focus and the choice of word order complicated when the cued referent is the patient. In this case one should expect participants to avoid reassignment of the case form and use active voice SVO predominantly in both cueing conditions.

5.1.3. Method

Utterances were collected with the help of the Fish Film animation program (Tomlin, 1995). There were 32 trials in this version of the Fish Film. There were no filler trials between the target trials. In half of the trials, the cued element was the agent; in the other half it was the patient. The color of the fish as well as the ordering of the trials was randomly assigned. The total duration of the Fish Film was 4 minutes and 40 seconds.

The core element of each trial was the target event, in which one of the fish ate the other. During the preceding portion of each trial, the two fish approached each other proceeding simultaneously from both sides of the screen (See Figure 2). An audio signal accompanied and coincided with, the presentation of the target event. The audio signal was recorded alongside the participants' narratives for further chronometric analysis.

5.1.3.1. Cueing

Participants' attention was manipulated via the presentation of an exogenous visual cue – an arrow to the spatial location of one of the fish (Posner, 1980). The experimental instruction was to look at the cued fish for the entire duration of each trial and extemporaneously describe the unfolding events. The cue appeared twice during each trial: First, shortly after the appearance of the both fish, second, shortly before the presentation of the target event. Once presented for the second time, the cue stayed above the cued fish until the trial ended.

5.1.3.2. Experimental task

The participants were instructed to describe the events of the Fish Film extemporaneously. The fact that participants had to describe how the two fish approached each other resulted in the intervention of 4-5 sentences between the descriptions of the target event. These intervening sentences were structurally very different form the transitive sentences used to describe the target event; therefore they can be treated as effective fillers separating the target sentences. I discuss the nature of these sentences in Results section.

5.1.3.3. Participants

15 monolingual native speakers of Russian (5 males) participated in the study. All the participants were recruited from the population of undergraduate students of Moscow State University. All the participants had normal or corrected-to normal vision. Their age ranged from 20 to 25 years and averaged 21.48 years. Participants took part in the study voluntarily, and they did not receive any material reward for their participation.

5.1.3.4. Apparatus

Participants were tested in a laboratory setting with the help of a DELL Inspiron 8200 laptop computer. The utterances were collected with the help of the microphone built into the laptop. The participants did not receive any specific instruction regarding the sound marker.

The computer program Adobe Audition 2.0 was used for the analysis of the collected utterances. The target utterances were the descriptions of the dynamic event. They were analyzed in terms of their syntactic structure as well as the speech onset latencies measured as reaction times to initiate the descriptions of the target event. The audio marker of the dynamic event appeared as a clear peak with the duration of approximately 20 msec. The offset of this component marked the starting point for the reaction time analysis. The onset of the corresponding narrative indicated the finishing point of the analyzed latencies.

The investigator measured reaction times manually. This involved a subjective factor that might have influenced the results. In order to avoid possible errors, the measurer remained blind to the cueing map of each trial; therefore he could not know whether the coded trial was the agent- or the patient-cued. In addition, an independent native speaker coded one participant's recording. Correlation of measurements by the investigator and those performed by the independent measurer was significant ($r=.98$, $p\leq.001$). This analysis confirmed reliability of the reaction time measure.

5.1.4. Results

In all experiments reported in this thesis, separate tests were performed with participants and items as random factors (F1 and F2, respectively). Only F values statistically significant at the 5% level are reported, otherwise the result is reported as non-significant.

5.1.4.1. Word order analysis

477 out of 480 collected utterances were subjected to statistical analysis. Three utterances were dismissed because the participants failed to describe the target event. Each target description was analyzed with respect to its word-order structure. Tables 2 and 3 represents the distribution of syntactic patterns by the trial type (the items in brackets were optional) and by the specified word order types. Table 2 clearly demonstrates that participants used agent-initial word orders in over 81% of all trials, while patient-initial structures were used only in 19% of all trials. In the agent-cued condition, the participants used agent-initial structures in virtually 100 % of their narratives: There were only 4 patient-initial descriptions in this condition. Although patient-initial structures accounted only for 36.4 % of utterances in the patient-cued condition, this increase was reliable. An ANOVA on the probabilities of observing agent-initial responses revealed a reliable effect of Cueing ($F1(1,14)=12.560$; $F2(1,31)=25.043$).

5.1.4.2. Error rates

The results of the error-rate analysis are presented in the Table 4. All the errors except one were made when the speakers attempted to produce agent-initial word orders in the patient-cued condition. All the speech errors were associated with choosing a proper case marker: The participants tried to produce the accusative form for the cued patient and then switched to nominative one, appropriate for the agent.

5.1.4.3. Sentence onset latencies

The sentence onset latencies were measured as the time between the onset of the audio signal that accompanied the presentation of the target event and the moment the participant started to describe the target event. It took participants longer to generate sentences in the patient-cued condition than in the agent-cued condition (444 msec and 300 msec, accordingly). An ANOVA performed on the sentence onset latencies confirmed that this difference was significant ($F1(1,14) = 25.853$; $F2(1,31) = 32.525$). The mean speech onset latencies by cueing conditions are presented in Table 5. The reaction time data were informally compared between the following word order groups:

1. agent-initial(the agent-cued) vs. agent-initial(the patient-cued);
2. agent-initial(the agent-cued) vs. patient-initial(the patient-cued);
3. agent-initial(the patient-cued) vs. patient-initial(the patient-cued).

It took participants 182 msec longer to start the agent-initial sentences in the patient-cued condition than in the agent-cued one. It took participants 82 msec longer to initiate patient-initial sentences in the patient-cued condition than the agent-initial sentences in the agent-cued condition. Finally, it took the participants 101 msec. longer to initiate the agent-initial sentences in the patient-cued condition than the patient-initial sentences in the same condition.

As Table 2 demonstrates, the Object in some OVS constructions used in the patient-cued condition was described with the help of a pronominal form that does not assume a case marker in Russian (*ee* [Pr(Acc)]). Similarly, in some SVO utterances produced in the agent-cued condition the pronominal form was used for the Subject of the sentence (*ona* [Pr(Nom)]). Because constructing these utterances does not involve case marking, the onset latencies to initiate such sentences may be informative as to whether the range of word order options in Russian facilitated or complicated participants' performance. I compared the speech onset latencies between these types of utterances. Participants were 118 msec faster to initiate case-marked sentences than the sentences with the pronominal NPs.

5.1.5. Possible effects of structural priming

One concern about interpreting the results of a Fish Film study is related to structural priming – a tendency to recycle the syntactic structure that was used in earlier discourse (Bock, 1986). A larger dependence of Russian speakers on canonical SVO may reflect the fact that they were influenced by previously produced SVO sentences. I find it necessary to discuss why and how structural priming does not present a problem for the interpretation of the experimental results.

First, structural priming cannot explain the difference between the Russian data reported here and the English data obtained by Tomlin. Performing on exactly the same task, English speakers consistently varied assignment of the grammatical voice between conditions while Russian speakers did it to a much lesser degree. The obvious question then is: if Russians get primed by their canonical word order in the Fish Film task, why doesn't the same happen to English speakers? Second, the target sentences in the Fish Film task are separated by the non-target descriptions of the interactions that happen before the target event (e.g. *Here come blue fish and red fish. They approach each other. The blue fish gets closer to the red one.* etc.). So, if structural priming were to occur in the experiment, it would have to occur over a substantial lag – at least 3-4 intervening sentences of varying structure: Conjoined NP sentences, intransitive sentences, sentences with symmetrical predicates, etc. Similar sentences are often used as legitimate fillers in structural priming studies.

However, I tested the collected data for possible effects of structural priming. Two effects might be expected if structural priming had an independent effect on the data. (1) The likelihood of producing a particular structure should increase if the last produced sentence contained the same structure, and this effect should be *independent* of the cue location. Second, one should expect reaction time facilitation when the same word order was used repeatedly (cf. Corley & Scheepers, 2002; Smith & Wheeldon, 2001). To address the first issue, I tested whether the probability of producing SVO was affected by the fact that the previous structure was also SVO. This hypothesis was not supported. A one-sample t test on the calculated probabilities against 50% distribution returned a non-significant result ($t(14)$=.743, ns). To address the second issue, I separated the 32 experimental trials into (1) the matching cue trials – the consecutive pairs that cued participants' attention to the same referent, and (2) mismatching cue pairs – the consecutive pairs that cued attention to different referents. I then analyzed the corresponding pairs of narratives. In the matching cue pairs, the word order was "recycled" in 97 % of cases. In the mismatching pairs it happened in 55 % of cases. In the matching pairs, perceptual and structural priming are indistinguishable because they effectively prime the same structure. In the mismatching pairs, perceptual and structural effects should work in opposition of each other if structural priming contributed to the result. The reaction time benefit to initiate the second of the two matching structures was observed only in matching cue trial pairs (138 msec). But it did not occur in the mismatching cue pairs – the speech onset latencies were on average 8 msec longer for the second sentence even when the same word order was produced. These analyses show that the choice of word order in Fish Film task cannot be attributed to the effects of structural priming.

5.1.6. Discussion

The results of Experiment 1 suggest the following. Attention to the patient of the event significantly improved its chances of assuming starting point in target sentence. Therefore, the choice of word order in Russian was influenced by the location of the perceptual cue. This, for the first time, provides evidence for consistent attention-syntax interplay in a language dramatically different from English. This provides further support to the idea of a consistent functional relationship between distribution of attention in a perceived event and the choice of the sentential starting point as the determinant of the resulting word order.

On the other hand, the data revealed a pattern of attention-syntax interaction somewhat different from English. The agent-initial structures were largely preferred in both cueing conditions. It is an important difference from the Fish Film results using English (Tomlin, 1995) as it suggests that the effect of attentional focus on the choice of word order may not be uniform across languages. This, in turn, suggests a stronger tendency for speakers of Russian to rely on the canonical grammar and to a possibility of reciprocal effects from language to perception during online sentence processing.

One theory that predicts this type of interaction is *Thinking for Speaking* (Slobin, 1996; 2003). Thinking for speaking relates the dynamics of online language processing to a set of constant bi-directional interactions between language and thought. As such, it predicts "pervasive effects of language on selective attention and memory for particular event characteristics" (Slobin, 2003: 158). Two recent cross-linguistic studies supported this idea by using visually mediated production tasks. One study (Trueswell, 2005) tested whether the visual scanning patterns produced by speakers of English and Greek would differ depending on the preferences to encode the manner of motion accommodated in the language grammar. Participants received one of the two instructions: (1) to describe visually presented events with the verbs of manner or (2) silently inspect the pictures of these events In The results confirmed that when *describing* the pictures, speakers of English and Greek attended differently to the content of the events as a function of the differences in how their languages tend to encode motion: English participants preferred to visually interrogate the goal of the perceived event while their Greek counterparts continuously scanned the motion trajectory. However, there were no such language-related differences when English and Greek participants silently inspected the pictures. The second study (Bock, et al., 2003) analyzed English and Dutch descriptions of analogue and digital time displays. Speakers of English and Dutch differ in how they prefer to report time. For example, speakers of American English prefer to report time using absolute expressions, such as *"five thirty"* while Dutch speakers are more likely to report the same time using relative expression such as *"half zes"* (or "half six" in English rendition). The tendency of American undergraduates to use absolute expressions correlated with several problems in their timetelling performance with analogue clocks. The eye-movement analysis supported this result. Visual scanning patterns mirrored the timetelling preferences encoded in English and Dutch. These reports support conclusions about Russian speakers' performance in Experiment 1.

The reaction-time analysis revealed that it took Russian speakers longer to initiate non-canonical patient-initial constructions than the canonical agent-initial constructions. This supports the idea that canonical sentences are in general easier to produce than non-canonicals (e.g., Tannenbaum & Williams, 1968). Surprisingly, Russian speakers did not produce many passives in the patient-cued condition although given the early commitments to the Nominative forms and the time pressure of the task, one could expect this structure to be used more frequently regardless of its rarity in regular Russian discourse. It is possible that an "alien" nature of this structure for Russian grammar and its rare appearance in regular discourse makes it exceptionally difficult to retrieve even when the discourse situation prompts its usage (cf. Gennari, et al., 2005).

At the same time, Russian speakers were slower to produce agent-initial constructions in the patient-cued condition than the same type of constructions in the agent-cued condition. This delay probably reflects the time penalty associated with switching from early commitment to the Nominative to the Accusative word form. Together with the higher error rates and the nature of the SVO-related speech errors in the patient-cued trials, this suggests the reluctance to abandon the initial case commitments associated with the preferred SVO. The reaction time penalty may also reflect an operational conflict between the cueing scheme and the word order. One possible scenario suggests that the commitment to the agent-related word form in patient-cued trials resulted in shifting attention from the patient to the agent in order to arrive at the preferred SVO. The magnitude of the effect is comparable to the latencies reported in other studies that involved oculomotor saccades related to shifting attention during sentence production (e.g. Altmann & Kamide, 2004).

Finally, initiation of agent-initial sentences in the agent-cued condition occurred faster than the initiation of patient-initial sentences in the patient-cued condition. This was true both when the latter structures started with a case-marked word form (*ryb-a* [N(Acc)]) and a non-marked word form (*ee* [Pr(Acc)]). This suggests that the Russian participants did not benefit from the availability of a wider range of syntactic choices and effectively contradicts the results of some earlier studies (Ferreira V., 1997). It is, however, noteworthy that in Ferreira (1997), the structural alternatives were roughly equi-biased while SVO is associated with a clear dominance status in Russian.

5.2. Experiments 2 and 3. Regular transitive sentence production in English and Russian

Experiment 1 provided initial insights into the performance of the language-perception interface across languages with different grammatical systems. These results seem to be comparable with the results of some other studies and predicted by some theories. However, the resulting picture is incomplete and suggestive at best. First, the Fish Film paradigm does not provide any independent control of attention. Second, the use of the same materials from trial to trial may influence the referent's memorial activation status and as such is a potential confound to the observed results. Third, virtual absence of fillers between trials makes routinization a possible strategy and, together with event repetition, may lead to structural priming from trial to trial. Most importantly, Fish Film relies only on two types of data: the probability of making a particular structural choice and the name onset latencies during sentence production. This leaves out every psychological operation that occurs before the sentence formulation has started.

In the experiments that follow I tried to avoid these problems. First, Experiments 2-9 all used eye-tracking methodology. Eye tracking provides accurate control of eye movements that accompany participants' behaviour in experimental tasks. Because visual search typically accompanies shifting of the attentional focus, control of eye movements can reflect distribution of participants' attention in the described display. From the processing point of view, eye-tracking provides an early window into the speaker's visual interrogation patterns during sentence production. The ordering of the initial gazes, the time spent looking at a particular referent, and visual scanning patterns throughout the course of sentence production allow a more comprehensive and complete picture of what happens in the speakers' minds before and during the formulation of sentences about visually presented events. Also, in each of the following experiments, participants described a series of novel events that differed from one target trial to another. The use of black-and-white line drawings helped to avoid possible problems with colour prominence. Finally, the presence of at least two fillers between the target trials reduced possible effects of residual activation from one target trial to the next.

In a way, the introduction of eye-tracking methodology led to the decision to "start from scratch". Encouraged by the results of Experiment 1, I decided to step back and collect the norming data using the whole flexibility provided by eye-tracking. Thus, the next two experiments utilize a no-cue picture description task in order to compare regular sentence production in English and Russian.

5.2.1. Experimental design

I used a very simple design in Experiments 2 and 3. English and Russian participants extemporaneously described pictures of simple transitive events using any word order available in their language. The results of this study provide a no-manipulation baseline for further comparisons with the data obtained under various combinations of perceptual, semantic, and syntactic cues. To obtain accurate base-line data, it was important to have no independent cueing procedure, perceptual or otherwise. This way, speakers could arrive at any word order available in their language. I, however, expected the referent's agenthood and event causality to act as implicit cues for the choice of the sentential subject as well as the subsequent word order. To avoid other salience factors, the described entities were controlled for size, colour, animacy, humanness, and the position of referent presentation (e.g. left or right of the picture).

5.2.2. Method

5.2.2.1. Materials

Figure 8 (a, b) provides examples of the materials used in the both experimental tasks. Two sets of materials were used. The first set contained 17 pictures of single entities not involved in any action. These pictures were used in the Naming task. The second set contained simple black and white line drawings of transitive interactions between the characters portrayed in the Naming set. The Event set was used in the event description task. The event set consisted of 48 materials using the following verbs: *kick, pull, punch, push, touch,* and *shoot* (Russian renditions *pinat' tjanut', bit', tolkat', trogat', and ubivat'*). Fillers were 98 pictures of intransitive events (e.g. *sleep, run, yawn, clap*). A minimum of 2 fillers separated the target trials. There were 4 filler trials at the beginning of each experimental session, and there were no filler trials after the last target trial. The total number of trials was 146 in both experiments.

For the purposes of eye-tracking analysis, I pre-coded one interest area in each of the Naming pictures and three interest areas in each of the event pictures for further fixation analysis. The interest area for referent in the Naming pictures included the single entity presented on the screen and the surrounding area of approximately 2^0. The interest areas for agent and patient of the event included the corresponding entity together with a surrounding area of approximately 2^0. The interest area for the action encompassed an instrument of an action or the area between the two protagonists. Although such interpretation involves a certain theoretical leap of faith, the current studies used the logic proposed by Griffin and Bock (2000). Also, as our analysis demonstrates, speakers do tend to produce distinct clusters of fixations to the specified action areas during the apprehension of the event, and they tend to fixate these shortly before the production of the sentential verbs. The same parameters for the interest area coding were used in all the remaining experiments.

5.2.2.2. Participants

15 Native speakers of English (7 female) and 15 native speakers of Russian (10 female) participated in the study. All the participants had normal or corrected-to normal vision. All English speaking participants were undergraduate students at the Department of Psychology University of Glasgow; Russian speaking participants were recruited from the staff at Moscow Institute of Neurorehabilitation. The mean age was 21,43 years for the English participants and 27,1 for the Russian participants.

5.2.2.3. Apparatus

The English sample was collected using an SMI Eye Link II head-mounted tracker. An SMI iView tracker was used to collect the Russian sample. In both experiments, materials were presented on a 17’ monitor with the 75 Hz refresh rate. A SONY DAT recorder was used to collect the speech data in both cases.

5.2.2.4. Procedure

Before the session began, participants were instructed that the main purpose of the study is to analyze what people say when they describe events. Participants were seated in front of the monitor at a distance of approximately 60 cm between the eyes and the monitor. Prior to the experimental session, participants had a practice session, during which they viewed pictures of the referent characters presented alone. The characters’ names appeared alongside the pictures. The instruction was to view the referents’ pictures, say the names presented with them out loud, and remember these names and the pictures for the further tasks. After that, participants practiced describing sample event pictures – one for each event. There was no specific instruction as to how to describe the practice event pictures.

There were two consecutive tasks to perform during the experimental session. For the first task, participants were instructed to name the referents presented alone. This provided us with the time measure of independent lexical access and a pattern of eye fixations to the referents not involved in action. Also, this task helped to reduce the competition between synonymous names known to complicate production (Griffin, 2001) and to control for the accessibility of the referent names. For the second task, participants were asked to describe extemporaneously simple transitive interactions between the referents they named in the naming task.

Each trial started with the presentation of a fixation mark in the centre of the screen. After the participant successfully fixated it, the fixation mark moved to the lower area of the screen to insure that the saccade was made to one of the three interest areas when the target picture was displayed. The presentation of the target display was gaze-contingent with the presentation of the previous fixation mark. The fixation mark that preceded the target picture was equally distant from all the three interest areas. The participants had to look at the second fixation mark for a minimum of 200 msec. in order for the target picture to be displayed. The target picture then appeared in the centre of the screen. The participants were instructed to use a single sentence to describe the target picture. They were also instructed to describe the target pictures extemporaneously; right after the picture onset. Once they finished describing the picture, participants pushed the spacebar key, and the next trial would begin. If a participant failed to produce a description, the target picture was removed from the screen after 7700 msec offset. An auditory signal accompanied presentation of each target picture; therefore participants' speech was recorded time-locked to their eye-tracking data.

5.2.3. Results

Because the distribution of the reaction time data is typically asymmetrical, I calculated harmonic means to normalize the data distribution for the analysis of the data in this and all the remaining experiments. The missing values for each data set were replaced by the mean values for the corresponding condition.

5.2.3.1. Naming

The participants used the names suggested to them during the practice session in 97.6 % of the trials. The syllabic word length was measured and compared between the two languages (see Table 6). A paired sample t-test returned a non-significant result, therefore a slight difference between English and Russian nominative word forms was not used as a covariate factor for further analyses.

Figure 9 compares the main reaction times from the Naming task for both languages. The mean value for the first fixation onset in both languages was around 200 msec. Analyses of Variance (ANOVA) revealed no effect of Experiment on the first fixation latencies. However, Russian speakers were slower to initiate the names than their English counterparts (1226 msec vs 887 msec). This difference was reliable ($F1(1, 29)=21.755$; $F2(1,16)=44.439$).

Unsurprisingly, there was no difference between the two languages in the latencies to initiate the first fixations to the referents as first fixations reflect attentional processing independent of linguistic operations. Although statistically insignificant, the difference in the syllabic length between the two languages may explain a reliable delay in the initiation of the Russian names. However, it might also be due to the necessity to access Nominative case markers in Russian – a property not represented in English (unless, of course, Russian word forms are stored with the nominative case markers and not as roots). This result supports the findings reported in Experiment 1.

5.2.3.2. Event Descriptions

The names practiced during the Naming stage were used in 94.25 % of all trials during the Event stage. The syllabic complexity of those names was once again measured and compared between the two languages (see Table 7). Constructing a Russian transitive sentence requires assigning a specific ending for each sentence constituent: Nominative Case marker for Subject, Number and Gender marker for Verb, and Accusative Case marker for Object. Paired sample t-tests revealed that English and Russian names differed significantly in the syllabic word length for Verb and Object (Verb: $t(16)=3.425$; Object: $t(16)=9.670$,). To account for the syllabic length factor, I used syllabic complexity as a covariate in the corresponding comparisons between the languages.

5.2.3.3. Word order analysis

The descriptions of the events that did not conform to the specified instructions (e.g. *There are a boxer and a swimmer in the picture* or *A boxer and a swimmer are holding hands*) were counted as missing values in these and all the remaining experiments.

In Experiments 2 and 3 speech data were coded using the following categories: Active Voice and Passive Voice – for English; SVO, OVS, Passive Voice, and other – for Russian. The missing values comprised 2.65% of the data in Experiment 2 and 4% of data in Experiment 3. Table 8 illustrates the distribution of word order choices between the languages in the rest of the data. It is clear that Active Voice SVO order, canonical for both languages, was greatly preferred to all other structures in both experiments. Therefore, in the absence of other cues, the event's causal structure (*agent-event-patient*) was used to assign the word order.

5.2.3.4. First fixation analysis

The values for the first fixation analysis were calculated as the temporal lag between the onset of the picture and the onset of the corresponding fixation to the interest area. When no fixation to the interest area occurred during the trial, the corresponding value for the area was coded as missing value. The threshold for the individual fixations in all experiments was set at 50 msec. Cumulative fixations to the same area for longer than 150 msec. were aggregated into gazes for the eye-voice span analysis. The same logic for the fixation onset analysis was used in all the remaining studies.

In Experiment 2, missing values accounted for 3.4% of initial fixation; in experiment 3, there were 2.9% of such values. Figure 10 summarizes the first fixation data for each of the target regions during the event description phase and compares it between the languages. The first important observation follows from the fact that the early perceptually-driven apprehension of the event proceeded from the action area to the agent, and finally – to the patient. A similar process with initial looks concentrated on the action was reported in one comprehension study (Knoeferle, et al., 2005). Seemingly, it is essential for both speakers and listeners to address the area containing the information about the action to be able to disambiguate the referent roles of the interacting protagonists. Then, participants looked at the agent, and finally at the patient.

ANCOVAs with First Fixations as the dependent variable and Experiment as the independent factor demonstrated that there was no significant difference between the two languages in the first fixation onsets at the action area and at the agent area. However, a 175 msec slow-down at the patient area observed in Experiment 3 was reliable ($F1(1, 29)=14.149$; $F2(1, 47)=5.093$): Russian participants lagged behind their English counterparts when visually interrogating the event's patient.

5.2.3.5. Eye-voice span analysis

The basic logic for extraction of the eye-voice spans remains the same in all the remaining experiments of this thesis. Following Griffin and Bock's (2000) procedure, I calculated eye-voice spans as the temporal lags between the onsets of the last fixations to the referents immediately preceding the production of the corresponding names and the onsets of corresponding names themselves. Because these "last fixations" were only used for the calculation of eye-voice spans, they are not reported separately. If no fixation to the referent occurred during the target trial, or there was no name onset value available, the EVS value was coded as missing. Finally, the protocol for extracting eye-voice spans "evolved" from one experiment to another. This is why the developments in eye-voice span extraction are discussed separately where necessary.

In Experiments 2 and 3 a coder blind to the experimental hypothesis used two worksheets in order to obtain the EVS measurements. One sheet contained the name onset data, and the other – the gaze protocol fro each participant. The EVS values, therefore, were calculated "backwards": first, the coder identified the name onset signature for a particular sentence constituent and then used it to locate the fixation to the corresponding interest area immediately preceding this onset. The resulting temporal lag represents the corresponding EVS.

Figure 11 summarizes mean EVS values observed in the event description task in both studies. Informal examination of the EVS data demonstrates that EVS values for the Subjects were 242 msec smaller than those for the Objects in Experiment 2 (English), while in Experiment 3 (Russian) this difference was reversed: The EVS values for the Subjects were 117 msec larger than those for the Objects. Also, EVS values for the Verbs were substantially larger in Experiment 3 than in Experiment 2.

ANOVAs performed on the mean EVS values with Constituent as the independent factor revealed a reliable effect of the latter on the EVS values in Experiment 2 ($F1(2,28)=36.629$; $F2(2,94)=9.922$). Test of within-participant and within-item contrasts revealed a significant linear trend for the EVS magnitude to grow from the Subject to the Verb to the Object ($F1(1,14)=61.754$; $F2(1,47)=23.297$). Comparisons between sentence constituents revealed a reliable increase from the Subject to the Verb ($F1(1,14)=38.749$; $F2(1.47)=6.135$) and from the Subject to the Object ($F1(1,14)=61.754$;

$F2(1.47)=23.297$); the increase from the Verb to the Object was not significant. This pattern demonstrates that the processing difficulties associated with relating referents to the corresponding constituent names increased linearly during the production of English transitive sentences from the Subject to the Verb and then stayed flat from the Verb to the Object.

The same analysis in Experiment 3 also returned a reliable effect of Constituent on the EVS values during Russian sentence production ($F1(2,28)=20.557$; $F2(2,94)=31.155$). However, linear trend proved to be insignificant this time while quadratic trend was significant ($F1(1,14)=24.870$; $F2(1.47)=36.313$). The comparisons between the data for each sentence constituent revealed a reliable increase from the Subject to the Verb ($F1(1,14)=20.956$; $F2(1.47)=20.822$), a reliable decrease from the Verb to the Object ($F1(1,14)=24.279$; $F2(1.47)=48.584$), and a reliable difference between the Subject and the Object ($F1(1,14)=8.159$; $F2(1,47)=13.815$). The EVS analysis of the Russian data in Experiment 3 demonstrated that (1) it was easier to process Objects than Subjects during the production of Russian transitive sentences and that (2) the processing load rapidly increased from the Subject to the Verb, and then dropped to its lowest at the Object.

ANCOVAs using the mean EVS values as the dependant variable, Experiment as the independent variable, and Syllabic Complexity as a Covariate revealed significant differences between the two languages for the Subject ($F1(1,29)=19.050$; $F2(1, 47)=48.774$) and the Verb ($F1(1,29)=19.050$; $F2(1, 47)= 21.264$) but not for the Object. Once again, there was no effect of syllabic length on the eye-voice spans for any of the constituents. This result suggests that the increasing difficulty experienced by the Russian participants was only observed at the Subject and the Verb, and that the EVS performance of both English and Russian speakers at the end of the sentence was comparable.

5.2.3.6. Name onset analysis

Speech data in all the remaining studies were analyzed in terms of the latencies to produce sentence constituents: subject, verb, and object. These onsets were calculated as the temporal lags between the onset of the target picture and the onsets of the corresponding constituent's name. The starting point responses with onset latencies less than 300 msec. and more than 5000 msec. were coded as missing values. In Experiment 2 this resulted in 2.9 % of data loss; in Experiment 3 – 3.8% of data loss.

Figure 12 presents the mean name onset latencies for each of the sentence constituents for the two languages. Informal examination of the name onset data collected in the two studies demonstrates that Russian participants were slower than their English colleagues to initiate every sentence constituent (Subjects: 320 msec, Verbs: 332 msec, and Objects: 410 msec). ANCOVAs with Syllabic Length as a Covariate demonstrated an effect of Language on name onset latencies for Subject ($F1(1, 29)=14.376$; $F2(1, 47)=43.804$), Verb ($F1(1, 29)=10.893$; $F2(1, 47)=27.847$), and Object ($F1(1, 29)=12.221$; $F2(1, 47)=6.311$). There were no separate effects of syllabic length on the name onset latencies for any of the sentence constituents. The analysis of the temporal lags between the constituents revealed no effect of Experiment on the lag from the Subject to the Verb and a reliable difference for the lag between the Verb and the Object ($F1(1,29)=6.620$;$F2(1,47)=12.011$). This pattern of results demonstrates that the production of Russian transitive sentences consistently lagged behind the production of the English equivalents slowly growing from the Subject to the verb and then to the Object.

5.2.5. Discussion

In the absence of specific perceptual, conceptual, and linguistic cues, speakers of both English and Russian overwhelmingly chose the canonical event causality (from the action initiator – to the action – to the action experiencer) and the corresponding SVO word order when describe the pictures of transitive events in a non-manipulated environment.

Analysis of the initial fixations to the event's information-laden areas revealed the dynamics of the early perceptual processes involved in the preparation of sentences. Speakers of both languages preferred to first look at the action area, which provided them with the raw resolution of the event's referential scheme. This preference was revealed by a non-arbitrary ordering and the distinct clustering of fixations to the instrument of the event and/or to the area between the referents as corresponding to the information conveyed by the verb. At the same time, one has to acknowledge that the current experiments do not provide enough evidence to strongly confirm such explanation. Speakers tend to preferentially interrogate centrally located information in the display, and that is exactly where the instrument or the area between the referents appeared to be. The best way to tease apart linguistic and purely visual accounts of the action information extraction would be to place the instrument of the event outside of the central region of the picture.

The fact that the agent area consistently attracted speakers' attention after the action area suggests that already during the perceptual analysis the production system was biased toward interpreting the event according to the canonical causality. It indicates that speakers do not interrogate events randomly, at least not if they know that they will need to speak about them. Instead, they attempt to quickly apprehend the event in a fashion that saves them cognitive effort during early processing stages. This property demonstrates that although the speaker's conceptualization of the described event can be under conscious control, in default or decontexctualized setting it relies on a certain amount of automaticity prompted by the naturally assumed event causality. (cf. Garrod & Pickering, 2006). The experiments discussed in the remaining part of the thesis will demonstrate that the preference for early automatic conceptualization based on the preferred causality is very persistent, even when the surrounding environment suggests otherwise.

The chronometric analysis of the initial fixations revealed that their timeline at the first two areas was similar between English and Russian. However, a small time penalty for Russian speakers was revealed at the Patient, which was fixated last by speakers of both languages. This fact may suggest a partial overlap between the perceptual analysis of the event and the later operations of conceptual preparations for sentence formulation.

The analysis of the eye-voice spans provides understanding of operations that belong to both conceptual and linguistic processing. On one hand, eye-voice spans reflect the completion of the conceptual analysis; on the other, they tap into the lexical processing associated with the incremental production of names in a sentence. The pattern of the EVS results in Experiment 2 demonstrates that relating referential candidates to the corresponding grammatical roles in an English sentence followed a pattern of incremental growth of the processing load until the Verb was selected. This result provides support for the results of an earlier study (Ferreira, 1997), which demonstrated that sentence formulation is easier when speakers can entertain more choices than when such choices are limited. In other words, structural ambiguity facilitates rapid selection of forms, while the necessity to conform to one existing variant hinders it. This fact is usually taken as supporting an incremental account of sentence formulation. Producing an English transitive sentence involves the largest degree of ambiguity at the Subject, disambiguation happens at the Verb, and after that there is no more ambiguity. The direction of the EVS magnitude change observed in Experiment 2 mirrors such progressive narrowing of choice.

In Russian, a very different pattern was observed: (1) processing load was already quite high at the Subject, peaking at Verb, and rapidly dropping at Object to a value similar to the one observed in the English participants' performance. This pattern suggests that in Russian sentence production, the full event analysis and the early selection of the sentential frame is necessary before sentence production can start. On one hand, a larger inventory of Russian word order should logically lead to a facilitation similar to the one reported by Ferreira. On the other, Russian word order relies on the assignment of specific lexical markers, which effectively constrains the word order flexibility effect already at Subject: Once the nominative case marker is assigned, the word cannot perform any other role than that of the Subject of an active voice SVO. In this situation, speakers cannot be sure that everything goes well until at least the Subject NP and Verb lexical forms have been selected. So, greater flexibility of word order combined with obligatory case marking led to the necessity to analyze the event in its full complexity before sentence formulation could start. However, an alternative explanation is also possible: There may simply be a stronger tendency to use SVO in Russian, whereas the use of SVO and Passive in English is more equi-biased. This point of view is, although theoretically possible, is hard to accept for two reasons. First, it is unclear why speakers of one SVO language would be more biased to use it than the speakers of another SVO language. Second, the eye-voice span data comparison speaks in favour of the idea that Russian speakers experience larger processing load at the starting point of the sentence formulation. The latter result supports the case-marking interpretation of the stronger propensity of Russians to use SVO.

Comparison between the two languages revealed longer eye-voice spans for the sentential Subjects and Verbs in Russian sentence production but not for the Objects. Larger processing loads at the essential sentence constituents in Russian are not surprising, because processing is evidently more difficult when the speaker cannot analyze the event and formulate a corresponding sentence in parallel, like in English. However, once the Subject and the Verb are selected, processing the rest of the referential structure is determined in both languages. Therefore, the difference between the two languages disappears at the Object.

Interestingly the difference in the EVS values between the two NP constituents followed different directions in the two languages: Processing the English Subjects was easier than processing the Objects while there was an opposite effect in Russian. This result strengthens my earlier argument about different degree of event analysis essential for the formulation of sentences in these two languages. Smaller EVS values for the

Subject than for the Object reflect shallower processing at the beginning of the sentence; because more conceptual and lexical processing is necessary later, the EVS values at the Object become inflated. A reverse pattern in Russian demonstrates a greater amount of processing at the beginning than at the end of the sentence.

The analysis of name onset latencies revealed a general tendency for Russian speakers to lag behind their English counterparts at each stage of sentence formulation. This difference may in principle reflect both longer words used by Russian participants and the eye-voice span difference between the two samples. Lack of a separate effect of the word length on the name onset latencies dismisses the first possibility. As I have noted before, the EVS difference between the two languages disappeared at the Object, but the name onset related difference increased. In fact, a comparison of the time lags from the Verbs to the Objects between the two languages returned a reliable result. This result reveals a separate effect of a greater lexical access difficulty that accompanied Russian sentence production. This observation supports one of the conclusions in Experiment 1: Case marking should be treated as an additional processing operation. Such marking-dependent difference remained noticeable even when EVS difference between the two languages disappeared. Separate analysis of the EVS and the name onset data helps to tease apart the word order effects related to difference in the conceptual analysis from the case marking effects in Russian sentence production.

5.3. Experiments 4 and 5. Perceptually-primed transitive and ditransitive sentence production in Finnish

Experiments 1, 2 and 3 together provided important insights into the functional and chronometric properties of visually-mediated sentence production systems in English and Russian. So far, it seems possible to conclude that grammatical features such as case marking and the relative combinatorial freedom may complicate making rapid choices between word orders in the situation when one of the many referents tries to assume an early placement in a sentence in accordance with its salience. A greater dependence of Russian speakers on the canonical word order in Experiment 1, slower production rates and longer eye-voice spans at the initial stages of sentence formulation in Experiment 3 demonstrated that grammatical features supporting a wider inventory of word order come at a cost of slower mapping of the message semantics onto the language syntax. At the same time, Experiment 1 confirmed a strong coupling between the assignment of the sentential starting point and the attentional focus. Speakers of Russian produced reliably more non-canonical OVS sentences when their attention was directed to the event's patient. The general preference to produce canonical SVO resulted in a specific attentional switch and associated reaction time penalty when the cue directed participants' attention to the patient.

However, the reliable effect of cueing in Experiment 1 does not provide direct evidence for the attentional account of word order assignment. One reason to be cautious is a relative weakness of the effect itself. Speakers of Russian seemed to be quite stubborn in their preference for the canonical SVO even when their attention was strongly attracted to the event's Patient. Also, the non-arbitrary pattern of the progression of gazes during the rapid apprehension in Experiments 2 and 3 suggests that speakers may make early commitments to the structure they are going to produce based on the event causality and referential properties that support the canonical grammar of their language rather than the salience map. So, the question of whether it is the attentional focus that predicts the order of mentioning or the early commitment to a particular structure that biases preferential looks to the referents remains open. I designed two eye-tracking experiments using Finnish to further investigate this issue.

The two experiments reported below continue using the perceptual priming paradigm. However, there are some important differences in (1) the inventory of the elicited syntactic structures and (2) the ways of controlling participants' attention. In previous studies I have tried to address the production of transitive sentences. The visual support for this type of sentences typically involves relatively simple two-referent scenes. If one were to assume that sentence production is not 100% incremental and that speakers have to do a certain amount of conceptual analysis and syntactic pre-planning before sentence formulation begins, the difference in the referential complexity might become an important issue: The more information there is to process before sentence formulation starts, the more complex and delayed the rapid apprehension of the event should be. A greater amount of semantic processing should make coupling between the element in the attentional focus and the sentential starting point problematic. This should be the case because the perceptual effects take their toil on the speaker's performance very early; if the event is simple, they are more likely to survive through conceptual and syntactic stages of formulation. On the other hand, if sentence production is strictly incremental, one could expect the focused referent to continue to appear before the rest of the sentence regardless of the event's complexity. If so, referential complexity should not produce any effect on the "survival" of the attentional effects during conceptual analysis and sentence formulation. According to this scenario the probability of encountering the attended referent at the beginning of a sentence should not suffer from the relative complexity of the described scene.

Also starting with Experiments 4 and 5, I introduce a different cueing paradigm. Instead of a strong explicit cue like in Experiment 1, an implicit perceptual cue (Posner, 1980) is used to attract speakers' attention. An implicit cue constitutes a cue that effectively captures attention while the participants remain unaware of the manipulation. Such attraction of attention is possible via presenting a cue for a very short duration of time, e.g. 50-70 msec. Participants typically do not *explicitly* notice the presence of such a cue while it successfully captures their attention[3]. The importance of implicit attentional cueing derives from the issue of strategic planning. When the cue remains on the screen for a time sufficient for conscious processing, the cue itself and its location may become informative about the placement and, sometimes, the nature of the subsequently presented stimuli.

Two experiments reported below use an implicit cueing paradigm in order to compare performance of Finnish speakers on transitive (Experiment 4) and ditransitive (Experiment 5) sentence production while their attention is implicitly attracted to one of the referents of the visually presented event.

5.3.1. Some facts about Finnish grammar

---

[3] One needs to be cautious with the "not noticing" issue. What really happens in implicit cueing tasks is that participants notice *something* when the cue is presented. When asked to report on whether they saw the cue or not they typically report that they thought something was wrong with the monitor or something briefly flickered on the screen. At the same time, they remain naïve about the purposes of the cueing manipulation.

Finnish belongs to the Finno-Ugric family of languages. Like Russian, it uses a case marking system to denote the sentential roles. It is a typical SVO language although all possible scrambling alternations are possible (Vilkuna, 1995). Production of non-canonical sentences is usually ascribed to discourse factors, such as theme vs. rheme or old vs. new information. As with Russian, old information tends to appear before new in a Finnish sentence. Some recent studies support this idea. For example, manipulations of the discourse status of the referents were demonstrated to lead to faster comprehension of Finnish non-canonical sentences (Kaiser & Trueswell, 2004).

Finnish nouns receive one of the 14 grammatical cases. For a transitive event, the Agent receives Nominative case marking while the Patient receives Accusative case marking. For example, an event portrayed in Picture xx can be described with the help of Finnish SVO – "*Nyrkkeilijä lyö karjapaimenta*" or OVS – "*Karjapaimenta lyö nyrkkeilijä*" with other scrambling options also available. Passive constructions in Finnish are also available but they are typically agent-less. For example, one could say in Finnish: *"Karjapaimenta lyötiin"* (The cowboy was hit). The sentence indicates that, indeed, somebody intentionally hit the cowboy, but the actor in not explicitly mentioned.

As far as the ditransitive sentence production goes, Finnish supports general Subject-Verb-Object organization of the sentence, and it permits frontal placement of both direct and indirect object (or *the theme*). I will refer to these two types of constructions as SVTO and SVOT accordingly (DO-first and IO-first in (Kaiser, 2002). The direct object in both cases receives the accusative case marker while the indirect object – the allative case marker. Allative case works as a substitution to the Dative case with the general meaning of a movement or transfer *to someone.* In pragmatically neutral contexts, SVTO order is typically preferred to SVOT order. Pragmatic factors such as "old vs. new" information make SVOT more likely if the indirect object was mentioned in the preceding discourse (Kaiser, 2002).

5.3.2. Method

5.3.2.1. Materials

Two sets of the materials were used in Experiments 4 and 5 accordingly. Both sets contained simple black and white line drawings of interactions between two human characters. The set for Experiment 4 contained pictures of transitive events, and the set for Experiment 5 contained pictures of ditransitive events.

The transitive set was virtually the same as the one used in Experiments 2 and 3. It contained 64 pictures of simple transitive events counterbalanced for left-right presentation, size, animacy, color, and referent role suggestibility. Materials were not controlled for corpus frequency; therefore I provided a comprehensive preview and practice session for the participants to get familiarized with the referents they would encounter in the task. There were 17 human referents used in the target trials. 64 target materials comprised 8 different transitive events (CHASE, HIT, KICK, PULL, PUNCH, PUSH, SCOLD, SHOOT, TICKLE, and TOUCH) with 8 different combinations of referents. 130 filler pictures of intransitive events (e.g. SLEEP, RUN, YAWN, and CLAP) separated target trials. There were 4 filler trials at the beginning of each session, a minimum of 2 fillers between the target trials, and no filler trials after the last target trial. In 32 of the materials the cue was on the agent, while the other 32 had the cue on the patient of the event.

The materials for Experiment 5 were black-and-white pictures of ditransitive interactions between the same human characters that appeared in the previous experiments. A total of 96 materials comprised 6 different ditransitive events (HAND, THROW, SELL, GIVE, SHOW, and OFFER); therefore there were 16 pictures of each event with different referents in each event. There were 32 materials in the agent-cued, 32 – in the patient-cued, and 32 in theme-cued conditions. The same kind of fillers separated the target trials in the same fashion as in Experiment 4.

5.3.2.2. Participants

24 native speakers of Finnish (8 male) with normal or corrected-to-normal vision participated in each study. All the participants were undergraduate students at the Department of Psychology at the University of Turku, Finland. The mean age of the participants was 20.51 in the first and 21.88 in the second study. The participants took part in the study voluntarily for two psychology course credits. There was no material reward for the participation.

5.3.2.3. Apparatus

Experiments were created with the help of Experiment Builder software (SR Research, 2006). The eye data were extracted and filtered using Data Viewer software (SR Research, 2006). Participants' speech was recorded using SONY DAT digital recorder. Adobe Audition 2.0 software was used for the analysis of the speech data.

The experimental data in both experiments were collected with the help of SMI Eye Link II head-mounted eye tracker. The experimental materials were presented on a 17' monitor with the help of a DELL Optiplex GX 270 personal computer with the display refresh rate of 75 Hertz. The same apparatus was used in all the remaining studies.

5.3.2.4. Procedure

Figures 13 and 14 present examples of target trial sequences in Experiments 4 and 5. The participants were positioned in front of the monitor approximately 60 centimeters from it. They had a direct view of the screen throughout the trial. Before the experimental session began, the participants had a practice session. During the practice session participants saw the pictures of the referent characters that would later interact in the target trials of the experimental session. The referents appeared in the center of the screen one at a time. Characters' names were presented with the pictures, and the participants were instructed to pronounce those names out loud and remember them for the further tasks. Also, the participants practiced describing sample event pictures (one for each event) during the practice session. The pictures of the events were presented in the middle of the screen with the practice sentences beneath them. To make sure that both preferred and non-preferred structures were pre-activated before the experiment began, both SVO and OVS (for Experiment 4) and SVTO and SVOT (for Experiment 5) sentences had to be practiced by the participants. The participants had to produce both types of sentences out loud while looking at the same picture.

The instruction for the experimental session was to describe extemporaneously transitive (Experiment 4) or ditransitive (Experiment 5) interactions between the referents viewed by the participants in the practice session. Each trial began with the presentation of a fixation mark in the centre of the screen. After the participant successfully fixated the fixation mark, the verb which described the subsequently presented picture appeared on the screen. Participants read the verb out loud. Soon, the fixation mark reappeared in the lower area of the screen. This insured that the participants would make a saccade from the second fixation mark to the visually cued area of the screen once the cue is presented or, if the cue was overlooked, to the target picture. The fixation mark that preceded the cue was equally distant from all the three interest areas of the subsequently presented event. The presentation of the cue was gaze-contingent with the presentation of the second fixation mark. The participants had to look at the second fixation mark for a minimum of 200 msec. in order for the cue to be displayed.

The cueing was operationalized via the presentation of a red circle in one of the event's interest areas – Agent or Patient in Experiment 4 and Agent, Patient, or Event – in Experiment 5. The cue was 1 centimeter in the diameter, it was presented for 60 msec, and its presentation corresponded with the centre of each interest area. After the cue disappeared from the screen, the target picture was displayed in the centre of the screen. Then participants described the target picture using a single sentence. Once finished describing the picture, the participant pushed the spacebar button, and the next trial would begin. If the participants did not produce a sentence, the target picture disappeared from the screen after 7700 msec time-out. An auditory signal accompanied the onset of each target picture; therefore the participants' narratives were recorded time-locked to their eye-tracking data.

I pre-coded three visual interest areas in each of the materials for further fixation analysis. The interest areas for the *agent* and the *patient* included the corresponding referent together with a surrounding area of approximately 2^0 of visual angle. The interest area for the *event* in Experiment 4 (transitive sentence production) encompassed an instrument of an event or the area between the two protagonists. The interest area for both the *event* and the *theme* in Experiment 5 (ditransitive sentence production) included the instrument of the event and conformed to the same surrounding area parameters as the areas for Agent and Patient in Experiment 4.

5.3.3. Results. Experiment 4

5.3.3.1. Cueing efficiency

To assess the efficiency of the visual cueing procedure, I performed an analysis of the probabilities to fixate the cued region as a factor of the cue location. This analysis confirmed that the implicit visual cue used in the study successfully attracted the attentional focus in 83 % of agent- and 77% of the patient-cued trials. Two-tailed t-tests performed on the mean proportions as compared to chance (.5) returned a reliable result ($t1(23)=46.929$, $t2(63)=63.781$).

5.3.3.2. Word order analysis

The speech data were transcribed according to their syntactic structure using the following categories: SVO, OVS, Passive Voice, and other. The missing values comprised 1.03% of the data. Table 9 illustrates the distribution of word order choices in Experiment 4. It is obvious that Finnish participants preferred Active Voice SVO order to all other options regardless of the location of the cue when describing pictures of transitive events. This preference was slightly less expressed in the patient-cued trials, but this difference was not significant. Therefore, regardless of the location of the visual cue, the preference for the canonical word order overrode the preference for making the cued referent the starting point of the sentence.

5.3.3.3. First fixation analysis

Raw fixation data were filtered with the help of the Data Viewer program using previously specified thresholds. Missing values accounted for 3.75% of the data. Table 10 summarizes the harmonic mean times for first fixations to each of the interest areas as a factor of the cueing scheme.

The average onset to fixate the cued referent is very similar between the agent-cued and the patient-cued conditions. The average estimation of these first fixation onsets is on a typical saccadic scale – around 150 msec. This is by no means a surprising result because the cue used in this experiment was uninformative, and the only purpose of the cue was to direct participants' attention to the area where one of the referents would shortly appear. As expected, the comparison between these reaction times did not return a significant result (both $Fs<1$).

However, first fixations to the *non-cued referent* differed significantly between the cueing conditions so that the first fixation to the patient in the agent-cued condition occurred 130 msec. later than the first fixation to the agent in the patient-cued condition. This difference was significant ($F1(1,23)=72.979$; $F2(1,63)=4.666$).

An important observation about the pattern of first fixations comes from the analysis of the initial fixations to the event area. The experimental procedure allowed participants to preview the verb related to the event they were preparing to describe before the target picture itself was displayed. It would be natural to assume that this valuable information would result in very few looks to the *event* area during the description phase. However, this was only true for the agent-cued condition: participants fixated the event area 15% of the time. When it happened, it happened late in the sentence preparation process averaging at 838 msec., which is around the time the sentence production would begin. When the patient was cued, participants fixated the event area 63.4% of the time. These fixations occurred much earlier than the fixations to the event area in the agent-cued condition – at 198.23 msec. on average.

5.3.3.4. Eye-voice span analysis

I performed the EVS analysis for the Subjects and the Objects only. The eye-voice spans for the Verbs were not separately analyzed because participants fixated the event area very infrequently (especially in the agent-cued condition).

Although the chronometrical concept of EVS remained the same, starting with this experiment I utilized a very different procedure for calculating the EVS. A Pearl-based script made it possible to extract EVS automatically. The script used the text files containing a participant's data for name and gaze onset latencies for each trial. Each gaze onset corresponded to one of the pre-coded interest area: Agent, Event, or Patient. The name onsets were marked as corresponding to Subject, Verb, or Object of the event; each produced sentence also received a word order code, e.g. SVO[4]. The script used this set of markers in order to perform a loop search and eventually match a particular name onset to the relevant onset of the last gaze to the corresponding interest area. Because the

[4] In the current study virtually 100% of the produced sentences were SVO, so this marker was not particularly important for the EVS analysis. However, it will become more important for proper association of the interest area and the corresponding sentence constituent in further experiments with more considerable syntactic alternations.

search was automatic, it was necessary to manually check the extracted EVS data for (1) values less than 100 msec and (2) missing values. The EVS values of less than 100 msec were replaced with the ones calculated manually using the fixation immediately preceding the one detected by the script. The missing values accounted for 5.74% of the total number of trials. Such values were replaced with the mean EVS value for the corresponding condition.

Table 11 summarizes mean eye-voice span values across the sentence constituents. An informal examination of the data summarized in Table 11 suggests that (1) EVS values for the Subjects are on average 110 msec longer than those for the Objects and (2) the EVS values in the agent-cued condition are on average 16 msec less than those in the patient-cued condition. Univariate ANOVA performed with the Cue Location as the independent factor confirmed that the latter difference was not reliable. The lack of effect means that the participants did not experience any additional processing load associated with the production of the canonical SVO sentences when their attention was directed to the non-preferred starting point – the patient – although the EVS values for both sentence constituents were slightly larger in the patient-cued condition.

Another important result is that the cumulative eye-voice spans for Subject and Object differed significantly so that the eye-voice spans for the Subjects were reliably longer. The ANOVA with Constituent and Cue as independent factors supported this observation. It revealed a reliable effect of Constituent ($F1(1,23)=46.349$, $F2(1,62)=46.997$) and no interaction between the two factors. This pattern of results mirrors the increase in the EVS values at Object observed in Experiment 3 using Russian.

5.3.3.5. Name onset analysis

The missing values for the name onset analysis comprised in 2.3 % of data. Table 12 reports harmonic mean values for the name onset latencies in Experiment 4 at each sentence constituent.

The production of sentences in the patient-cued condition was delayed at each part of the sentence by a consistent lag averaging 137 msec. This difference was reliable for Subject (F1(1,23)=59.696; F2(1,63)=11.567), Verb (F1(1,23)=67.727; F2(1,63)=9.931), and Object (F1(1,23)=15.469; F2(1,63)=4.351). This pattern confirms that the preference for the canonical word order in languages like Russian and Finish may come in conflict with the biases promoted by the salience map of the described event. However, a tight coupling between what the speaker says and what he looks at results in a specific attentional switch from the salient but non-preferred starting point (the patient) to the non-salient but highly preferred starting point (the agent). This switch allows the processor to analyze the scene according to the top-down operational preferences promoted by the canonical grammar.

5.3.4. Results. Experiment 5

The same performance parameters were analyzed in Experiment 5. This time speakers of Finnish described the pictures of ditransitive interactions between the human characters. The cueing procedure was identical to the one used in the previous experiment. Three regions received an implicit visual cue – *the agent* (the initiator of the event), *the patient* (the recipient of the event), and *the theme* (the referent being transferred from the agent to the patient).

One important difference with the previous experiments is the inclusion of the third explicit referent – the theme. The pictures used in the current study were designed so that the theme was always "detached" from both the agent and the recipient (to avoid terminological ambiguity I continue calling it *the patient*) (see Figure 15). The latter feature is important because the area around each referent had to be coded as a distinct region for further eye-tracking analysis. However, separate area coding for the theme creates a potential problem. In Experiments 2, 3, and 4 the region carrying the event information included the area between the referents or the instrument. Experiments 2 and 3 demonstrated that speakers tended to look at the event region immediately before producing the verb of a transitive sentence although the instrument itself was not explicitly mentioned. However, in our ditransitive materials the same area is coded as *theme* (the direct object) of the event. One consequence is that the producers of ditransitive sentences will have to look at the theme area in order to extract information about both event and theme itself. This hypothesis is not new. For example, Government and Binding theory (Chomsky, 1981) suggests that the nodes assigned to the VP and the Direct Object, which is immediately adjacent to the VP, form more or less the same informational unit. If this interpretation is correct, one might expect longer eye-voice spans for the direct object (the theme) than the subject and the indirect object (the patient).

5.3.4.1. Cueing efficiency

The analysis of the initial fixations relative to the presentation of the visual cue confirmed that the cue was successful in attracting the participants' attentional focus in 93 % of agent, 86% of theme, and 75% of the patient-cued trials. A two-tailed t-tests revealed a reliable difference of these proportions from chance (.33) ($t1(23)=37.513$, $t2(95)=55.38$).

5.3.4.2. Word order analysis

The speech data were coded according to the syntactic structure of the produced sentences using the following categories: SVTO (Subject-Verb-Theme-Object), SVOT (Subject-Verb-Object-Theme), and Other. The missing data comprised .50% of the data. Table 13 demonstrates that although Finnish participants could potentially use other scrambling options to describe the experimental materials, they resorted only to two orders: SVTO and SVOT. The largely preferred word order was SVTO regardless of the location of the cue. The participants produced SVOT sentences in around 14 % of all descriptions, but the distribution of SVOT cases was equal across the experimental conditions. The ANOVA test using Cue Location as the independent and the probability to produce SVTO word order as the dependent variable returned did not reveal any reliable differences.

5.3.4.3. First fixation analysis

The missing data comprised 2.89% of the data. Table 14 summarizes the harmonic means for first fixation onsets for each of the interest areas as a factor of the location of the visual cue. The average onsets to fixate the cued referent were very similar across the three cueing conditions. The average estimation of the first fixation onsets was on a typical saccadic scale – around 150 msec. These results further support the saccadic nature of the first fixations driven by the implicit visual cue. The comparison between these reaction times did not return a significant result (all three *Fs*<1). It is more interesting to separately analyze the progression of first fixations from the cued to the other two referents across the cueing conditions.

The pattern of initial visual interrogation is surprisingly similar across the experimental conditions. Regardless of the location of the cue, the main purpose of the initial visual search seems to be to locate the agent as soon as possible in order to arrive at the preferred SVTO word order. In the agent-cued condition, once the agent was successfully located, the speakers moved their gazes to the theme area and then to the patient area. This pattern of fixation mirrors the predominantly produced word order – SVTO, which supports our previous findings for English and Russian (see Experiments 2 and 3). Statistical analysis supports such likelihood of fixations. A pair-wise comparison between the first fixations to the theme and to the patient in the agent-cued condition revealed that the participants were reliably more likely to fixate the theme than the patient after the agent ($F1(1,23)= 39.753$; $F2(1,31)= 109.004$). It is noteworthy that the delay between the fixations increases as the speakers proceed with the initial interrogation of the picture: the delay from the agent to the theme is 261 msec, while the delay from the theme to the patient is 423 msec. When the cue was on the theme of the event, the participants preferred to first move their eyes to the agent and then to the patient. After fixating the theme, participants proceeded with fixating the agent 236 msec. later; 71 msec. after that they fixated the patient. Although the latter delay reflects certain ambivalence, a pair-wise comparison between the first fixations to the agent and to the patient in the theme-cueing condition returned reliable result ($F1(1,31)=35.657$; $F2(1,31)=48.286$). Finally in the patient-cued condition participants preferred to fixate the theme and then the agent after their attention was attracted to the patient: they fixated the theme 245 msec. after the patient, and the agent – 47 msec after they fixated the theme. The latter delay was significant by the items ($F2(1,31)=4.946$) but not by the subjects ($F1(1,23)<1$). Given the general preference for agent-driven SVTO, one would expect the series of fixations from the patient to the agent to the theme in the patient-cued condition. However, on the saccadic trajectory from the patient to the agent the gaze would inevitably land at the intermediate area of the theme. This is exactly what happened, but typically for a very short time: the interrogation rapidly proceeded to the agent after a very short visit to the theme. Another important observation is that while in the agent-cued condition the delay between the fixations to the non-cued referents increases while the same delay decreases in the other two cueing conditions. At the same time the first delay is similar across the conditions (247 msec on average). The second delay, however, increases to 423 msec. in the agent-cued condition, but it decreases dramatically in the other two conditions. This pattern of results demonstrates an early semantic bias to the event analysis. Already at the very early stages of visual interrogation, speakers preferred to rapidly interpret events consistent with the preferred causality reflected in the canonical word order of their language.

5.3.4.4. Eye-voice span analysis

The current study used the same script for the automatic calculation of the EVS values as described in Experiment 4. Also, the same manual filtering was used: the EVS values of less than 100 msec were recalculated using the next available fixation onset; the missing EVS values (3.91%) were replaced with the mean value for the corresponding condition. The previous study revealed that Finnish speakers used canonical SVO in virtually 100 % of sentences when describing transitive events. When describing ditransitive events, they produced a considerably larger proportion of alternative structures with indirect object immediately following the verb (SVOT) (circa 14%). This increase did not differ between the experimental conditions, but the very presence of more alternative word orders becomes important when calculating EVS values. The script had to use a word order marker in order to properly associate the name onset signatures with their gaze counterparts. It became possible because the worksheet containing the name onset data for each trial also carried coding information about the word order choice made by the speaker. Therefore, the script first identified the word order associated with the trial, then matched the constituents to the corresponding visual interest areas, finally looked for the fixation onsets relevant for the calculation of the EVSs.

One specific parameter for calculating eye-voice spans in the situation of ditransitive sentence production is the fact that the interest area related to both the event and the theme is the single region – the one around the event's THEME. Due to this specific feature, the eye-voice spans for both sentential Verb and Theme had to be calculated using the same visual region. It is possible that a speaker has to interrogate the same part of a visually presented event in order to extract different types of discourse-relevant information. In fact, speakers quite often refixate referents areas in order to collect additional information or correct the information that has already been processed. Also, our previous studies confirmed the importance of the region that contains the event's instrument (transferable theme in this case) for the extraction of the visual information necessary to arrive at the sentential verb.

Table 15 presents the mean EVS values across the experimental conditions. Table 5 does not separate the EVS values according to the produced word order – SVTO or SVOT. If the EVS measurement in some way represents the lexical access processes associated with filling the slots of the formulated sentence (e.g. Griffin & Bock, 2000), the positioning of the referent in a sentence should have no effect on the magnitude of the EVS. However, what exactly eye-voice span measurement is good for is one of the questions I attempt to answer in this thesis.

Informal analysis of the data illustrated in Table 15 leads to the following conclusions: (1) the mean magnitude of the EVS values at Subject and Verb were quite comparable, (2) there is an observable slow-down in processing of the sentential Theme, (3) processing of the Object seems to be the easiest among the four referents, (4) the difference between the mean EVS values for the agent- and the theme-cued conditions seems quite negligible (20 msec); however, the 43 msec longer EVS mean value for the patient-cued condition may denote a certain processing difficulty associated with the formulation of the sentential theme. A series of ANOVAs performed on the EVS data revealed a main effect of Cue for Subject ($F1(2,46)=3.837$), Verb ($F1(2,46)=4.722$; $F2(2,93)=9.806$), and Theme ($F1(2,46)=3.539$; $F2(2,93)=2.884$, $p=.061$), but not for Object. Pair-wise comparisons confirmed that EVS values in the theme-cueing condition were reliably longer than in the other two conditions for Subject and Verb. At Theme, the general effect of Cue was carried by the difference between EVS values in the patient-cued and the other two conditions. In general, participants demonstrated comparable EVS performance at Subject and Verb, while EVS values for Theme were reliably longer than for the other three constituents. EVS values for Object were observably shorter than for the other constituents but reliably different only from Verb. In any case, the most outstanding observation is that processing Theme presented participants with certain problems. It is, however, unclear whether this effect means that there is something special about processing the Theme as a referent or whether it has to do with filling in the position immediately following the Verb as a vast majority of the produced structures were in fact SVTO.

A separate analysis using constituents as a within-item and within-subject factor revealed a reliable effect of the constituent ($F1(2,46)=29.928$; $F2(1,93)=48.831$). This effect was largely due to the outstandingly large EVS values for Theme and the relatively small values for the Object as compared to the other constituents. Based on the results, it is possible to conclude that during the formulation of the Finnish ditransitive sentences (1) processing of the first two sentence constituents – Subject and Verb was associated with a similar amount of cognitive load while (2) processing the theme was associated with arguably the “most difficult” and (3) finalizing the sentence with the Object was the easiest part of the task.

In order to get a better understanding of the EVS dynamics, I separated the EVS data for the Theme and the Object according to the produced word orders (see Table 16). Although proper statistical analysis of the data represented in Table 16 was difficult due to unequal numbers of SVTO and SVOT observations, the overall picture prompts further conclusions about the processing loads associated with the construction of ditransitive argument structure as revealed by EVS measurements. First important observation is that the increase in the magnitude of the EVS associated with the sentential Theme persists in both word orders. Hence, there is something special in processing the Theme of a ditransitive sentence in Finish regardless of its position in a sentence. Second, a decrease in EVS observed for the Object was confirmed only for SVTO structure – the structure, in which the object comes last. When the Object was inserted before the Theme, like in SVOT structure, an observable increase in EVS occurred instead. This pattern supports the positional explanation of the effect reported above. It is, therefore, possible that the participants experienced additional processing load at two stages of structure building: when filling the slot immediately following the assignment of the main constituents (the head NP and the Verb), and when processing the Theme of the sentence.

5.3.4.5. Name onset analysis

The logic of measuring the name onset latencies was identical to the one used in the previous experiments. Because there was no significant difference in the use of the non-canonical SVOT order as a factor of the location of the cue and because this study is not primarily interested in the relative difficulty of producing SVOT in Finnish as compared to SVTO, all recorded narratives received the following coding – Subject-Verb-Object 1-Object 2 regardless of whether the direct object came first. This aggregation is further justified by the fact that word order scrambling in Finnish does not require the use of prepositions; therefore the general structure of the sentence remains the same, the only thing that changes is the position of the constituents. Finally, there are theoretical reasons to treat these two word orders in Finnish as functionally similar (Kaiser, 2002). The responses with Subject onset latencies less than 300 msec. and more than 4000 msec. were excluded from the analysis as well as the sentences that did not include all three of the sentence constituents. This resulted in 1.1 % of data loss. The missing values were replaced with the mean value for the corresponding condition. Table 17 reports harmonic mean values for the name onset latencies in Experiment 5 at each sentence constituent. The first important observation resulting from the pattern of the results is that the cueing manipulations did not lead to any reliable reaction time penalty although, similarly to the Experiments 1 and 4, participants tended to heavily rely upon agent-driven syntactic structures regardless of the location of the cue. This conclusion was supported by the lack of any reliable effect of cue on name onset latencies at either of the sentence constituents with the exception of Object 1 ($F1(2,46)=4.203$). However, the direction of the local effects was quite different from the previous studies. Pair-wise comparisons confirmed that the effect at Object 1 was due to the 73 msec. longer latencies in the theme-cueing than in the patient-cued condition ($F(1,23)=8.759$). Interestingly, a test of within-subject contrasts revealed a reliable quadratic trend at Subject ($F(1,23)=5.616$), Object 1($F(1,23)=6.205$), and Object 2 ($F(1,23)=4.490$). This pattern suggests that while the production rates did not differ between the agent- and the patient-cued conditions, the production proceeded relatively slower when the cued referent was the theme. Although this slow-down was not substantial (40 msec on average), it was persistent throughout the course of sentence production and already observable at the Subject.

5.3.5. Discussion

The current set of experiments used a variant of the implicit perceptual priming paradigm in order to test whether the location of the speaker's attentional focus predicts consistent alternation of the assignment of the sentential starting point and, as a result, biases subsequent choice of the word order during Finnish sentence production. Finnish participants produced transitive and ditransitive sentences while viewing line drawings of the corresponding events.

The implicit visual cue used in the current experiments successfully attracted participants' attention to the cued referents; however, this manipulation did not result in any reliable word order alternation during the production of either transitive or ditransitive sentences. Results of Experiment 4 did not reveal any alternation of word order during Finnish transitive sentence production, and results of Experiment 5 demonstrated that although speakers used both SVTO and SVOT orders, their distribution did not rely on the Cue Location factor. On one hand, the distribution of ditransitive orders in Experiment 5 observed merely supports the dominant status of the DO-first order – SVTO – in Finnish. On the other, the current result contrasts with similar studies using English (e.g. Gleitman, et al., in press), who demonstrated how virtually the same cue can be efficient in driving the choice of the sentential starting point and the subsequent syntactic structure in English. Together, these results confirm that a successful mapping of the visually focused referent onto the sentential starting point can sometimes be problematic. Table 18 provides simple comparisons of the proportion of agent-driven word orders as a factor of cue location in two pairs of studies that used virtually the same cueing protocols but employed languages that differed on case-marking and word order parameters.

In the presence of an explicit and persistent visual cue combined with a restricting instruction to preferentially attend to the cued referent, speakers of English mapped subject's position onto the cued referent in virtually 100% of the cases (Tomlin, 1995). The very same cueing procedure led to a much weaker result in Russian: The participants continued producing the canonical SVO in over 60% of the patient-cued trials (Experiment 1, this thesis). Implicit cueing of attention in Gleitman, et al., (in press) led a much weaker yet reliable alternation pattern than the one observed by Tomlin. The same cueing paradigm used in the current set of studies with Finish did not result in *any* word order alternation. Results for grammatical voice alternation in English resemble the pattern observed in Russian using a much stronger Fish Film cueing protocol. Together these comparisons support the idea that active case marking and wider word order optionality come at a larger processing cost when the speakers of the languages with such properties have to map the attentionally focused referent onto the initial position in a sentence.

The first fixation analysis in Experiment 4 revealed very similar onsets to fixate the cued referent between the cueing conditions. On the other hand onset of the first fixations to the non-cued referents varied as a factor of the Cue Location: the first fixations to the patient in the agent-cued condition occurred 130 msec. later than the first fixations to the agent in the patient-cued condition This pattern demonstrates that Finnish speakers were more likely to quickly switch attention from the cued patient to the non-cued agent in the patient-cued condition in order to quickly extract the gist of the described event in accordance with the agent-driven canonical causality. Gaze shifting in this situation possibly reflects an active response to a conflict between the two semi-automatic tasks: to attend to the visually-cued stimulus and to select a referent for further lexical encoding (cf. Roelofs, 2007). When the cued referent was the agent, this early switching was not necessary because the cueing scheme accorded with the conceptualization of the subsequently presented picture in accordance with the preferred causality. Given the fact that first fixations for the most part represent early perceptually-motivated event analysis, these results suggest that perceptual processing of the event does not have to be completed before operations at subsequent processing stages can be activated if the cue supports the preferred interpretation of the event. If, on the other hand, the cue contradicts this interpretation, a complete perceptual analysis becomes necessary before the processor can be feed the information forward for deeper conceptual processing.

The analysis of the initial fixations to the event area revealed very few looks in the agent-cued condition. When the patient was cued, participants fixated the event area substantially more often. This suggests that speakers of Finnish tended to use the information provided by the verb preview when the agent of the event was cued while the event reanalysis was necessary when the patient was cued, which further supports an early bias toward the event causality that supports the canonical SVO structure. A similar picture of the initial fixations to the cued referent was observed in Experiment 5: The average onsets to fixate the cued referent were very similar across the three cueing conditions. The pattern of the initial fixations following the fixations to the cued referent also revealed a picture similar to the one observed in Experiment 4: Regardless of the location of the cue, the main purpose of the initial visual search was to locate the agent as soon as possible in order to arrive at the preferred event causality.

Eye-voice span analysis provides an independent signature of the processing load or cognitive effort associated with the selection of a referent for a particular sentential role. The analysis of EVS data in Experiment 4 prompted three important inferences. (1) The EVS magnitude did not vary as a factor of the cueing scheme, which suggests that a tacit cueing manipulation used in the experiment did not lead to any additional processing load in "miscued" trials. (2) The EVS values for the Object were much smaller than the ones for the Subject, which suggests a higher processing load, associated with the selection of the initial sentence constituents. This confirms my interpretation of the importance of eye-voice spans as indicators of the difficulty associated with filling-in of the lots in the argument structure. If structural preplanning occurs before the event can to be described, it should be more difficult to identify and assign the initial nodes of the structure whereas the assignment of the final constituents should be more automatic. (3) The EVS values observed in the current study were much shorter than the ones reported in other visually-mediated production studies (e.g. Griffin & Bock, 2000, Experiments 2 and 3, this thesis). It is very likely that the average magnitude of the EVS values depends on the amount of information about the event available to the speaker prior to the initial analysis. If this is the case, an interpretation of EVS measurement as a signature of lexical access during incremental formulation of sentences can not be completely correct. It is more likely that EVS reflects more general processes associated with the selection and the construction of the argument structure selection *and* probably accessing the lexical forms during the instantiation of the chosen structure in a spoken sentence.

The analysis of the EVS data observed in Experiment 5 revealed that (1) Finnish speakers experienced the least cognitive load when their attention was directed to the Agent. This is by no means a surprising result because the cue to the agent supports the canonical S-V causality preferred by the participants. (2) Participants experienced comparable cognitive load when relating referents to Subjects and Verbs. Both structures used by the participants in the current experiment had agent-driven S-V as their starting components, so it is not surprising that it was equally difficult to build the structure up to the point where the argument structure construction diverge dependent on whether the next NP contains the Theme or the Object. (3) The magnitude of the EVS values was observably larger at the Theme and smaller at the Object. A separate analysis of EVS patterns for SVTO and SVOT sentences revealed that the increase in the magnitude of the EVS associated with the Theme persisted in both word orders. Hence, there is something special in processing the Theme in a Finish sentence regardless of its sequential position. A decrease in EVS observed at the Object was confirmed only for SVTO structure – the structure, in which the object comes last. When the Object was inserted before the Theme, (SVOT), an observable increase in EVS occurred instead. This pattern supports the positional explanation of the effect reported above. It is, therefore, possible that the participants experienced additional processing load twice during structure building: When filling the slot immediately following the assignment of the notional verb and when processing the Theme of the sentence.

The pattern of the name onset latencies during the production of Finnish ditransitive sentences was quite different from the one observed in transitive sentence production. In Experiment 4, Finnish speakers demonstrated a specific reaction time penalty associated with the necessity to switch attention from the cued patient to the non-cued agent in order to arrive at the preferred SVO word order. A chronometric trade quite similar to the one observed in Experiment 1 with Russian demonstrates once again the presence of a specific attentional shift from the cued but non-preferred starting point to the preferred starting point in order to arrive at a well-automated canonical SVO. This attentional shift helps maintain correspondence between the linearity of the chosen word order and the sequence of fixations immediately preceding the production of the corresponding constituents.

However there was no such delay during ditransitive sentence production, although the same manipulation of the attentional focus was used. In other words, although attention was successfully captured in all the three experimental conditions and although participants preferred the agent-driven canonical SVTO, the name onset data did not reflect the expected switch of attention in the patient- and the theme-cueing conditions. One important difference between the previous and the current experiments that might help in interpreting the data is the change in the referential complexity of the described events. Experiment 4 used transitive events with only two referents interacting. Experiment 5 made the referential scheme more complex adding the third referent. If the perceptual effects on the choice of word order are relatively weak and transient it is not surprising that the increase in the semantic complexity of the event resulted in even weaker representation of the salience parameters across the whole spectrum of data. Apparently, perceptual biases become available to the speaker very early. If the cue is presented together with the presentation of the event, they probably become one of the driving forces of the conceptual analysis. If the cue is presented before the event information is available, the influence of the former may be even weaker. In any case, unless the bias suggested by the perceptual asymmetries is confirmed during further stages of sentence formulation, it becomes replaced by other biases that originate from semantic and syntactic processing. It is quite obvious that in the case of a "miscued" (the trials whose cueing scheme did not support the inclusion of the agent as the sentential starting point) transitive sentence production, the event reanalysis toward the preferred agent-driven causality is minimal: The speaker's attention is rapidly switched to the agent, and then the sentence production begins especially if the event information is available from the verb preview as in Experiment 4. This explains the presence of the observable reaction time penalty associated with switching attention from the patient to the agent. As the referential complexity of the event increases, so do the time and effort necessary for the event reanalysis. With more processing during rapid apprehension stage, the perceptual effects become completely "washed out" already at very early steps of the event (re)analysis.

Experiments 4 and 5 together provide a fine-grained dynamic picture of how the processor operates during the production of Finnish transitive and ditransitive sentences. The results support the idea that speakers of free-word-order case marking languages experience additional processing load when they need to rapidly map the perceptual asymmetries embedded in the scene onto the available word order structures. Because of this additional load, they tend to rely more upon the canonical word order automated in the language use. Both experiments revealed speakers' tendency to seek agent-driven event interpretation as soon as possible in order to minimize the time and effort necessary to arrive at the canonical syntactic structures. A relatively new view of the eye-voice span measurement suggests that the latter reflect relative processing load associated with relating the event's referents to the sentence constituent slots. Together the latter results provide strong evidence for early argument structure planning the speakers undertake well before the sentence formulation process commences.

5.4. Chapter Conclusions

Experiments 1-5 revealed some properties of the perception-grammar interface established online during the production of sentences. The four most important conclusions from these experiments are: (1) Conceptual analysis of the visually presented events commences very early: Speakers' order of visual interrogation of the relevant entities is already biased at the first-pass scanning of the event; (2) Increasing the referent' salience may result in higher prominence of this referent among its counterparts and bias the production system toward using it as the starting point of a spoken sentence; (3) Only explicit attentional cueing leads to a similar result in languages with free combinatorial properties and rich case grammars; and (4) The effects of perceptual salience on the choice of word order are indirect and effective only at the rapid apprehension stage.

This last conclusion is justified by the fact that perceptual biases were more likely to be observed when the described event and the resulting syntactic structure were relatively simple. When perceptual biases were not strong enough or did not receive further support at the later stages of sentence formulation, they were replaced by the preferences accommodated within the levels following event conceptualization. Logically, the propagation of perceptual effects to the stage of word order selection may be more likely when perceptual cues are combined with conceptual, lexical-semantic, and/or grammatical cues. Availability of conceptual or lexical-semantic information prior to the event's presentation should help the attentionally focused patient to carry on its competitive status through conceptual analysis and lemma selection. Previously encountered syntax might also help promote the grammatical role assignment to the already attentionally-cued and semantically supported referent during formulating of a novel sentence.

The background research for the conceptually motivated selection of word order was outlined in Chapters 1 and 2. Previous studies using referential priming demonstrated that the primed referent receives the given entity status, is processed early during lemma selection, and, consequently, tends to be inserted early into the sentence frame. Although this effect can be viewed as functionally similar to the effect of attentional focus on the starting point assignment, the driving mechanisms behind these two effects cannot be exactly the same. One reason is that while both types of priming contain an attentional focusing component, only referential priming provides information about the referent's identity that can result in the establishment of the given/new contrast further biasing conceptual and linguistic analysis of the event[5]. Because of this additional power, the effect of referential priming on the choice of word order should be greater than the isolated effect of perceptual priming. The impact of the cue on choice of the word order should increase as a factor of the amount of conceptual information the cue provides. The set of experiments discussed in the next chapter provides such comparison.

---

5 If availability of this information leads to active lexicalization of the primed referent then the combinatory power of the cue becomes yet stronger; however, the two Experiments discussed in this chapter do not provide any evidence for the existence of lexicalization component in referential priming.

Chapter 6. Perceptual and conceptual priming in English transitive sentence production

In many respects, Experiments 6 and 7 are methodologically different from existing perceptual and referential priming paradigms. For example, the methodological justification for the experiments reported in the preceding part of the thesis was that (1) earlier visually-cued production research used English and (2) very few studies used eye-voice span as an indicator of the processing dynamics in sentence production. Otherwise, Experiments 1-5 use a well-established paradigm. Studies in Chapter 6 have novel methodological features that allow us to understand how perceptual constraints improve the likelihood of the focused element being promoted to the starting point of a sentence and (2) how access to *conceptual* information (manipulated through referential priming) interacts with the attentional focus during sentence production.

6.1. Experiments 6 and 7. Perceptual and conceptual priming with relative SOA and constraining power in English transitive sentence production

6.1.1. Experimental design and hypotheses

Three independent factors were manipulated in both experiments: (1) the location of the cue, (2) the duration of the cue, (3) the experimental instruction. The difference between the two studies was in the type of the priming paradigm used. Like Experiments 4 and 5, Experiment 6 used a variant of the perceptual priming paradigm. A red dot presented to participants in one of the display’s regions acted as the attentional cue. In Experiment 7, a variant of the referential priming paradigm was used. Cueing was achieved via preview of one of the event referents – agent or patient.

Experimental materials in both studies were the pictures of transitive events used in previous experiments (see example in Figure 1). Cueing was done prior to the target picture onset in both studies. In half of the materials the cue was on the event's agent; in the other half – on the patient. This constituted the first independent factor – *the Cue Location*. One prediction is that referential priming should produce a greater bias toward use of the cued referent as the sentential starting point. This should be true because, in addition to spatial indexing, a referential cue reveals important conceptual information about the cued referent.

The Cue SOA factor was manipulated as the stimulus onset asynchrony (SOA) – the time the cue remained on the screen before the target event was displayed. This resulted in either *implicit* or *explicit* cueing. The explicit cue was displayed for 700 msec.; the implicit cue - for 70 msec. Speakers should be more sensitive to the cueing manipulation when the cue is (1) explicit as explicit cueing provides firmer indexing of the cued location of the referent, and (2) referential because referential priming reveals the referent's identity – an important part of conceptual information about the referent. On the other hand, the advantage of referential priming may disappear in the implicit cue condition because speakers may not have enough time to extract the information provided by the referent preview and treat it just as a visual cue. However, some recent studies (Dobel & Gumnior, 2004; Glanemann, et al., in press) demonstrated that speakers can apprehend both events and entities with very short presentation SOAs – around 200 msec for the events and around 100 msec for the entities.

There were two different *Experimental Instructions* about how to treat the cue. One instruction – *non-constraining* – was to direct gaze to the cue and then describe the subsequent picture while freely scanning it. The *constraining* instruction was to direct gaze to the cue and keep looking *only* at the cued referent while describing the target event. The instruction factor was manipulated across 4 blocks of trials. One half of the participants received the non-constraining instruction first; the other half – the constraining instruction first. Of course, the most obvious prediction is that speakers' performance in the constraining condition should be hindered by the necessity to interrogate the elements outside of their visual focus covertly – without looking at them. One consequence of this restriction might be speakers' higher dependency to alternate word order as a factor of the cue location (cf. Tomlin, 1995). A more interesting question, however, is whether processing difficulties will be observed only for the referents outside of the visual focus? If complete event analysis is a prerequisite for incremental sentence formulation, eye-voice spans should reveal processing difficulties due to the instruction effect already at Subject; if, however, sentence production can commence without such complete analysis, the instruction effect should be observed at the Object but not at the Subject. Finally, a similar question can be asked about the latencies to produce sentence constituents. For example, if production does not depend on full event analysis, there should be no onset delays for the names corresponding to the cued referents.

Hence, in both studies I administered a 2x2x2 design with the following independent factors: Cue Location (agent/patient), Cue SOA (70 msec/700 msec), and Experimental Instruction (non-constraining/constraining). The dependent variables in both studies were (1) the probability to produce Active Voice sentences, (2) the name onset latencies for each sentence constituent, (3) the onset of the first fixations to the referent regions, (4) the eye-voice spans for each sentence constituent.

6.1.2. Materials

For the purposes of the current studies, I continued using the same set of simple black-and-white line drawings of transitive events (see example in Figure 1). The materials were counterbalanced for left-right agent presentation, size, animacy, color, and referent role suggestibility. The materials were not controlled for corpus frequency; therefore participants in both studies had an opportunity to preview the single pictures of each referent during the practice session and get familiarized with the referents they encountered in the experimental session. There were total of 17 referents used in the transitive interaction portrayed in the target trials. The experimental materials consisted of a set of 64 pictures: 8 pictures x 8 events (CHASE, KICK, PULL, PUNCH, PUSH, SCOLD, SHOOT, and TOUCH). 130 filler pictures were used in the experiments. The filler materials were pictures of locative events each containing 2 geometrical shapes presented simultaneously in different regions of the screen (see Figure 16). There were four possible orientations of the geometrical figures in the filler materials: (1) vertical, (2) horizontal, (3) diagonal ascending (from left to right), and (4) diagonal descending (from left to right). The instruction to treat the fillers was to describe the fillers in any possible way. For example, a typical way to describe the event in Figure 16 was to say: *"A heart is to the right and above a circle"*. A minimum of 2 fillers separated the target trials. There were 4 filler trials at the beginning of each session and no filler trials after the last target trial. Therefore, each experimental session consisted of total of 194 trials separated in 4 blocks, 16 target trials in each block.

I pre-coded one interest area for each referent portrayed in the target pictures. These interest areas included the referent itself and the surrounding area of approximately 2^0 of visual angle. The interest area for the *action* encompassed an instrument of an action or the area between the two protagonists.

6.1.3. Participants

A separate group of 24 native speakers of English with normal or corrected-to-normal vision and no language-related conditions took part in each study. All the participants were undergraduate students at the Department of Psychology of the University of Glasgow. The mean age of the participants was 19.5 years in Experiment 6 and 20,3 in Experiment 7.

6.1.4. Apparatus

Same as before.

6.1.5. Experimental procedure

At the beginning of each experimental session, the participant was positioned in front of the monitor approximately 60 centimetres from the screen. They had a direct view of the monitor throughout the session. Then a practice session followed, during which the participants performed two tasks: (1) they viewed and named the pictures of the single referent characters and (2) practiced describing sample event pictures. In single referent naming task, the characters' names appeared on the screen together with the pictures, and the participants were asked to say those names out loud, associate them with the picture of each referent, and remember the names and the referents' appearances for the further tasks. The event description part of the practice session offered the participants a chance to practice describing 4 sample event pictures interspersed with 4 sample filler pictures. Each sample picture appeared on the screen without any cue with a sentence that could be potentially used to describe it: An active or passive voice sentence accompanied each target event picture, and one possible locative sentence accompanied each filler picture. Participants were not aware of the true nature (target or filler) of the practiced materials. They viewed the pictures and read those sentences out loud. The experimenter then explained that either of the suggested structures could be used to describe the events during the experimental session. Practicing both active and passive voice sample sentences during the practice session ensured that a typically infrequent structure (passive voice) was pre-activated alongside more frequent canonical active voice frame. In other words, if passivization typically observed in visually cued production studies really depends on the location of the cue, the practice session gave the passive voice all the chances for fair competition with its more frequent syntactic alternative – the active voice.

Typical experimental trial sequences for Experiments 6 and 7 are portrayed in Figures 17 and 18. Each trial started with the presentation of a fixation mark in the centre of the screen. After the participant successfully fixated the fixation mark, the verb related to the event in the target display appeared on the screen. The presentation of the second fixation mark followed, now in the lower area of the screen to insure that the saccade to one of the interest areas once the target picture was displayed. Presentation of the target display was gaze-contingent with presentation of the previous fixation mark. The fixation mark that preceded the target picture was equally distant from all the three interest areas. The participants had to look at the fixation mark for a minimum of 200 msec. in order for the target picture to be displayed. The target picture then appeared in the centre of the screen. The participants were instructed to use a single sentence to describe the target picture. They were also instructed to begin describing the picture extemporaneously right after the picture was displayed. Once they finished describing the picture, the participants pushed the spacebar key, and the next trial would begin. If the participants failed to produce a sentence, the target picture was removed from the screen after 7700 msec. An audio marker accompanied presentation of each target picture; therefore the participants' speech was recorded time-locked to their eye-tracking data. The audio marker appeared as a clear peak with the duration of approximately 20 msec. The offset of this component marked the starting point for the reaction time analysis. The onset of the corresponding narrative was used as the finishing point of the analyzed latencies.

6.1.6. Results. Experiment 6

Because the statistical analysis in Experiments 6-9 is more complex and involves a set of interactions, the F values are presented in ANOVA tables instead of being reported in the text. Separate analyses for the possible effects of the Block Order and the position of the agent (left/right) revealed no effect of either factor on any of the dependent variables; hence these variables were removed from the reported analysis.

6.1.6.1. Cueing efficiency

The cueing efficiency analysis confirmed that both implicit and explicit visual cue successfully attracted the attentional focus. Table 19 summarizes the cueing efficiency measures for Experiment 6. Two-tailed t-tests performed on the mean proportions as compared to chance (.5) returned reliable cueing results for all 8 experimental conditions.

6.1.6.2. Word order analysis

I transcribed the speech data according to the syntactic structure of the collected sentences using the following categories: Active Voice, Passive Voice, and other. Table 20 summarizes the observed probabilities to produce Active Voice descriptions across the experimental conditions. The results of the ANOVA tests conducted on the observed probabilities are reported in Table 21. This analysis suggests that the cueing manipulation was efficient in driving the starting point assignment and the alternation of the resulting syntactic structure: The participants were 39 % more likely to produce passive voice sentences when their attention was directed to the patient. This effect was slightly amplified when the cue was explicit, which was revealed by both a significant effect of the cue SOA (4%) and the interaction between the two factors (see Figure 19). Pair-wise comparison confirmed that the significant interaction between the cue location and the cue SOA factors was due to the difference in the patient-cued condition ($F1(1,23)=5.612$; $F2(1,31)=12.387$). Also, English speakers appeared to be sensitive to the instruction manipulation: They were 15% more likely to alternate the produced word order when the instruction constrained their visual focus to the cued referent. The instruction factor interacted reliably with the cue location factor largely due to the difference in the patient-cued condition (F1(1,23)=75.836; F2(1,31)=119.333) (see Figure 20). It is also notable that, regardless of the observed effects, English speakers tended to use the canonical active voice syntactic frame even in the conditions that strongly promoted the patient as the sentential starting point producing up to 36% of cue-conflicting word order in the patient-cued/explicit/ constraining condition. In non-constraining condition there was only 20 % lower probability to use active voice order regardless of the cue location. This pattern of results further supports the importance of the global constraints imposed by the automated grammatical defaults and a relative weakness of the ability of the perceptual factors to influence the choice of the word order during sentence production.

One possible consequence of the instruction manipulation was the inability of the speakers to identify the referent outside their visual focus. This assumption was not confirmed. The analysis of the participants' ability to identify the referents out of their visual focus in the constraining instruction condition revealed that such identification was successful in 83% of constraining condition trials. Therefore, the effect of the visual constraint on the selection of the sentential starting point has to be attributed to speakers' inability to successfully assign grammatical functions to the referents they could not interrogate visually.

6.1.6.3. First fixation analysis

The missing values accounted for 2.3% of the data. I separately analyzed cases when participants had not complied with the constraining instruction and looked at the non-cued referent during the trial. Such cases accounted for 21 % of the total data. One sample t-test confirmed that the compliance with the instruction was better than chance (t1(23) = 7.289, p<.000, t2(31) = 12.36, p<.000). When the "unauthorized" fixations occurred, they were executed late – well into the sentence formulation process – averaging 2882 msec after target picture onset. Table 22 summarizes the harmonic mean reaction times for first fixation onsets to each interest area across the experimental conditions. Table 23 presents the results of ANOVA performed on the latencies to fixate the *cued* referent. Because of the nature of the experimental instruction, the first fixation latencies to the non-cued referent were only analyzed for the non-constraining condition.

The analysis confirmed that there were no reliable effects of the cue location and the cue SOA on the first fixation latencies. However, participants were reliably slower to execute the initial saccade to the cued referent when they were under the constraining condition. It is quite likely that such a "slow-down" is an artefact of the preparatory processes associated with performance on quite a difficult task: To describe a dynamic event without looking at one of its referents. On the other hand, the onset of the initial fixation within the cued referent region in a condition when a participant's glance is bound to the cued location with instruction not to look at the other referent does not exactly reflect a saccade from another part of the display: By the time the target picture is displayed, the speaker's glance is already in the location, from which it cannot move due to the nature of the instruction. This property might also add to longer first fixation onsets in the constraining condition.

Results of the analysis of the first fixation onsets to the *non-cued* referent in the non-constraining condition revealed a picture similar to the one observed in Experiments 4 and 5 (see Table 24 and Figure 21). The interaction between the cue location and the cue SOA factors was carried by both differences in the implicit (F1(1,23)=28.303;

F2(1,31)=30.216) and the explicit cue conditions (F1(1,23)=17.846; F2(1,31)=10.349). Participants tended to fixate the non-cued patients about 300 msec later in the agent-cued condition, than the non-cued agent in the patient-cued condition. This difference is probably due to the fact that attracting attention to the agent resulted in speakers' assuming the event's canonical causality without further examination of the event, which was overwhelmingly confirmed by the choice of the active voice sentence frame in this condition. Yet an even larger difference was observed between the average values for the two levels of the Cue SOA factor: Speakers tended to switch their attention to the non-cued referent on average 600 msec earlier in the implicit cue condition. It is quite likely that explicit attraction of attention to the location resulted in deeper processing of the latter and a higher chance of early biasing toward the canonical causality. The explicit cue and/or the cue on the agent were sufficient for the earliest commitment to agent-event-patient causality without an interrogation of the rest of the scene. If the cue was implicit and/or on the patient, an event (re)analysis was more likely; as a consequence speakers fixated the non-cued referent much earlier in these conditions. These results reveal speakers' sensitivity to the cue duration and the existence of very early biases to interpret the visually presented events according to the canonical causality.

6.1.6.4. Eye-voice span analysis

I need to make two points of explanation before I discuss the eye-voice data for Experiments 6 and 7. First, because of the specific nature of the constraining instruction there were virtually no looks to the event interest area (less than 1%) in this condition. Also because of verb preview, participants rarely fixated the event area, especially in the agent-cued condition (7% of the time). In the patient-cued condition such gazes were more frequent – around 21% of the total trials. This result corroborates findings from Experiment 4, in which I found a higher probability of event reanalysis when the patient of the event was visually cued.

Second, in the constraining instruction condition participants were not allowed to look at the non-cued referent, and they coped with this restriction quite well. However, because of this feature, the EVS values for one of the referents were regularly unavailable. Hence, I report and discuss the EVS mean values for the constraining condition only informally (red values in Table 25). To further substantiate quasi-experimental comparisons, I report a separate matrix for the EVS values in the trials with Passive Voice descriptions.

In the non-constraining condition, the procedure for calculating EVS values remained the same as in Experiment 5. The values less than 100 msec were replaced with the ones calculated manually using the fixation immediately preceding the one detected by the script. The missing values accounted for 4.3% of the total number of trials. Such values were replaced with the mean EVS value for the corresponding condition. Tables 25 and 26 summarize mean eye-voice span values across the sentence constituents in both active and passive voice sentences as a factor of experimental conditions.

Informal examination of the data in Tables 25 and 26 led to the following observations: (1) The mean EVS values for both sentence constituents are much greater in the constraining condition, more so for the Objects, (2) EVS values for the Object in the non-constraining condition were on average longer than those for the Subject (57 msec), (3) Processing of the both sentence constituents was easier in the agent-cued condition as revealed by the EVS data (Subject: 62 msec difference; Object: 26 msec difference), and (4) Cue SOA effect was only noticeable for the Subjects: There was a 47 msec facilitation effect observed in the explicit cue condition.

Informal analysis of the EVS data for the passive voice sentences in the non-constraining condition reveals that: (1) Average EVS values for The Subjects were about 20 msec smaller than those for Objects, (2) the overall mean EVS value for The Subjects in the passive voice sentences is about 60 msec greater than in the active voice sentences, (3) the same is true for the Objects (47 msec), and (4) Processing of the referents in the passive voice sentences was easier in the agent-cued condition – (396 vs 486 msec average) and in the implicit cue condition (417 vs. 464 msec average). Although most of the reported differences for the passive voice sentences are minimal, and a formal analysis of the EVS data in the passive voice sentences is not possible, informal conducted examinations confirm that the processing load associated with the formulation of the passive voice sentences is at least larger than during the active voice sentence formulation. Also, the previously reported facilitation effects of cue location and SOA were also confirmed by the analysis of the EVS signatures in the passive voice sentences.

Statistical analyses of the Subject-related EVS data in the non-constraining condition are summarized in Table 27. A separate analysis of the EVS data for the Objects did not return any reliable results. A series of ANOVAs confirmed the effect of the cue location on the relative processing difficulties associated with the formulation of the sentential Subjects: Speakers experienced reliably fewer problems during formulation of the Subjects when their attention was directed to the event's agent. A comparison between the EVS values for the Subjects and the Objects returned a non-significant result.

6.1.6.5. Name onset analysis

The missing values accounted for 3.7 % of data loss. Table 28 reports harmonic mean values for the name onset latencies in Experiment 6 at each sentence constituent.

An informal examination of the results summarized in Table 28 reveals that speakers were faster to initiate sentences when their attention was attracted to the agent, when the cue was explicit, and when the instruction was non-constraining. Table 29 shows the name onset latencies in the passive voice descriptions. The mean values reported for both explicit- and implicit-cue constraining/AC condition are not particularly informative as these means are based on a very small number of observations. The distribution of the latencies in the rest of the table partially supports our preliminary analysis: The initiation of passive voice sentences was the slowest in the constraining/patient-cued/implicit condition. However, the fastest production rate for the passive voice sentences was observed in the non-constraining/agent-cued/implicit condition. Although there were not many cases of passive voice sentences in this condition, their faster initiation may be a result of a rapid event reanalysis toward the passive causality when processing is not hindered either by a strong indexing of the cued location or the limited ability for visual interrogation. It is also important to note that even at its fastest production of passive voice sentences was delayed by almost 100 msec as compared to the fastest active voice sentences. These observations suggest that the production of passive voice sentences was associated with a more thorough conceptual analysis and a larger degree of preplanning. Active voice sentence production, on the other hand, can begin as soon as the starting point is assigned and follow a more incremental production protocol.

Statistical analyses performed on the name onset latencies largely supported these observations (see Tables 30-32). At all three constituents, the location of the cue produced a reliable facilitation effect with the agent-cued sentences being initiated on average 240 msec. faster than the patient-cued ones. This effect was almost equal across the constituents. Varying the cue SOA led to a similar result: The descriptions were started on average 110 msec. earlier if the cue was explicit than when the cue was implicit. However, this difference decreased linearly from 148 msec. at Subject to 65 msec. at Object. This result suggests that the stronger indexing advantage associated with longer cue SOA is important for the selection of the starting point referent; that is why the effect slowly vanishes during the rest of sentence production. Finally, the effect of the experimental instruction led to a faster constituent production when the instruction was non-constraining. This effect also changed linearly over the course of sentence production from the average of 73 msec. at the starting point (Subject) to the 369 msec. effect at the end of a sentence (Object). This pattern suggests that the uptake of the information about the elements outside of the visual focus was more difficult when the freedom of visual interrogation was limited. This result provides support to the earlier research by Griffin and Bock (2000), which demonstrated that speakers need to look at the referents shortly before naming them during the production of sentences.

The pattern of interactions largely supported the idea that the perceptual effects observed in this study are restricted to the identification of sentential starting points.

Although the interaction between the Cue Location and the Cue SOA was only reliable by the F1 analysis (see example in Figure 22), the result supports my prediction about the increased effect of longer SOAs on cue location: Speakers were more likely to alternate word order as a factor of the cue location when the cue was explicit. At Subject, this effect was reflected in the difference between the implicit and the explicit SOAs in both the agent-cued and the patient-cued condition (agent-cued: $F1(1,23)=5.778$; patient-cued: $F1(1,23)=41.009$). However, both at Verb and Object the same effect was driven solely by the difference in the patient-cued condition (Verb: $F1(1,23)=24.386$; Object: $F1(1,23)=13.487$). This pattern of results suggests that cue SOA has a greater effect on the time to initiate sentence constituents when the visual cue favours the non-preferred starting point. When the cue is on the agent – the element that typically starts the canonical English transitive frame – the difference in the cue SOA is only important at the beginning of a sentence.

Similarly, instruction interacted with the cue location; the cueing effect was stronger when the instruction was constraining (see Figure 23). This interaction was significant only at Subject and Verb suggesting that the effect slowly decays over the course of sentence production. For both constituents, the interaction was due to facilitation in the patient-cued condition (Subject: $F1(1,23)=4.589$; $F2(1,31)=8.916$; Verb: $F1(1,23)=17.073$; $F2(1,31)=24.902$). Again, we can conclude that the constraining instruction produced facilitation effect when the cue is associated with the patient; the production of the agent-driven canonical frame did not benefit from speaker's visual interrogation freedom.

Finally, the interaction between the experimental instruction and the Cue SOA also gradually diminished in the course of sentence production. The pattern of this interaction at Subject is portrayed in Figure 24a. Speakers' rates of constituent production did not differ as a factor of the constraining instruction when the cue was explicit. However, when the cue was implicit, the initiation of the first two sentence constituents proceeded much faster if the instruction was non-constraining (Subject: $F1(1,23)=5.494$; $F2(1,31)=13.806$; Verb: $F1(1,23)=9.543$; $F2(1,31)=21.623$). This result suggests that a stronger indexing of the cued location attenuates processing difficulties associated with sentence production in a visually constrained environment. A stronger cue, although semantically uninformative, seems to provide a better attentional state, which facilitates rapid apprehension and, as a result, quicker assignment of grammatical roles.

Because the experimental manipulations in Experiment 6 did not vary speakers' access to the conceptual information about the target events, the power and the persistence of the cueing effect can only be explained as a factor of perceptual salience of the cue. In this respect, the results of Experiment 6 support the hypothesis that perceptual effects in sentence production are localized primarily to the starting point selection regardless of the relative salience associated with the perceptually primed referent. In other words, making a referent increasingly salient does not improve the chances of the perceptual effects continuing far beyond the rapid apprehension stage. At the same time, cueing manipulations only made a difference in the processing of non-canonical sentences – the passive voice in our situation. When the cue was on the agent, the combination of cueing factors neither helped nor complicated the formulation of the sentence. It prompts the following conclusion: Strictly incremental formulation of sentences is associated with the production of highly automated canonical structures; when a speaker embarks on the formulation of a non-canonical sentence, a larger degree of event (re)analysis and structural pre-planning is necessary. It has to be noted that this conclusion should be limited to the production of sentences without any discourse support because none of the cue factors changed the given/new status of either referent and there was no previous discourse information available to the participants in the current study.

6.1.7. Results. Experiment 7

6.1.7.1. Cueing efficiency

The analysis of cueing efficiency confirmed that both implicit and explicit visual cues successfully attracted participants' attentional focus. Table 33 summarizes the cueing efficiency pattern observed in Experiment 7. The t-tests compared the cueing success probabilities to chance (.5) and confirmed that the cueing manipulation was successful in all combination of experimental factors.

6.1.7.2. Word order analysis

The data loss due to the failure to describe a target picture or providing descriptions that did not conform to the experimental instructions comprised 1.7% of the total data. The good data were coded using the same word order types as in Experiment 6. Table 34 summarizes the observed probabilities to produce Active Voice sentences across the experimental conditions. Examination of the data in Table 34 reveal a pattern very similar to the one observed in Experiment 6: Speakers were 35% more likely to produce passive voice sentence in the patient-cued condition, 4% more likely to do so when the cue was explicit, and 19% - if the instruction was constraining. However, although they were sensitive to the manipulations of the cueing factors, participants continued to rely upon the canonical active voice SVO even when the strongest combination of the cueing factors suggesting otherwise. An informal comparison of the word order data in Experiments 6 and 7 suggests that the difference in the priming paradigms used in the two studies did not result in any substantial contrast. The three slots marked red in Table 35 illustrate some noticeable difference: Participants in Experiment 7 produced 8% more passive voice sentences in the constraining/implicit/agent-cued condition, they were 5% more likely to produce passive voice sentences in the constraining/explicit/patient-cued condition, and they were surprisingly less likely to produce passive voice sentences in the non-constraining/implicit/patient-cued condition.

Table 35 reports the results of a series of ANOVA tests conducted on the observed probabilities. Confirming my informal examination, a picture surprisingly similar to the results of Experiment 6 emerged. Virtually the same main effects and interactions were observed. Participants reliably alternated their word order choices dependent on the cue location, the cue SOA, and the experimental instruction. Active voice canonical sentence persisted as the speakers' default choice, more so with the non-constraining instruction. One noticeable difference is a larger magnitude of Cue Location x Cue SOA interaction, which may be a result of the referential cue informativity. The same can be said about Cue Location x Instruction interaction: a referential cue was associated with a better indexing power when participants had a chance to identify the cued entity in explicit cueing condition.

A separate ANOVA analysis using Experiment as a between-subject and a within-item variable revealed no separate effect of Experiment on participants' choice of word order and no reliable interactions between Experiment and other factors.

6.1.7.3. First fixation analysis

Participants failed to comply with the constraining instruction and looked at the non-cued referent in the constraining condition 18 % of the total cases. One sample t-test confirmed that the compliance with the instruction was better than chance ($t1(23) = 7.289$, $p<.000$, $t2(31) = 12.36$, $p<.000$).

Table 36 summarizes the harmonic mean reaction times for first fixation onsets to each of the referent's interest areas across the experimental conditions. Although the emerging picture is once again similar to the results of Experiment 6, there are some noticeable differences (marked red in Table 37) in the reaction times associated with the looks to the non-cued referents. (1) Participants in Experiment 6 were faster to initiate their first fixations to the non-cued referents in the non-constraining/implicit/ condition; (2) they were also faster to switch attention to the non-cued patients in the non-constraining/explicit/agent-cued condition; (3) they were slower to execute initial looks to the non-cued agents in the non-constraining/explicit/patient-cued condition.

Table 37 presents the results of ANOVA tests performed on the latencies to fixate the *cued* referent. Because of the nature of the experimental instruction, the first fixation latencies to the non-cued referent were only analyzed for the non-constraining condition. The data pattern is virtually the same as in the previous study. Once again, there were no reliable effects of the cue location and the cue SOA, but there was an identical effect of instruction: First fixation tended to be initiated later in the constraining instruction condition. More likely than not, such slow-down does not reflect anything but the fact that the speaker's visual focus is bound to the cued referent before the target picture appears on the screen, therefore, the reported latencies for constraining condition reflect just another local saccade within an already cued and fixated region. A comparison between first fixation onsets to the cued referent across Experiments 6 and 7 revealed no significant effect of Experiment and no interaction between Experiment and other factors.

Results of the analysis of the first fixation onsets to the *non-cued* referent in the non-constraining condition are summarized in Table 38. Like in Experiment 6, there were facilitating effects of both Cue Location and Cue SOA on the latencies to initiate first fixations to the non-cued referents: Speakers were more likely to use the cued referent as "good-enough" anchor for when the cue was on the agent and when it was explicit. However, this time these two factors did not reliably interact; also, while the effect of Cue Location is noticeably smaller than in Experiment 6, the effect of Cue SOA is slightly larger. This pattern of results confirms that an attentional cue to the event's agent results in a stronger tendency to use the cued referent as the indexing point for further conceptual analysis and sentence formulation, more so in the situation when the cue confirms the expectation for the canonical causality. This effect becomes even stronger when the cue is explicit. In such a scenario, the initial fixations to the non-cued referent are delayed beyond the average time of sentence onset, which means that the incremental sentence formulation begins immediately as the "strongly" cued referent' role is revealed to the speaker by the target picture presentation. Interestingly, the tendency to delay the first fixations to the non-cued referents was more pronounced in Experiment 6, in which the cue was uninformative.

A comparison between the two experiments revealed a reliable effect of Experiment ($F1(1,46)=14.358$, $p<.000$, $F2(1,60)=138.993$, $p<.000$) and a reliable interaction between the Cue Location and Experiment ($F1(1,46)=12.179$, $p<.001$, $F2(1,60)=23.216$, $p<.000$) (see Figure 25). Participants in Experiment 6 were on average 191 msec slower to initiate the initial fixations to the non-cued referents than in Experiment 7. The cross-over interaction effect was carried by a 445 msec slow-down in the agent-cued condition in Experiment 6 ($F1(1,47)=24.303$; $F2(1,31)=127.343$).

6.1.7.4. Eye-voice span analysis

Eye-voice span data are reported for Subjects and Objects only. The procedure for calculating EVS values remained the same as in previous experiments. The missing values accounted for 3.1% of the total number of trials. Using the same reasoning as in Experiment 6, only EVS values for the non-constraining condition were formally analyzed. Tables 39 and 40 summarize mean EVS values across the sentence constituents and experimental conditions for both types of the produced structures.

The general pattern of the EVS data distribution is similar to that observed in Experiment 6. The EVS values in the constraining condition are much larger, which supports our conclusion about the cognitively more difficult processing associated with production of sentences in a visually constrained environment. The fact that it was true for both Objects and Subjects in an experiment using a referential priming paradigm further supports my interpretation of EVS measurement as an indicator of more distributed processes including both argument structure formulation and lexical access.

One important observation following from the data summarized in Table 40 is that the difference between the EVS data for Subjects and Objects is larger (91 msec) than in Experiment 6. Also, EVS values in the non-constraining condition in Experiment 7 are substantially larger for the Subjects and slightly smaller for the Objects than in Experiment 6. This may suggest that early access to both perceptual and conceptual information about the candidate for the Subject results in deeper processing of the corresponding referent. Note that a reverse effect of the EVS inflation from Subject to Object was observed during normative English sentence production in Experiment 2. This interpretation receives support from an informal examination of the passive voice EVS data: EVS values for Subject are noticeably larger in Experiment 7.

The results of ANOVA tests performed on the mean EVS values at Subject are summarized in Tables 47. Statistical analysis of the EVS data for Object did not reveal any reliable effects. The only reliable main effect established for the Subject EVS values was the general effect (62 msec) of the cue location: Processing of Subject-related referents were easier when the cue was on the agent. A comparison of the EVS data for Subjects and Objects confirmed a reliable difference between the constituents ($F1(1,23)=8.641$; $F2(1,63)=12.876$): Processing of Objects was in general easier than processing Subjects. Comparisons between the two studies revealed no general effect of Experiment and a reliable Experiment x Cue SOA interaction ($F1(1,46)=11.929$, $p=.001$; $F2(1,60)=15.391$, $p<.000$) (see Figure 26). The post-hoc analysis confirmed that this interaction was due to the difference in the EVS performance in the explicit cue condition

($F1(1,47)=10.137$; $F2(1,31)=28.315$). Finally, analysis with Experiment and Constituent as independent variables revealed a reliable interaction between these two factors ($F1(1,46)=7.548$;$F2(1,63)=15.498$) (see Figure 27). Separate comparisons between Subject and Object EVS data between the two experiments confirmed that the aforementioned interaction was due to a reliable difference at Subject ($F1(1,47)=7.518$; $F2(1,63)=32.514$) but not at Object.

6.1.7.5. Name onset analysis

This data loss in the name onset analysis was 2.9 %. Tables 42 and 43 report harmonic the mean name onset values in Experiment 7 at each sentence constituent and for both syntactic structures. Tables 44-46 summarize the results of the statistical analysis performed on the name onset latencies in Experiment 7. Main effects of the cueing factors remain almost the same as in Experiment6, but the effect of the cue SOA is much stronger. This observation suggests that, although word order analysis revealed no facilitation affect for the referent preview, the opportunity to identify the cued referent before the rest of the event is available may lead to faster production rates. Interactions between the experimental factors at Subject and Verb were now limited to the interaction between the cue SOA and the experimental instruction, and there were no interactions at Object. The Instruction x Cue SOA interaction at Subject and Verb followed the same pattern as in Experiment 6 (see Figure 24b): Speakers' rate of sentence production did not differ as a factor of the constraining instruction when the cue was explicit. However, when the cue was implicit, the initiation of the first two sentence constituents proceeded much faster if the instruction was non-constraining (Subject: $F1(1,23)=6.198$;

$F2(1,31)=7.771$; Verb: $F1(1,23)=5.311$; $F2(1,31)=9.504$). This result confirms that a stronger indexing of the cued location attenuates processing difficulties associated with sentence production in a visually constrained environment. A stronger cue, semantically informative or not, seems to result in a better attentional state, facilitating conceptual analysis and, as a result, a more rapid assignment of the grammatical roles in a spoken sentence. When the cue is implicit, the potential problems of the restrictions on visual interrogation of the event become more important because a thorough event analysis is necessary in both experimental instruction conditions.

An informal comparison of the name onset data patterns between the two experiments reveals that, quite noticeably, the production of sentences in the implicit cue condition was initiated later in Experiment 7; a reverse pattern, however, is observed in the explicit cue condition: Speakers were faster to initiate their sentences in Experiment 7 where the cue was referential. Also, comparison of the passive voice sentence production between the two studies demonstrates that this operation was generally slower when the cue was referential. Such difference is probably due to a deeper referent processing associated with the referential priming procedure. Once again, these observations were put to an ANOVA test using Experiment as a between-group factor. There was no main effect of Experiment at either of the sentence constituents, but there was a reliable interaction between Experiment x Cue SOA at Subject ($F1(1,46)=19.590$; $F2(1,60)=19.873$), Verb ($F1(1,46)=12.422$; $F2(1,60)=11.107$), and Object ($F1(1,46)=12.214$; $F2(1,60)=9.798$) (see example in Figure 28 for Verb). The detailed analysis of this interaction (see Table 47) confirmed that the cross-over effect of this interaction was due to a faster processing in the explicit cue condition and a slower processing in the implicit cue condition in Experiment 7.

6.2. Discussion

The two studies reported in Chapter 6 explored extemporaneous production of English transitive sentences in an environment associated with varying degrees of (1) salience of the event's referents and (2) freedom of visual interrogation of the scene elements. In addition, Experiment 7 analyzed how speakers use a combination of perceptual and conceptual information provided by the cue when choosing the starting point and the resulting grammatical structure of a spoken sentence.

In general, attentional manipulations were equally successful regardless of whether the deployed cue was on the agent or on the patient, whether it was implicit or explicit, and whether the visual focus was constrained or not. However, varying power dimensions of the deployed cue resulted in consistent difference in speakers' behaviour as was revealed by a battery of tests.

Varying the cue location from agent to patient resulted in a higher probability of producing passive voice sentences in both studies, and speakers' word order choices relied on the cueing manipulation more when the freedom of visual interrogation was constrained to the cued referent. This effect was mirrored in the name onset latencies, which reliably varied as a factor of cue location: The cue to the agent facilitated faster sentence production. However, there was also a strong reliance of the processor on the top-down constraints of grammatical canonicality. Even when the combination of cueing manipulations strongly favoured passive voice speakers quite often produced active voice sentences. This result seriously limits the validity of perceptually motivated accounts of word order assignment even for the languages with rigid word order systems like English. Together with Experiments 1-5, Experiments 6 and 7 provide compelling evidence for a substantial reliance of speakers on the canonical grammar, probably more so in the free word order languages.

Contradicting the prediction about a stronger cueing effect in the referential priming setting, the conceptual information provided by the cue in Experiment 7 did not lead to a noticeably stronger alternation of the produced word order or facilitation of name production. This suggests that the given/new contrast known to enhance the conceptual accessibility of the referent and bias the assignment of the sentential starting point and the subsequent choice of the word order is not promoted via referential priming. It means that the given-before-new effect is probably established later during lexical processing and supports previous findings that speakers do not necessarily lexicalize visually presented referents even when they have enough time unless they have to do so (Bloem and La Heij, 2003; Roelofs, 2003, Roelofs, 2006). However, better indexing associated with explicit cueing had a facilitating effect on the identification of the sentential starting point, the ease, and the speed of conceptual analysis and sentence formulation.

The use of secondary behavioral measurements revealed a more detailed picture of how the information uptake at different stages of event examination and sentence formulation varied as a factor of experimental manipulations. The time-course of the first fixations to the referents confirmed the success of attentional manipulations and revealed that the onsets of initial fixations to the non-cued referents differed as a factor of the cue location. When the cue was on the agent, speakers tended to dwell on the cued referent longer effectively delaying their initial interrogation of the rest of the scene almost to the point of the sentence formulation. Delayed fixations to the non-cued referent (a condition that supports the preferred causality) suggest that perceptual processing does not have to be fully completed before conceptual processing can begin. When the cue was referential, this delay was reliably smaller suggesting that the conceptual information provided by the cue triggered faster identification. This pattern suggests that cue to the agent prompted speakers in both studies to immediately assume the canonical causality without further examination of the event. Supposedly, full conceptual analysis in such case happened in a cascade fashion – alongside incremental sentence formulation. This finding supports my previous claim about a certain degree of automaticity involved in the conceptualization of the event.

When the cue favoured the patient, speakers quickly switched their attention to the preferred starting point – the agent –proceeding with rapid apprehension. Early switching to the non-cued referent in the condition that does not support the preferred causality suggests that perceptual processing of the event under such circumstance has to be completed before conceptual analysis can begin. If the perceptual and/or conceptual environment did not warrant the early starting point assignment, full conceptual analysis became necessary before sentence formulation could commence. This result extends the omnipresence of incrementality in sentence production on one hand and limits it, on the other. On one hand, it demonstrates that sentence formulation can start even before the whole event is apprehended. On the other, the apprehension of less preferred referential schemes and the production of less common syntax may depend on a much higher degree of full event analysis. This processing difference provides further detail to the model proposed by Griffin and Bock (2000), which separates rapid apprehension and incremental sentence production of transitive sentences.

The distribution of speakers' eye-voice spans reflects the transfer of processing from rapid apprehension to the completion of the sentence plan. Most of the EVS effects in both experiments were limited to processing subjects. In Experiment 6, the magnitude of the Subject- and the Object-related EVS values was comparable suggesting that perceptual cueing helped attenuate the increase of processing load associated with English transitive sentence production observed in Experiment 2. In Experiment 7, the EVS values for the Object were reliably smaller than for the Subject, which suggests that early access to perceptual information results in deeper processing of the corresponding referent, which helps to reverse the effect of the EVS inflation from Subject to Object observed during normative English sentence production in Experiment 2.

Cue to the agent significantly reduced the processing load associated with the promotion of the cued referent to the sentential Subject position in both experiments. However, the analysis of the Object-related EVS data did not reveal any difference. It suggests that the referential scheme is realized prior to the mapping of the grammatical roles in a sentence. The cognitive load reflected by EVS was insensitive to the explicit/implicit cue contrast, which shows that the advantage of having referential information was not used until sentence formulation started. However, the presence of a reliable interaction between Experiment and SOA observed both in name onset latencies and the EVS data in Experiment 7 confirms that the availability of referential information starts to take effect when the conceptual plan of the sentence is being laid out and that this effect continues later while naming sentence constituents. Limited interaction effects between the cueing parameters in Experiment 7 suggests that availability of conceptual information rapidly replaced perceptual effects biasing the processor away from reliance on salience factors. Finally, EVS data revealed that speakers started to experience processing problems even during the preparation of the Subjects and these problems only increased at Object. The latter is unsurprising because speakers could not visually examine the non-cued referents in the constraining condition, and more often than not these referents were subsequent Objects. The fact that conceptual processing of *Subject-related referents* was also quite difficult in the constraining condition suggests that having full access to the event information is a partial prerequisite for sentence generation.

Analysis of the name onset data showed that the Experiment by Cue Location interaction observed in the first fixation analysis did not continue beyond referent identification as its was not associated with any facilitation of lexical access. On the other hand, speakers in both studies were faster to initiate sentences when the cue was explicit, with this effect being much stronger when a referential cue was used. Better indexing provided by explicit cueing improved lexical access speeding sentence production. However, a cross-over interaction between Experiment and the cue SOA provides additional information about how and when the benefit of the referential cueing was used. Participants in Experiment 7 were faster to initiate their sentences than their counterparts in Experiment 6 in the explicit cue condition and they were slower to do so in the implicit cue condition. It is an interesting contrast because it suggests limitations to the rapid information uptake account found in (Dobel & Gumnior, 2004). Participants in the Dobel and Gumnior's study did not have to provide full sentences; they simply needed to observe the referents and interactions at very short SOAs and then answer post-hoc questions like *"What happened to X?"* When providing full-fledged extemporaneous sentences about the events, speakers tended to "dislike" implicit referential cueing (Experiment 7) possibly because a brief exposure to a more complex cue confused them: They had enough time to notice it was something different from a simple flash but not enough time to guess what it was. This noticeably slowed down sentence production. On the other hand, explicit referential cueing facilitated production because speakers had enough time to benefit from the conceptual information provided by the cue. Rapid recognition of both referents and events is apparently quite possible even at very short presentation times as Dobel and Gumnior's study demonstrates, but the full-fledged sentence production becomes complicated when a potentially informative cue appears on the screen for a time too short for its recognition.

Sentence production was also complicated by the inability to visually interrogate the scene starting from the very sentence onset. Although non-canonical sentence production was generally less likely, it proceeded faster when there was less constraint on the visual event interrogation. When visually constrained, non-canonical sentence formulation had to rely on explicit cueing; if it was not provided, such production was slow and unlikely to happen at all.

6.3. Chapter Conclusions

The overall pattern of results in Experiments 6 and 7 suggest that visual cueing of a referent results in stable assignment of the sentential starting point, facilitates processing of the associated grammatical role (Subject), and results in consistent alternation of word order. This has now been confirmed by a whole battery of behavioral data ranging form the probability of alternating the available syntax to initial fixation analysis, to name onset latencies, finally to a measurement specifically related to the processing difficulties associated with online formulation of a sentence. Explicit cueing of attention substantially improves the cueing effect making it more powerful both in its ability to bias speakers' grammatical choices and to drive earlier selection of argument structure paired with a more rapid speech production.

Analysis of the EVS and name onset data suggested that apprehension of the event and the sentence production was increasingly more difficult *at every point* when participants could not freely examine the event. On one hand, this result supports the idea about a tight link between the looks to the entities and the production of their names in a sentence (cf. Griffn & Bock, 2000). On the other, the fact that speakers experienced processing difficulties already at the very beginning of a sentence suggests that a full apprehension of the event may be necessary before sentence formulation can start. Finally, referential priming advantage, although not particularly useful for the starting point assignment and the word order alternation, sped up processing conceptual and lexical information about the referents making the overall constituent onsets faster. This suggests a certain degree of independence of conceptual effects from syntactic processing.

Chapter 7. Perceptual, conceptual, and syntactic priming in English transitive sentence production

Experiments in the preceding part of the thesis analysed how speakers' behavior differs depending on the details of the perceptual map of the cognised event, its conceptual properties, and the grammar of the language in use. In addition to the already established perceptual and referential priming environments, experiments reported in this chapter manipulate speakers' access to lexical and structural information during sentence production. Therefore, one research question for the studies reported below is how perceptual priming effects will interact with priming lexical and syntactic information.

The second research question has to do with how, in this tripartite priming environment, speakers' behavior will differ as a factor of the cue informativity. Similarly to Experiments 6 and 7, the studies in Chapter 7 differ in the visual cueing procedure: Experiment 8 uses perceptual and Experiment 9 – referential cueing. Increase in conceptual accessibility established without further lexical support in Experiment 7 did not improve the cueing effect although it facilitated conceptual analysis and overt sentence generation. Comparisons between Experiments 6 and 7 showed that conceptual effects can extend well beyond the event apprehension phase and influence both eye-voice spans and the latencies to produce the sentence constituents. It is possible that the relative inability of the referential priming to increase *structural alternation* in Experiment 7 was due to the lack of supporting lexical and/or structural information. It is possible that, when such information is provided via lexical or structural priming, the processor might accommodate conceptual effects better.

7.1. Experiments 8 and 9

7.1.1. General structure

The main goal of the two studies reported below was to analyze effects of perceptual, semantic, and syntactic priming during extemporaneous production of English transitive sentences. Both experiments continued to use the Visual World paradigm: The materials were presented to the participants on the computer screen, and the spoken sentences were collected in real time as participants viewed and described the pictures. Three factors were independently manipulated at two levels each in both experiments: (1) the perceptual prime, (2) the lexical prime (verb), and (3) the syntactic prime (active or passive voice). Hence, both experiments used a 2x2x2 design with all three factors manipulated within participants and between the experimental materials.

Like in the two previous experiments, the difference between the experimental protocols in Experiments 8 and 9 was the nature of the visual cue used to attract participants' attention to either the agent or the patient of the target event. The cue used in experiment 8 was similar to the one used in Experiment 6: A red dot with the same parameters as in Experiment 6 appeared for 500 msec. in either the Agent or the Patient region prior to the presentation of the target display. Hence, cueing in Experiment 8 was always perceptual and explicit. Cueing in Experiment 9 was similar to the one used in Experiment 7: One of the event's referents appeared on the screen for 500 msec. prior to the target display presentation, therefore, cueing in Experiment 9 was always referential and explicit. The referential cue always appeared in the same area as it would appear within the target event, which ensured that a participant was looking at the cued referent once the target picture appeared on the screen.

Syntactic priming (*Prime*) was achieved with the help of participants reading out loud either an active (AV) or a passive voice (PV) sentence prior to describing a target event (Bock 1986). The names of the referents used in the priming sentences never corresponded to those in the subsequent target picture. In 50% of the priming materials the sentential verb corresponded to the verb necessary to describe the target event. Therefore an effect of partial semantic overlap was achieved (Branigan & Pickering, 1998). The verb-overlap (*Match*) factor was manipulated between two levels: match (VM) and no-match (VN) conditions. The dependent variables in both studies were (1) the probability to produce Active Voice sentences, (2) the onset of the first fixations to the interest areas, (3) the eye-voice spans for each sentence constituents, and (4) the name onset latencies for each sentence constituent.

7.1.2. Materials

For the purposes of the current studies, I continued using the same set of 64 simple black-and-white line drawings of transitive events (see example in Figure 1). I also created a separate set of syntactic priming materials. Table 48 lists priming materials according to the experimental conditions. The same set of filler pictures as in Experiments 6 and 7 ensured separations of the target trials by at least two filler presentation (see Figure 16). In addition, a set of 130 filler locative sentences was created to make filler trials similar to the experimental trials. Table 49 contains the complete set of filler sentences. The coding of interest areas followed the already established logic.

7.1.3. Participants

A separate group of 24 native English speakers with normal or corrected-to-normal vision and no language-related impairments took part in each study. All the participants were undergraduate students at the Department of Psychology University of Glasgow. There were 12 male and 12 female participants in Experiment 8; 10 male and 14 female participants in Experiment 9. The mean age of the participants was 20.3 years in Experiment 8 and 21.2 in Experiment 9.

7.1.4. Apparatus

Same as in Experiments 4-7.

7.1.5. Experimental procedure

Participants were not informed about the nature of experimental manipulations or the exact purpose of the studies. They were told that the study was concerned with speaking about what you see on the computer screen. Before the experimental session began, participants practised describing 5 target trials. Participants did not practice describing filler trials. To avoid the surprise of seeing completely unpractised trials, participants were told that they would sometimes see novel pictures that contain geometrical shapes. Similarly to Experiments 6 and 7, the instruction was to simply describe these pictures in one sentence using locatives like *above, below, right,* or *left.*

Typical target trial sequences in Experiments 8 and 9 are illustrated in Figures 29 and 30. The experimental procedure was similar to the one reported for the previous studies. Each trial began with the presentation of a fixation mark in the center of the screen. Once the participant visually located and fixated the fixation mark, the priming sentence appeared on the screen; therefore the structural prime was available to participants before the visual cue was. The prime was either an active or a passive voice sentence written in Arial font, size 16 and always presented in the middle of the screen in the direct view of the participant. Participants read the sentence out loud as soon as it appeared on the screen and pressed the spacebar key. Another fixation mark appeared on the screen, now in its lower area. Once a participant successfully fixated the second fixation mark, the cue to one of the referents appeared on the screen for duration of 500 msec. Finally, the target picture appeared in the centre of the screen. The parameters of the target picture presentation were the same as described for Experiments 6 and 7. Participants described the target pictures using a single sentence extemporaneously as soon as the pictures appeared on the screen. Once participants finished their descriptions, they pressed space bar key again and a new trial sequence started. If no description was collected for the trial, an automatic 7700 msec timeout triggered the commencement of the next trial sequence.

The filler trials followed a similar sequence: A filler sentence (see Table 49) appeared on the screen after the initial fixation mark presentation. The second fixation mark appeared on the screen after that; then – the cue in either right or left part of the screen; finally – the filler picture was displayed to the participants and they described it.

7.1.6. Results. Experiment 8

The same procedures for data coding and dealing with the missing values as in Experiments 6 and 7 were used.

7.1.6.1. Cueing efficiency

Analysis of the cueing efficiency confirmed that the visual cue successfully attracted participants' attention. Table 50 summarizes the cueing efficiency results for Experiment 8. Two-tailed t-tests performed on the mean proportions as compared to chance (.5) returned reliable cueing results for all 8 experimental conditions.

7.1.6.2. Word order analysis

The loss of the data due to the lack of description or providing sentences that did not conform to the experimental instruction was 0.9%. Table 51 summarizes the observed probabilities to produce Active Voice sentences across the experimental conditions. The results of the ANOVA tests conducted on the observed probabilities are reported in Table 52.

Examination of Tables 51 and 52 reveals that (1) speakers were 14% more likely to produce passive voice sentences when they read a passive voice prime, (2) they were 9% more likely to alternate the produced word order when the verbs between the prime and the target sentences were matched, and they were 19% more likely to use the passive voice when the visual cue was on the patient. All the three main effects were reliable by both subject and item analyses. The analysis of observed interactions (Figures 31 and 32) confirmed reliable interactions between the Prime and the Match and between the Cue and the Match factors. There was no reliable interaction between the Cue and the Prime. Comparisons between relevant data groups showed that, when primed to produce passive voice sentences, speakers were 14% more likely to do so when the verbs in the prime and the target were matched ($F1(1,23)=11,527$; $F2(1,31)=8.592$). Matching verbs in the active voice priming condition did not lead to any considerable difference in the word order choice. The latter results effectively replicate earlier findings by Pickering and Branigan (1998). The Cue x Verb Match interaction required a more exhaustive analysis because, although the existence of such an interaction is predicted by the model, I had no specific expectation about the nature of this interaction. Figure 32 illustrates the interaction pattern along the Match condition. Comparison along the Match condition showed reliable differences in both groups: participants alternated word order as a factor of cue in both Match ($F1(1,23)=4.093$; $F2(1,31)=7.912$) and No Match ($F1(1,23)=15.167$; $F2(1,31)=49.422$) conditions. However, the size of effect in No Match condition (25%) is twice as large as in Match condition (12%), which demonstrates that when no verb information was available, speakers relied more on the direction of the visual cue; when such information was available, they tended to use the latter instead of alternating the word order as a function of the cue direction. Comparison along the cueing condition showed that the same interaction was carried by the difference in the agent-cued but not in the patient-cued condition. When the cue was on the patient, the difference in the Match factor did not lead to any noticeable alternation of word order; when the cue was on the agent, participants were 15% more likely to produce passive voice sentences when the verbs between the prime and the target were matched ($F1(1,23)=23.083$; $F2(1,31)=15.161$).

7.1.6.3. First fixation analysis

Raw fixation data were analyzed with the help of the Data Viewer program. The same coding, filtering, and missing value detection procedures as in Experiments 6 and 7 were used. The missing data accounted for 1.3% of the data; their values were replaced with the means of the corresponding conditions. Table 53 summarizes the harmonic mean reaction times for first fixation onsets to each of the interest areas across the experimental conditions. ANOVA performed on the initial fixations to the *cued* referent did not reveal any reliable effects or interaction. Results of the ANOVA performed on the initial fixations to the *non-cued* referents are summarized in Table 54. Examination of Tables 53 and 54 reveal that participants were 130 msec faster to switch attention from the cued to the non-cued referent in the patient-cued condition and that they were 165 msec faster to do so when the verbs in the prime and the target sentences were matched. The former result replicates my earlier findings that the cue to the agent leads to earlier commitment to the canonical event apprehension and, as a result, an earlier commitment to the canonical word order. The latter result demonstrates that speakers are faster to proceed with the referential analysis when they have previewed the verb.

7.1.6.4. Eye-voice span analysis

The procedure for calculating EVS values remained the same as in Experiments 6 and 7. The values less than 100 msec were replaced with the ones calculated manually using the fixation immediately preceding the one detected by the script. The missing values accounted for 4.3% of the total number of trials. Such values were replaced with the mean EVS value for the corresponding condition. Table 55 summarizes mean eye-voice span values across the sentence constituents and experimental conditions.

Once again, statistical analyses are reported only for Subject and Object. Participants fixated the verb area only in 20 out of 64 experimental trials on average.

Very few of these infrequent fixations occurred before the sentence formulation started, therefore calculating eye-voice spans for the Verb was only possible for the trials on which participants fixated the verb interest area prior to producing the verb's name. There were 17 such trials per participant on average. The maximum number of such trials for a single participant was 25 and the minimum number 9. Therefore, I only report the verb eye-voice span means for the trials that yielded such observations without formal statistical analysis of the latter. Table 56 summarizes the EVS values in the trials with the passive voice sentences. Tables 57 and 58 report the results of ANOVA tests for the EVS values at Subjects and Objects.

Informal examination of Table 55 reveals that the EVS values for the Subjects were slightly longer (26 msec) than for the Objects. However, this difference was not reliable. Statistical analysis of participants' EVS showed that the only experimental manipulation that resulted in a reliable variation of the EVS performance was the Match factor: Eye-voice spans were 48 msec shorter for the Subjects and 66 msec shorter for the Objects when the verbs were matched. No reliable effect of any other experimental factor was detected. There were two reliable interactions in both the Subject and the Object EVS data: between the Cue and the Prime factors and between the Prime and the Match factors. Detailed analyses of these interactions revealed that the participants' performance did not differ as a function of the Prime manipulation when the cue was on the agent. However, when the patient was cued, there was a 88 msec facilitation in the patient-cued condition when the prime sentence was a passive voice one ($F1(1,23)=11.242$). Analysis of the Prime x Match interaction confirmed 78 msec facilitation observed when the active voice prime was combined with matching verbs from the prime to the target

(F1(1,23)=19.548; F2(1,31)=3.075). Very similar results were revealed for the EVS data at the Object: There was a 128 msec facilitation effect in the patient-cued/passive-voice condition (F1(1,23)=24.356) and a 132 msec facilitation effect in the active-voice/verb-match condition (F1(1,23)=29.441; F2(1,31)=4.038).

7.1.6.5. Name onset analysis

The same coding and filtering procedure as in previous experiments was applied. The name-onset-related data loss in Experiment 8 was 1.9 %. Table 59 reports harmonic mean values for the name onset latencies in Experiment 8 at each sentence constituent. Table 60 summarizes the name-onset data for the passive voice sentences. Tables 61-63 report the results of ANOVA tests at each sentence constituent.

Informal examination of the results summarized in Tables 59 and 61-63 demonstrates that the effect of the syntactic prime on the latencies to initiate sentence constituents was only observed at the Verb: participants were 27 msec faster to initiate verbs after reading the active voice primes. Therefore, the structural information provided by the structural prime was incorporated during the sentence formulation at the stage of the verb production. At the same time, matching the verbs between the prime and the target sentences resulted in consistent facilitation at all three sentence constituents so that the production was around 80 msec faster when the verbs were matched. This result further confirms an independent facilitation effect of verb overlap on a variety of behavioral measures representing sequential stages of sentence formulation starting with the early perceptual analysis of the event through to naming sentence constituents. Similarly, cueing the agent facilitated production at each step: Speakers were on average

65 msec faster to produce sentence constituents when their attention was directed to the agent. No interactions between the experimental factors were revealed by the analysis of the name onset latencies.

7.1.7. Results. Experiment 9

As discussed above, Experiment 9 uses the same logic, the same experimental protocol, and the same coding procedure as in Experiment 8. The only difference between the two studies is the type of visual cue used to attract participants' attention. Similarly to Experiment 7, Experiment 9 used a preview of one of the referents as a cue to either the agent or the patient. The time the referent was viewed by the participants was 500 msec.

7.1.7.1. Cueing efficiency

The analysis of cueing efficiency confirmed the visual cue successfully attracted the participants' attentional focus. Table 64 summarizes the cueing efficiency results for Experiment 7. Two-tailed t-tests performed on the mean proportions as compared to chance (.5) returned reliable cueing results for all 8 experimental conditions.

7.1.7.2. Word order analysis

Data loss due in Experiment 9 comprised 0.83% of trials. Table 65 summarizes the probabilities of producing Active Voice descriptions across the experimental conditions. Table 66 reports the results of the ANOVA tests conducted on these probabilities.

The pattern summarized in Tables 65 and 66 is similar to the one discussed above for Experiment 8: Participants were 28% more likely to produce passive voice sentences when they received a passive voice prime, 21% more likely to do so when the cue was on the patient, and they were 9% more likely to alternate the word order when the verbs matched between prime and target sentences. However, there are some interesting differences between the two experiments. Like in Experiment 8, the three experimental factors had separate effects on the probability of producing one of the two possible structures, but the priming effect was noticeably stronger this time (14% in Experiment 8). In contrast with the results of Experiment 8, there was only one reliable interaction: Between the Prime and the Match factors. This interaction followed the same pattern as in Experiment 8: Participants were 20% more likely to alternate the produced word order as a factor of the verb overlap if the prime was a passive voice ($F1(1,23)=40,013$; $F2(1,31)=21.338$). Matching verbs in the active voice priming condition did not lead to any considerable difference.

Comparison between the two studies reveals that participants in Experiment 9 were 14% more likely than their counterparts in Experiment 8 to produce a passive voice sentence when they were primed by a sentence with the same structure. However, the contributions of the Verb Match and the Cue factors were the same in the two studies. I performed a separate ANOVA using Experiment as a between-subject and a within-item variable. The effect of Experiment on participants' choice of word order was only reliable by the item analysis ($F2(1,56)=7.390$): Participants produced 5% more passive voice sentences in Experiment 9, and this improvement was due to the interaction between Experiment and Prime $(1,46)=3.328$; $F2(1,56)=23.786$) (see Figure 35). Analysis of this interaction confirmed that it resulted primarily from the difference in the passive voice priming condition ($F1(1,47)=4.085$; $F2(1,31)=21.061$): Participants were 11% more likely to produce passive voice sentences in the passive voice priming condition when the cue was referential as in Experiment 9. The same difference between the perceptual and the referential cueing paradigms did not matter in the active voice priming condition.

7.1.7.3. First fixation analysis

The same coding, filtering, and missing value detection procedures as in Experiment 8 were used. The missing data accounted for 1.65% of the data, and their values were replaced with the mean value for the corresponding condition. Table 67 summarizes the harmonic mean reaction times for first fixation onsets to each of the interest areas across the experimental conditions. ANOVAs performed on the latencies to fixate the *cued* referents did not reveal any main effects or interactions. The same analysis performed on the latencies to fixate the *non-cued* referents revealed a marginally reliable effect of the Cue ($F1(1,23)=3.226$, $p.=.086$); once again, speakers were faster to switch their attention to the non-cued agent when the cue was on the patient. A comparison between the first fixation onsets in Experiments 8 and 9 failed to reveal a reliable effect of Experiment. There was, however, a reliable interaction between Experiment and Verb Match conditions ($F1(1,46)=14.157$; $F2(1,56)=3.535$, $p=.065$). This interaction (see Figure 36) was due to the difference in performance in the No Match condition ($F1(1,47)=3.153$, $p=.082$, $F2(1,31)=17.654$, $p=.000$): Participants were later to initially fixate the non-cued referent in this condition when the cue was perceptual (Experiment 8) than when it was referential (Experiment 9). This result suggests that the verb overlap between the prime and the target sentences facilitated faster resolution of the referential map regardless of whether the cue was perceptual or referential. When participants did not have the valid verb information, they actively used the referential information provided by the cue in Experiment 9; this is why participants were 137 msec faster to initially fixate the non-cued referent in Experiment 9.

7.1.7.4. Eye-voice span analysis

Missing values accounted for 2.5% of the total number of trials. Such values were replaced with the mean EVS value for the corresponding condition. Table 68 summarizes mean eye-voice span values across the sentence constituents and experimental conditions. For the reasons discussed earlier, formal statistical analysis of eye-voice spans was performed only for Subject and Object. In Experiment 9, participants fixated the verb area on average in 18 out of 64 experimental trials. Most of these fixations occurred after the sentence formulation started; therefore calculating eye-voice spans for the Verb was only possible for the trials on which participants fixated the verb interest area prior to producing the verb's name. There were 7 of such trials per participant on average. The maximum number of such trials for a single participant was 21 and the minimum number 3. Therefore, the mean EVS values for the Verb are only reported informally. Table 69 summarizes the EVS values in the trials with Passive Voice sentences.

Similar to Experiment 8, the EVS values for the Subject were slightly longer (36 msec) than for the Object. The difference was reliable this time ($F1(1,23)=6.475$; $F2(1,63)=4.955$) confirming that processing of the Subjects in English can be associated with a larger load. Informal analysis of the data in Table 68 suggests that EVS values in the passive voice sentences were shorter in the passive voice priming condition. ANOVA analysis performed on the Subject EVS values revealed the main effect of the Prime

($F1(1,23)=16.927$). Interestingly, the EVS values were reliably shorter in the passive voice priming condition. No additional effects or interactions were observed. The results of similar analysis performed on the mean EVS values for Objects are summarized in Table 70. Reliable effects and interactions were only confirmed by the F1 analysis. However, participants were 66 msec faster to process Objects in the agent-cued condition and 49 msec faster to process Objects in the Verb Match condition. The details of interactions between the Cue and the Prime and between the Cue and the Verb Match are revealed in Figures 37 and 38. In both cases, the interaction between the variables occurred due to the reliable difference in the patient-cued condition: Participants' EVS behavior was facilitated in the latter condition when (1) the prime sentence was in passive voice ($F1(1,23)=9.919$) and (2) when the verbs between the prime and target sentences matched ($F1(1,23)=17.485$). Comparison between the Experiments along the EVS measurement did not reveal any reliable effect of Experiment for any of the two constituents. Also, no interactions between Experiment and other factors were observed.

7.1.7.5. Name onset analysis

The name-onset related data loss was 2.32 %. The missing values were replaced with the mean value for the corresponding condition. Table 71 reports harmonic mean values for the name onset latencies in Experiment 9 at each sentence constituent. Table 72 provides an additional informal analysis of the name onset latencies in the passive voice sentences. Tables 73-75 illustrate the results of ANOVA tests conducted on the latencies to produce each sentence constituent.

Participants were on average 66 msec faster to initiate constituents' names when primed by an active voice sentence, 85 msec faster when their attention was directed to the agent, and 139 msec faster when the verbs between the prime and the target sentences matched. The same pattern was observed at each sentence constituent with no interactions between the factors. Informal examination of Table 72 reveals that the passive voice sentences were produced faster when the speakers were primed by another passive voice sentence. It was especially true when in the Verb Match condition.

Comparison of the name onset data between the two studies revealed a reliable effect of Experiment at the first two sentence constituents: Participants in Experiment 9 were 116 msec faster to initiate the Subjects ($F1(1,46)=3.454$, $p=.07$; $F2(1,56)=42.635$) and 72 msec faster to initiate the Verbs ($F2(1,56)=16.658$) than their colleagues in Experiment 8. This difference between the experiments did not vary as a function of any of the experimental factors as there were not reliable interactions between Experiment and other factors this time. However, an informal comparison between the name onset data in Experiments 8 and 9 reveals that the participants' speaking performance was associated with a 46 msec Prime advantage, a 52 msec Verb Match advantage, and a 21 msec Cue advantage.

7.2. Discussion

Experiments 8 and 9 addressed for the first time the issue of the visually mediated sentence production in a tripartite priming environment – the environment, in which the availability of perceptual, conceptual, and syntactic information changes as a function of unfolding discourse. This experimental setting is a better approximation to the real world situation. When people talk about the visually presented events, they do not only take account of the information available to them visually; they also rely on the previously produced or comprehended discourse in all its semantic and structural complexity.

Similarly to the previous studies, participants in both current experiments speakers tended to commit to the canonical structure more firmly when their perceptual analysis of the scene was "automatic" (e.g., it did not involve early fixating of all the available information except the preferred starting point – the agent). The latter result reinforces the conclusion about relative automaticity of event apprehension when the perceptual cue promotes an early commitment.

At the same time, speakers were highly sensitive to the structural priming manipulation in the visual world setting especially when the verbs in the prime and the target sentences matched. The latter effect was stronger when the structural prime was in the passive voice. These results replicate earlier findings by Pickering and Branigan (1998) and Branigan, Pickering, and Clelland (2000) about the importance of lexical overlap in establishing persistent structural priming. On the other hand, structural priming – a considerable regulator of the word order choice – did not exert noticeable influence during generation of names in Experiment 8, but it was persistent almost at every stage in Experiment 9, which suggests that the conceptual information provided with referent preview boosted the priming effect through a stronger association between the cued referent and the sentential starting point. A more detailed examination revealed that the stronger priming effect was observed when the cue was on the patient whereas speakers' performance in the agent-cued condition did not differ between the two studies. This result was reinforced by the main effect of the prime and a reliable interaction between the Cue and the Prime on the eye-voice span and the name onset data in Experiment 9.

Apart from aiding structural priming, the verb match effect demonstrated a considerable degree of independence. For example, analysis of the initial fixations to the non-cued referents revealed that early perceptual apprehension proceeded faster with speakers' rapidly switching attention between the referents when the verbs in the prime and the target sentences matched. When participants did not have access to verb information, they used the referential information provided by the cue in Experiment 9 for faster perceptual analysis. Also, speakers experienced less processing load at both sentence constituents (as revealed by the EVS analysis) and they were faster to produce constituent names when the verbs matched and when the cue was on the agent. It is likely that the verb preview pre-activates associated argument structures, which leads to a more rapid event analysis and higher probability of structural alternation. This interpretation received further justification from the analysis of the EVS values. Together, these results suggest that information carried by the verb influences referential mapping independently of structural priming.

Therefore, linguistic manipulations, such as structural priming and verb match, administered early in the trial started constraining processing already during early perceptual analysis by creating biases at later stages of analysis. The primacy of linguistic biases resulted in the establishment of temporary constraints, which led to shallower processing at stages preceding overt sentence generation. This was revealed by scarcity of the observed effects and interactions before overt sentence generation started. This processing pattern is predicted by the model proposed in the beginning of the thesis. This model, among other things, assumes that preferential biasing at the processing stages closer to the output (e.g., assembly) would impose constraints on the depth of processing within the preceding stages.

Establishment of these temporary (or local) constraints, as well as existence of permanent (or global) constraints (e.g., preferences for canonical causality and word order) did not completely override propagation of perceptual and referential priming. As the proposed model assumes, these constraints can be violable: If biases created at corresponding processing levels (e.g., perceptual analysis or event apprehension) are strong enough, they may noticeably survive through the production sequence regardless of existing or established constraints. As a result, perceptual effects in the current set of experiments were reliable, although smaller, than in the experiments that relied exclusively on perceptual manipulations. At the same time, effectiveness of perceptual priming relied on support at later processing stages, such as event apprehension and lemma retrieval. Confirming this, Experiments 8 and 9 demonstrated that perceptual priming was stronger when the verbs in the prime and the target sentences matched. By contrast, perceptual priming did not interact with structural priming, which suggests that perceptual priming effects can be propagated only to the lemma level, and they do not have a direct connection to sentence assembly. One alternative explanation of the persistence of perceptual priming in the current studies is the presence of the recency effect: The visual cue was presented immediately before the target picture presentation, which might ensure higher chances of survival. The latter explanation can only be tested by reversing the order of the priming manipulations.

Finally, EVS values observed in both current studies were noticeably shorter than in the previous experiments. This confirms my earlier prediction about reduction of the eye-voce span values as a result of the amount of information known about the upcoming event. The more information available, the shorter the eye-voice spans become, which suggests that they do not only reveal lexical access effects but also reflect how difficult it is to resolve the referential scheme of the event and to map this scheme onto the word order in a spoken sentence. Processing of the Subject- and the Object-related referents was similar in Experiment 8 although the EVS values for Subjects were slightly larger. This difference was more pronounced in Experiment 9 confirming that processing Subjects in English is associated with a larger operational load (cf. Experiments 3, 6, and 7). Processing of both Subject- and Object-related referents was similarly difficult when the cue was on the agent; when the cue was on the patient, processing both referents was easier with the passive voice prime. This result provides evidence for the connection between the distribution of the referential roles in the event map and the constituent roles in the of sentence frame.

Chapter 8. Thesis conclusions

The experiments discussed in this thesis examined the nature of the interplay between linguistic, conceptual, and perceptual processes during production of sentences. The research tested speakers' performance in different languages using different priming techniques in combination with the Visual World paradigm. It investigated the necessary requirements for the establishment of perceptually driven word order selection in spoken discourse, the dependence of this selection on the availability of conceptual information about the event's referents, and the interaction of both perceptual and conceptual accessibility with structural processing. A variety of behavioral measures provided a fine-grained resolution of the processing operations at different stages of sentence production. The experimental data support the proposed model of the dynamic integration of the available perceptual, conceptual, and structural information at various stages of visually mediated sentence production.

The model described in detail in Chapter 4 of this thesis made specific predictions about interface between perceptual, conceptual, and linguistic processes during production of sentences in different languages. It may be helpful to reiterate the most relevant parameters of the model in order to relate the results of the experimental studies to predictions from the model.

The general architecture of the model assumes an operational division between the processing stages or levels: message, lemma, and assembly. Processing at the message level includes perceptual analysis and event apprehension[6]. Processing at lemma

[6] It is not certain at this point whether perceptual analysis should be separated from the message level as the level preceding it. Some of the data discussed here suggest that perceptual analysis and event apprehension can be virtually the same thing in that early detection of one of the referent can cancel full perceptual

level includes selection and retrieval of word lemmas. Processing at assembly level involves ordering the lemmas in the sentential plan. Informational inputs to the processing stages derive from either *global* or *local* levels. Global inputs originate from production-related preferences and automated routines stored in the long-term memory. They operate as a set of regular or global *constraints* that are treated by the processor as defaults. Inputs at local level operate as temporarily active *primes*. Primes are established at a corresponding level and are propagated from the lower (or the earlier) levels of processing to the higher (or the later) levels of processing. Constraints are established at each corresponding level and propagate control the other way around – from the higher levels over the lower levels. For example, constraints at assembly level can control the informational input from levels preceding so that this information is entered in the form best fitting the existing constraint. Priming parameters can override constraining parameters. Priming at each level can be affected by propagation of priming from the preceding level. Constraints at lemma and assembly level may vary from language to language so that some morphological and syntactic properties of one language may be more constraining than of another. The *types of information* the system may use to arrive at the alternating word order are perceptual, conceptual, lexical, and structural entered at corresponding processing levels. Each processing level is only connected to the level immediately adjacent to it; therefore processing interactions can only occur between the adjacent levels. Finally, primes or constraints operating at levels closer to the output have a higher chance of influencing the nature of the final product – the spoken sentence. The

---

analysis of the rest of the scene. In other cases, full perceptual analysis may be necessary before referential analysis takes over. For the sake of clarity I will for the time being continue to position both perceptual and conceptual processing within the same level – message.

following paragraphs will assess how the experimental data reported above fits with the proposed model.

8.1. Accommodation of perceptual effects

One simple lesson from the research reported in this thesis is that attention influences word order. It does so by biasing the processor to insert the perceptually more salient or intentionally attended referent as the starting point of a spoken sentence. In the case of simple transitive sentence production this means that when the agent of an event is attended to, the speaker tends to place it at the beginning of an upcoming sentence; if the speaker preferentially attends to the patient, it is the patient that tends to act as the sentence's starting point. However, this tendency is not deterministic: While cueing the agent almost inevitably predicts its early placement in a sentence, cueing the patient only increases the chances of its early placement by around 10-30%. The production of novel sentences without support from the preceding discourse relies heavily upon the structures speakers are most familiar with regardless of the perceptual properties of the described event.

What factors influence the persistence of perceptual effects during sentence production? The first group of factors belongs to the locally established discourse environment. The second has to do with the global constraints on the accommodation of the perceptual biases.

8.1.1. Local regulators of perceptual priming

First, the current studies demonstrate that manipulations of attentional focus have to be explicit and relatively strong in order to induce more or less consistent alternation of the word order choice. Implicit manipulations of attention (e.g., Gleitman, et al., in press) only lead to a minimal increase in the use of non-canonical structures, such as patient-driven passive voice in English. Experiments reported here demonstrate that word order alternations over 50% are only possible with an extremely powerful and visually restricting cueing procedure. At the same time, two experiments using Finnish – a language with flexible word order and noun case marking – failed to find any effect of implicitly directed attentional focus on sentence production.

The second factor influencing the promotion of the perceptual effects has to do with the fact that the perceptual information becomes available to the speaker at the earliest stages of analysis. Given that after perceptual analysis, sentence production involves a series of processing stages including conceptual analysis, lexical selection, and structural assembly, it is not surprising that the early biases have to survive competition from biases created at the subsequent processing stages. Powerful constraints of canonicality and the necessity to compete with other factors throughout the course of sentence production make perceptual effects on the speakers' structural choices relatively weak

The third important limitation on perceptually driven word order selection concerns the range of structures that support it. Production of structures that are conceptually and structurally relatively simple (this thesis) or those that do not require structural alternation (Forrest, 1997; Gleitman, et al., in press) make it more likely that perceptual effects will influence the assignment of sentential slots. Inclusion of one more referent (Experiment 5) greatly reduces these chances. Although implicit cueing of attention did not result in any structural alternation during Finnish transitive sentence production, its effects were detectable as attentional shifts during event apprehension. This effect was completely absent when Finnish speakers produced ditransitive sentences. With increased referential complexity, perceptual biases are swamped by the more complex conceptual map. Speakers ignore early perceptual biases as possible determiners of their analysis when they have to track and process more than two referents. Presumably, it saves them cognitive effort and facilitates automatic sentence formulation. Whether the same result would occur if the manipulation of attention was explicit is a good question. Using the already established logic, explicit manipulations of attention should increase the resilience of perceptual effects in ditransitive sentence production, probably to a lesser degree than in transitive sentence production. Also, it would be interesting to try to replicate the same result using English. Following from comparisons between English and Russian data (Experiments 1-3) one would expect to observe stronger perceptual effects in English ditransitive sentence production compared to the ones observed in Finnish.

The fourth requirement for the successful accommodation of the perceptual effects is support at later stages of processing. For example, if perceptual promotion of the patient is paired with the early priming of the verb, the chances of the patient assuming the starting point become higher (see in detail below). Hypothetically, the strongest lexical support for perceptual priming would be through lexical priming of the referent. However, this support is limited to the interaction between adjacent processing stages. For example, structural priming has an effect separate from the perceptual priming and does not aid the latter.

8.1.2. Global regulators of perceptual priming

8.1.2.1. Canonicality

One feature limiting the scope of perceptual effects is canonicality in event apprehension and word order grammar. In all the three languages examined the canonical word order is agent-driven SVO. It follows the preferred agent-action-patient event apprehension consistently mapping the agent onto the starting point in a subsequent sentence. Analysis of the initial fixations and eye-voice spans clearly showed that speakers prefer to rapidly detect and highlight the initiator of the action subsequently making it the sentential starting point, resolve the remainder of the referential scheme, and finally map the emerging conceptual plan onto the corresponding syntactic structure. Usually, the attentional focus quickly shifted from the cued patient to the non-cued agent – the referent typically occurring in the frontal position of the canonical sentence. Therefore the preferred event apprehension was almost as automatic as the selection of the canonical sentence frame. When it occurred, processing of non-prominent elements of the scene was optional and shallow as speakers rarely fixated non-cued patients when producing the agent-initial SVOs. When a non-preferred passive was selected, early analysis of the event was deeper, as speakers tended to rapidly fixate both available referents.

This result confirms one prediction about the performance of the model proposed in this thesis – that the global constraints are generally more powerful than the priming parameters. The tight coupling between the preferred event interpretation and the canonical word order creates a powerful set of processing defaults or constraints that have to be overridden when a non-preferred starting point is perceptually promoted. One could say that speakers are quite “lazy”: Unless there is a good reason to actively engage in the event’s analysis and map the resulting structure accordingly, speakers production behavior tends to be quite automated both during non-linguistic and linguistic stages of sentence formulation (cf. Garrod & Pickering, 2006). The information contradicting early commitment to the automated conception, lexicon, and/or grammar has to be considered in the context of automaticity and facilitated processing.

8.1.2.2. The effects of the language grammar

The second limitation on the promotion of perceptual effects has to do with the details of the language grammar. The grammar of language needs to “favor” or maximally allow such propagation. Dependence on canonical grammar was much more pronounced in case-marking languages with flexible word order systems: Seeming flexibility of choice led to narrower flexibility of performance. English – a language that relies on a relatively rigid system of word order and does not use case marking to denote referential roles – was more likely to accommodate perceptual effects in the word order (see also Tomlin, 1995; Gleitman, et al., in press). The lack of case marking on the nouns in English allows maintenance of early commitments to the starting point’s word form with alternating syntax. This option is usually not available to speakers of Russian and Finnish because the commitment to the starting point’s word form largely predetermines the referent’s grammatical function in the sentential frame. This means that changing the initial mapping requires repairing the whole sentence when the early commitment is not justified.[7] The data reported in this thesis demonstrate that *grammatical flexibility* has to be differentiated from *processing flexibility*. Theoretically, grammatical flexibility – the inventory of felicitous structures – can promote swift choices between the structural alternatives (Ferreira, 1997), but this is only possible when it is coupled with the maintenance of processing flexibility – an ability to alternate the produced structure beyond the starting point. Early commitments to the case-marked forms restrict processing flexibility biasing the speaker to the corresponding grammatical role early on. In English, processing flexibility is retained until the verb is selected. The eye-voice span data reported in this thesis confirm that the production of English transitive sentences follows a more flexible route until the verb is selected; the production of Russian transitive sentences is associated with a higher processing load early on because the selection of the starting point entails the commitment to the word form – typically the nominative case corresponding to the Subject.

All these features fit well with model proposed in Chapter 4 of this thesis. The first important restriction on the propagation of perceptual effects is that they can only be accommodated at the very first steps of event analysis. This predicts competition with factors important at the later stages of production. This competition starts already with the perceptual analysis. The distribution of the earliest looks in the scene clearly showed that speakers prefer to locate the preferred starting point – the agent as soon as they can regardless of where their attention has been previously directed. To prevent these early

[7] An attentive reader might notice that such cases can be found in Russian as well (see discussion of Jackobson's observations in Chapter 5). However, the fact that Russian allows a wide use of diminutives (which are always case-marked) makes Russian less than perfect candidate for the proposed analysis. Same is true about Finnish.

shifts of attention from the cued to the preferred starting point one needs to use quite powerful perceptual manipulations. The second model feature that predicts the observed pattern of results is the reliance of the processor on global and local constraints. Preferred event apprehension and canonical grammar are examples of global constraints, while effects of referential, lexical, or syntactic priming impose temporary or local constraints. These constraints have to be overridden in order for the perceptual effects to continue their influence to the point of the overt sentence generation. Third, the power of constraints may differ as a factor of the language: If lexical access in Russian comes with the early commitment to the grammatical role in a sentence, this constraint will be more difficult to override. Finally, the model predicts limited scope for the processing interactions between stages limiting them to the neighboring pairs. This prediction is supported by the lack of interactions between perceptual and structural priming, on one hand, and interactions between structural and lexical priming, on the other.

8.1.3. Extraction of Event information

An issue that needs to be discussed separately here is the importance of the looks to the *event area* during apprehension of events and formulation of sentences. It is difficult to establish which specific area of the display contains the information later conceptualized in the verb. On one hand, event information can be extracted from the areas containing referents typically adjacent to the verb phrase in the resulting sentence, for example, the instrument of the event. On the other, the event information can be more or less equally distributed among the referents making it impossible to suggest any specific area as related to the event *per se.* I initially used the logic for coding the event area suggested in previous studies (Griffin & Bock, 2000) treating the area around instrument as where event information is extracted. This logic proved to be successful. For example, experiments 2 and 3 did not provide participants with any information about the event prior to the target picture presentation. This resulted in persistent looks to what I decided to treat as the event area from trial to trial. The protocol used in Experiments 4-7 let participants preview the verb related to the subsequently described event. As a result, (1) there were substantially fewer looks to the same area when the target picture appeared on the screen and (2) such looks were practically absent when the attentional cue was on the agent, which almost inevitably led to the production of active voice SVO. These results justify, at least circumstantially, coding of the visual area around the instrument as the one related to the extraction of the event information later conceptualized in the sentential verb.

8.2. Accommodation of conceptual effects

As we already established, longer exposure to a cue associated with the *explicit* cueing of attention together with restricting the attentional focus to the cued referent leads to a better accommodation of perceptual priming. Some theories predict that referent preview should further increase depth of processing and result in the establishment of the given/new contrast subsequently leading to even stronger structural alternation. The experiments of the current thesis (almost) failed to support such a prediction. Speakers did not produce more non-canonical structures when the cue was referential than when it was perceptual in a predominantly perceptually-manipulated environment (Experiments 6 and 7). Analysis of the data in Experiments 8 and 9, in which both lexical and structural priming were involved, provided some counterevidence: speakers were 5% more likely to alternate word order when the cue was referential. For the most part, these results suggest that the given/new status known to predict the choice of word order is not boosted with the help of providing referential information. In other words, referential priming of word order works no better than perceptual priming. One conclusion is that the given/new effect might be enhanced with the help of lexical priming of the referents. Early results from the ongoing research (Myachykov, Garrod, & Coquillon, 2007) support this conclusion: Speakers are much more likely to actively alternate word order when perceptual priming is coupled with the lexical priming of the referents' names.

Although increasing conceptual accessibility without lexical support did not improve the progression of effects from perception to syntax it led to processing facilitation at stages preceding lexical access and during the formulation of sentences. For example, referent preview consistently sped up perceptual analysis of the event and the rate of overt sentence production. It also reduced the processing load reflected in the eye-voice span values. Together, these results suggest that the frame of the upcoming sentence is assigned early on, in accordance with the perspective derived from the event's causality. Conceptual and lexical computations that happen afterwards are more capable of facilitating the speed of processing, but they do not affect the prior referential frame analysis.

Finally, early availability of conceptual information about the referents enhanced the structural priming effect. A detailed examination of the data in Experiments 8 and 9 revealed that structural priming was stronger when the cue was referential. The referential boost of the priming effect was observed when the cue was on the patient whereas speakers' performance in the agent-cued condition did not differ between the two studies. This result suggests that priming of non-canonical passives may benefit from the availability of the referential information about the corresponding starting point – the patient.

These conceptual effects support the processing model proposed in this thesis. Although one might expect active lexicalization of visually presented material even when speakers do not have to name this material, this need not be the case. In fact, some studies suggest that when participants are instructed to silently examine pictures, they tend to avoid linguistic analysis of the scene (Griffin & Bock, 2000). Combined with the current studies, these results suggest that perceptual and conceptual analysis constitute a single processing stage, which is separated from lexical access. On the other hand, minimal yet reliable increase in the production of non-canonical passives in Experiment 9 suggests that increase in referential accessibility can promote preferential treatment of the primed referent at the stages following the perceptual and conceptual analyses.

8.3. Accommodation of structural effects

The contribution of structural information to production of sentences is well-studied. However, the effects of structural priming are typically studied without corresponding manipulations at the levels of perceptual and conceptual analysis. I tried to examine how structural priming effects interact with the effects of perceptual, referential, and lexical priming.

The first important conclusion supported both in studies using structural priming and those that did not is that active voice sentences are produced faster regardless of the nature of priming environment. For example, comparisons of the name onset latencies in Experiment 1 revealed that Russian speakers were faster to initiate the canonical SVO even when it contradicted the perceptual priming scheme. The same was true in Experiments using English; active voice sentences produced in conditions that strongly favored the production of passive voice were still produced faster than the alternative structure. This result confirms my earlier conclusion about the power of the canonicality constraint and that the canonical structures are generated in effortless and automatic manner.

The structural priming effect was replicated both in studies that used this paradigm: Speakers were more likely to produce non-canonical passives when they were primed by another sentence with the passive voice. This effect was substantially strengthened when the verbs between the prime and the target sentences matched. This finding replicates earlier findings about the importance of lexical overlap for the establishment of strong structural priming (Pickering & Branigan, 1998; Branigan, Pickering, & Clelland, 2000). A novel finding was that the verb match effect showed a considerable degree of independence. For example, analysis of the initial fixations to the non-cued referents revealed that early perceptual apprehension proceeded faster with speakers' rapidly switching attention between the referents when the verbs in the prime and the target sentences matched. When participants did not have access to verb information, they used the referential information provided by the cue for faster perceptual analysis. Also, speakers experienced less processing load at both sentence constituents (as revealed by the EVS analysis) and they were faster to produce constituent names when the verbs matched and when the cue was on the agent. It is likely that the verb preview pre-activates associated argument structures, which leads to a more rapid event analysis and higher probability of structural alternation. Together, these results suggest that information carried by the verb influences referential mapping independently of structural priming.

With regard to perceptual priming, the experiments reported here demonstrated that structural priming produces an effect independent of manipulations at the perceptual level: Neither of the two experiments revealed interactions between these two priming components. This result reinforces previous conclusions about the separation of perceptual and structural processing levels and supports the model's prediction about the localization of the interactions to the adjacent stages of processing. The effect of structural priming was also stronger when conceptual information about the cued referent was revealed to the speakers. This effect, discussed above, suggests some degree of accommodation of the conceptual information during structural preparation. Compared to the lack of the referential priming effect on the choice of word order in the experiments that did not use structural priming, this result means that referential priming promotes preferential assignment of the grammatical roles only when the structural frame of the upcoming sentence is also primed.

Providing syntactic information before the presentation of the target event leads to the establishment of powerful constraints on perceptual and conceptual analyses even at early stages of event analysis. This result also fits the model quite well. One of the model's features is that priming within the system occurs at corresponding levels, and the closer the level to the output the more likely the priming within the level will be successful. Because sentence assembly is practically the last stage of sentence formulation, structural priming is quite capable of attenuating all preceding effects and biasing the resulting sentence in accordance with the primed structure.

8.4. Eye-voice spans

One measurement that is not frequently used in psycholinguistic research is eye-voice span. In the Visual World studies, it represents the interval between the onset of the fixation to the referent immediately preceding the production of the referent's name and the onset of the corresponding name in a sentence. As such, it spans conceptual, lexical, and structural analyses and reflects generalized processing difficulty associated with relating the detected referent to its name and grammatical role. Based on this definition, EVS is different from another measurement used to analyze processing difficulty – latencies to produce constituent names. Although both measures are representative of the processing load, the EVS usually excludes the initial rapid apprehension of the event and, therefore reflects operations that a relatively less automatic.

This measurement proved to be extremely important for the purposes of my research. For example, it made it possible to distinguish between grammatical and processing flexibility associated with selection of word order. It also allowed separately analysis of the processing difficulty associated with the production of different sentence constituents.

In studies that use novel pictures and do not provide much information about the upcoming event, the magnitude of observed EVS tend to be similar to the times necessary for lexical access, which led to the initial conclusion that EVS represents lexical access. The series of experiments reported in this thesis show that EVS values tend to become shorter as more conceptual, lexical, and structural information about the event is revealed (see Table 76). This supports my proposal about the properties of this measure: EVS reveals general cognitive load associated with mapping of the referents onto their places in a sentence. On one hand, EVS spans a number of operations; on the other, with support of priming at different levels of processing, it takes on values much shorter than those observed when no or minimal priming is used. There may be two different explanations for this property of EVS. One explanation would be that processing at different levels may be partially parallel. According to this scenario, the simultaneous processing boost at a number of levels through priming might lead to the substantial decrease in the EVS magnitude. Hence, EVS do not represent the chronometric sum of sequential operations but reflect facilitation through parallel activations at different levels. The alternative explanation is that shorter EVS reflect a sum of facilitation effects at a number of levels of processing: As processing becomes faster at each of the EVS components, the EVS values become shorter. This interpretation would relate shorter EVS to the sequential facilitation effect (cf. Vigliocco & Hartsuiker, 2002).

Comparison of the EVS values for Subjects and Objects across the languages revealed a number of interesting observations. First, the tendency for the EVS values for Subject to be longer than those for the Object was confirmed in three experiments using free word order languages (Russian and Finnish). The EVS values in Experiments 4 and 5 with Finnish were much shorter than those reported in Experiment 3 using Russian due to the tendency of the EVS to become shorter as a factor of information access. However, the general tendency for a greater difficulty in processing the initial sentence constituent in two different languages with similar case marking properties reinforces my assumption about early processing inflexibility involved in the production of sentences in the languages that use case marking to denote grammatical roles in a sentence. Comparing EVS values across the English studies reveals that in Experiments 2 and 6 the EVS values for Object were longer than those for the Subject. This pattern reversed in Experiments 7-9. The first two studies provided minimal information about the upcoming event. This was not the case in the second group of studies. The results from Experiment 2 (possibly replicated in Experiment 6) suggest that a larger degree of processing flexibility during construction of English sentences is responsible for the fact that it is easier to process English Subjects than Objects when alternation can be entertained at least until the verb has been accessed. If there is more information available to speakers before sentence generation starts (Experiments 7-9), the production system is probably biased earlier to commit to one of the available structural options. This fact is reflected in the reversed pattern of EVS. This explanation is supported by a much shallower processing at early stages in Experiments 8 and 9: the imposition of procedural constraints through structural priming biases the processor to make the structural decision earlier, regardless of whether the decision follows or contradicts the prime.

8.5. Directions for further research

The research reported in this thesis and the model supported by the experimental data open many opportunities for further research. The current experiments provide compelling evidence about speakers' performance in the continuously changing sentence production environment. Some questions that I posed in the beginning of the thesis as well as new question prompted by the experimental results remain unanswered.

The most obvious route for further research is to compensate for the lack of data from languages like Russian and Finnish using multiple priming paradigms (Experiments 6-9). It is not clear that the priming parameters used in studies with English would interact with grammatical systems associated with more combinatorial freedom and active case marking.

Another direction is suggested by the results of Experiments 2 and 3. Among other things, these results prompted an important conclusion about the dependence of processing flexibility on the details of the language grammar, particularly on case marking. In short, I suggested that early commitments to the word form and the corresponding grammatical function in Russian reduces processing flexibility although the speaker's access to a wider word order inventory should facilitate processing. One way to test this assumption would be to conduct an experiment using German. Production of German transitive sentences sometimes is associated with distinctive Nominative and Accusative case marking for the Subjects and the Objects. At other times, marking these case forms is ambiguous; in this situation, the verb information is used to disambiguate the grammatical roles of the referents. To separate the effects of case marking from the effects of word order flexibility, one would need to compare production of the German case-sensitive sentences with the case-ambiguous ones with the verbs that promote word order flexibility (e.g., English *give*) and restrict it (e.g., English *donate*).

The third direction for further research is to continue testing visually-mediated sentence production by using languages with different rules of canonicality. It is possible that perceptually driven sentence production in Object- or Verb-frontal languages is regulated by parameters very different from those determined for the SVO languages tested in this thesis.

One more important alley is to use different attentional tasks in order to arrive at a finer-grained picture of attentional contributions during sentence production. Previous research in attention established the existence of three separate networks of attention: alerting, orienting, and conflict resolution (see Chapter 2). It is possible that there are different linguistic devices responsible for accommodating the separate contributions from these networks. This possibility also prompts testing other syntactic structures than transitive and ditransitive sentences analyzed in this thesis. For example, syntactic phenomena like clefts and dislocations may be responsible for orienting listener's attention within the unfolding discourse and/or repairing mistakenly assigned attentional foci. Also, such a phenomenon as linguistic focus typically interpreted as an attentional device has never been clearly linked to the attentional operations in an experimental study. To do so, one might need to independently manipulate attentional and linguistic focus in the same experimental setting looking for possible correlations between their effects.

There are three more ambitious directions to the research described here. The first direction has to do with the development of the interface between perception and language (also Chapter 2). Informal observations across the age groups during the collection of data and very limited pilot work with children suggest that children tend to rely upon attentional focus much more when constructing their sentences. This observation is supported by existing research on the importance of joint attention between the child and the adult for the development of the language facility. So, the link between attention to objects and their order of mention may be tighter during the early stages of language development. On the other hand, production performance of senior adults demonstrates higher reliance on the canonical grammar rather than attentional focus with the performance of young adults typically used in psycholinguistic studies falling somewhere in between. The deterioration of the link between attention and word order in senior years may result from both higher reliance on the long-term memory (automated syntax) and poorer ability of senior people to switch attention between stimuli. There is, therefore, a reason to hypothesise a complex function of the link between attention and word order changes across ages with the tightest link in childhood, moderate bias to automaticity rather than attentional focus in early adulthood and increasing of such bias in senior adults.

The same logic prompts another direction: Examining the language-perception interface in grammatical aphasia. It is possible that impaired attentional performance in grammatical aphasics (Cairns, et al., 2007; Stemmer, 1999) may show up in the less efficient mapping of the referential map onto the sentential frame. Finally, although there is an abundant literature about the development and the adult performance of the bilingual speakers, the relation between bilingual language processing and the attentional control has not been extensively studied (but see studies discussed in Chapter 2). On one hand, there is evidence for better performance of bilinguals on attentional tasks, which suggests that maintenance of two languages in mind may rely upon facilitated ability for switching attention from one language to the other. On the other hand, I would be interested to examine the reverse link – from attention to language – in bilinguals by administering the same perceptual priming task to elicit responses in two different languages from the same speaker.

These are only a small number of possible routes to further explore the functioning of the interface between perceptual and linguistic processing. The experimental studies of the sort described here are not numerous, which opens up possibilities for decades of research. I believe the research summarized in this thesis provides better understanding of the identified research questions and encourages better and deeper explorations within this research agenda.

References

Abdullaev, Y.G. & Posner, M.I. (1998). Event-related brain potential imaging of semantic encoding during processing single words. *Neuroimage, 7,* 1-13.

Adamec, P. (1966). *Porjadok slov v sovremennom russkom jazyke*, Prague: Academia.

Alario, F.-X., Costa, A., Ferreira, V.S., & Pickering, M.J. (2006). Architectures, representations, and processes of language production. *Language and Cognitive Processes, 21(7-8),* 777-789.

Altman, L.J.P. & Kemper, S. (2006). Effects of age, animacy and activation order on sentence production. *Language and Cognitive Processes, 21, 322-354.*

Altmann, G.T.M. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition, 93,* B79-B87.

Altmann, G.T.M. & Kamide Y. (2004). Now you see it, now you don't: mediating the mapping between language and the visual world. In J. Henderson & F. Ferreira, (Eds.) *The Integration of Language, Vision, And Action: Eye Movements And the Visual World* (pp. 311-347). New York: Psychology Press.

Altmann, G.T.M. & Kamide, Y. (in press). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language.*

Arai, M., Van Gompel, R.P.G., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology, 54,* 218-250.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics, 24,* 65–87.

Ariel, M. (2001). Accessibility Theory: An overview. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 29–87). Amsterdam: Benjamins.

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

Arnold, J., T. Wasow, A. Losongco, & Ginstrom, R. (2000). Heaviness vs. newness: The effects of complexity and information structure on constituent ordering. *Language*, *76,* 28–55.

Baldwin, D.A. (1995). Understanding the link between joint attention and language. In C. Moore and P.J. Dunham (Eds.), *Joint Attention: Its Origins And Role in Development* (pp. 131-158)*,* Hillsdale, NJ: Lawrence Erlbaum Associates.

Ballard, D., Hayhoe, M.M. & Pook, P.K. (1997). Deictic codes for the embodiment of cognition, *Behavioral and Brain Sciences, 20,* 723-742.

Barsalou, L.W. (1999). Perceptual symbol systems, *Behavioral and Brain Sciences, 22,* 577-660.

Bates, E., & Devescovi, A. (1989). Competition and sentence production. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 225-256). New York: Cambridge University Press.

Bates, E. & MacWhinney,B. (1982). Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173–218), Cambridge: Cambridge University Press.

Bates, E., & MacWhinney, B. (1987). Competition, variation and language learning. In B. MacWhinney, & E. Bates (Eds.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York: Cambridge University Press.

Bialystok, E., Craik, F. I. M., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effects of bilingualism on cognitive control in the Simon task: evidence from MEG. *Neuroimage, 24,* 40-49.

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language, 48,* 468–488.

Bock, J.K. (1977). The effect of pragmatic presupposition on syntactic structure in question answering. *Journal of Verbal Learning and Verbal Behavior, 16(6),* 723-734.

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

Bock, J.K. (1986a). Meaning, Sound, and Syntax: Lexical Priming in Sentence Production. *Journal of Experimental Psychology, 12(4),* 575-586.

Bock, J.K. (1986b). Syntactic persistence in language production. *Cognitive Psychology*, *18,* 355-387.

Bock, J.K. (1989). Closed-class immanence in sentence production. *Cognition, 31,* 163-186.

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

Bock, J.K. & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology General*, *129,* 177-192.

Bock, J.K. & Huitema, J. (1999). Language production. In S. Garrod & M. Pickering (Eds.), *Language Processing* (pp. 366-388). Hove, England: Psychology Press.

Bock, J.K. & Irwin, D.E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior, 19,* 467-484.

Bock, J.K., Irwin, D.E., & Davidson, D.J. (2004). Putting first things first. In J. Henderson & F. Ferreira, (Eds.) *The Integration of Language, Vision, And Action: Eye Movements And the Visual World* (pp. 249-278). New York: Psychology Press.

Bock, J.K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language, 48,* 653-685.

Bock, J.K. & Levelt, W.J.M. (1994). Language production: grammatical encoding. In M.Gernsbacher (Ed). *Handbook of Psycholinguistics* (pp. 945-984), New York: Academic Press.

Bock, J.K. & Loebell, H. (1990). Framing sentences. *Cognition, 35,* 1–39.

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.

Branigan, H.P. (2007). Syntactic priming. *Language and Linguistics Compass, 1(1-2),* 1-16.

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

Branigan, H.P., Pickering, M.J., & McLean, J.F. (2005). Priming prepositional-phrase attachment during comprehension. *Journal of Experimental Psychology: Learning, Memory & Cognition, 31,* 468-481.

Branigan, H.P., Pickering, M.J., McLean, J.F., & Stewart, A.J. (2006). The role of local and global syntactic structure in language production: Evidence from syntactic priming. *Language and Cognitive Processes, 21*, 974–1010.

Daneš, F. (1970). One Instance of Prague School Methodology: functional analysis of utterance and text. In P. L. Garvin (Ed.). *Method and Theory in Linguistics*. Paris:Mouton & Co.

Cairns, D., Marshall, J., Cairns, P., and Dipper. L. (2007). Event processing through naming: Investigating event focus in two people with aphasia. *Language and Cognitive Processes, 22(2),* 201-233.

Carpenter, M., Nagell, K. & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society of Research in Child Development, 63(4),* 1–143.

Caron, A.J., Kiel, E.J., Dayton, M. & Butler, S. C. (2002). Comprehension of the referential intent of looking and pointing between 12 and 15 months. *Journal of Cognition & Development, 3,* 445–64.

Carroll, J.B. (1958). Process and content in psycholinguistics. In R.A. Patton (Ed.) *Current trends in the description and analysis of behavior* (pp. 175-200). Pittsburgh: University of Pittsburgh Press.

Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. *Cognitive Science, 26,* 609-651.

Chang, F., Bock, K.J., and Goldberg, A.E. (2003). Can thematic roles leave traces of their places? *Cognition, 90,* 29-49.

Chang, F., Dell, G.S., Bock, J.K., & Griffin, Z.M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research, 29(2),* 217-229.

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

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on Government and Binding.* Dordrecht: Foris.

Christianson, K. & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition, 98,* 105-135.

Chafe, W.L. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View. In C.N. Li (Ed.), *Subject and Topic*, (pp. 25-55). New York: Academic Press.

Chafe, W.L. (1994). *Discourse, Consciousness, and Time.* Chicago & London: University of Chicago Press.

Chun, M.M., & Wolfe, J.M. (2001). Visual Attention. In B. Goldstein (Ed.), *Blackwell Handbook of Perception* (pp. 272-310). Oxford, UK: Blackwell Publishers Ltd.

Clark, H.H. (1965). Some structural properties of simple active and passive sentences. *Journal of Verbal Learning and Verbal Behavior, 4,* 365-370.

Clark, H.H. (1966). The prediction of recall patterns in simple active sentences. *Journal of Verbal Learning and Verbal Behavior, 5(2),* 99–106.

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

Clark, H.H. & Chase, W.G. (1974). Perceptual coding strategies in the formation and verification of descriptions. *Memory and Cognition, 2(1-A),* 101–111.

Clark, H. H. & Haviland, S. E. (1977). Comprehension and the Given-New contract. In R. O. Freedle (Ed.), *Discourse Production and Comprehension* (pp. 1–40). Norwood, NJ: Ablex.

Cleland, A. A. & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: evidence from the priming of noun phrase structure. *Journal of Memory and Language, 49,* 214–230.

Corbetta M. (1998) Frontoparietal Cortical Networks for Directing Attention and the Eye to Visual Locations: Identical, Independent, or Overlapping Neural Systems? *Proceedings of the National Academy of Sciences, 95(3)*, 831-838.

Corbetta, M., Kincade, J.M., & Shulman, G.L. (2002). Neural systems for visual orienting and their relationship to spatial working memory. *Journal of Cognitive Neuroscience, 14,* 508-523.

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

Cottrell, G. W. (1989). A Connectionist Approach to Word Sense Disambiguation. *Research Notes in Artificial Intelligence.* San Mateo: Morgan Kaufmann.

Dehaene, S. & Changeux J.P. (2004). Neural mechanisms for access to consciousness. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences III* (pp. 1145-1158). Cambridge, MA: MIT Press.

Dehaene, S., Kerszberg, M. & Changeux, J.P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences USA, 95,* 14529-14534.

Dehaene, S. & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition 79,* 1-37.

Dell, G.S. (1986). A spreading activation theory in sentence production. *Psychological Review, 93(3),* 283-321.

Dell, G.S., Chang, F., & Griffin, Z.M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science, 23(4),* 517-542.

Desmet, T. & Declercq, M. (2006). Cross-linguistic priming of syntactic hierarchical configuration information. *Journal of Memory and Language, 54,* 610-632.

Diderichsen, P. (2001) Selective attention in the development of the passive construction: a study of language acquisition in Danish children. In E. Engberg-Pedersen & P. Harder (Eds.), *Ikonicitet og struktur. Netvork for Funktionel Lingvistik.* Department of English: University of Copenhagen.

Dobel , C. & Gumnior, H. (2004). Naming without seeing. *Poster presented at 10th AMLAP conference*, Université de Provence.

Dominey, P.F. & Dodane, C. (2004). Indeterminacy in language acquisition: the role of child-directed speech and joint attention. *Journal of Neurolinguistics, 17,* 121-145.

Elman, J. (1990). Finding structure in time. *Cognitive Science, 14*, 179–211.

Elman, J. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning, 7,* 195-225.

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

Fan, J., McCandliss, B.D., Sommer, T., Raz, M., & Posner, M.I. (2002). Testing the efficiency and independence of attentional network. *Journal of Cognitive Neuroscience, 3,* 14 340-347.

Faroqi-Shah, Y. & Thompson, C.K. (2003). Effect of lexical cues on the production of active and passive sentences in Broca's and Wernicke's aphasia. *Brain and Language, 85,* 409-426.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33,* 715-736.

Ferreira, F. (2005). Psycholinguistics, formal grammars, and cognitive science. *The Linguistic Review, 22,* 365-380.

Ferreira, F. & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348-368.

Ferreira, V.S. (1997). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35*, 724-755.

Ferreira, V.S. (2003). The persistence of optional complementizer production: Why saying "that" is not saying "that" at all. *Journal of Memory and Language, 48,* 379-398.

Ferreira, V.S. & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes, 21(7-8),* 1011-1029.

Ferreira, V., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Human Learning and Memory, 28,* 1187–1199.

Ferreira, V.S. & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of psycholinguistic research, 32(6),* 669-692.

Fillmore, C. (1968). The case for case. In E. Bach & R. Harms (Eds.) *Universals in Linguistic Theory* (pp 1-90). New York, NY: Holt, Rinehart, & Winston.

Firbas, J. (1965). A note on transition proper in functional sentence analysis. *Philologica Pragensia, 8,* 170-176.

Fischer, B. (1998). Attention in saccades. In R. D. Wright (Ed.), *Visual attention* (pp. 289-305). New York, NY: Oxford University Press.

Flores d'Arcais, G.B. (1975). Some perceptual determinants of sentence construction. In G. . Flores d'Arcais (Ed.), *Studies in perception. Festschrift for Fabio Metelli* (pp. 344-373). Milan, Italy: Martello- Guinti.

Fodor, J.A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Forrest, L.B. (1996). Discourse goals and attentional processes in sentence production: the dynamic construal of events. In A.E. Goldberg (Ed.) *Conceptual Structure, Discourse and Language* (pp. 149-162). Stanford, CA: CSLI Publications.

Frazier, L. (1995). Constraint satisfaction as a theory of sentence processing. *Journal of Psycholinguistic Research, 24(6),* 437-468.

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

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

Garrod, S. & Pickering, M.J. (1999). Issues in Language Processing. In S. Garrod & M.J. Pickering (Eds.) *Language Processing* (pp. 1-11). Hove, England: Psychology Press.

Garrod, S. & Pickering, M.J. (2006). Automaticity of language production in monologue and dialogue. In A. Meyer, L. Wheeldon, & A. Krott (Eds.), *Automaticity and control in language processing* (pp. 1-20). Taylor and Francis.

Garson, J. (2007). Connectionism. In E.N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy.* URL: <http://plato.stanford.edu/archives/spr2007/entries/connectionism/>.

Gasser, M.E. (1988). *A connectionist model of sentence generation in a first and second language.* Unpublished doctoral dissertation, Computer Science Department, University of California, Los Angeles, CA.

Gennari, S.P., Mirkovic, J., & MacDonald, M. (2005). The role of animacy in relative clause production. *Paper presented at the 18th Annual CUNY Conference on Human Sentence Processing*.

Gershkoff-Stowe, L. & Goldin-Medow, S. (2002). Is there a natural order for expressing semantic relations? *Cognitive Psychology, 45,* 375-412.

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

Glanemann, R., Zwitserlood, P., Bölte, J., Kreysa, H., & Dobel, C. (in press). Event conceptualization in free view and at an eyeblink. *Cognition.*

Gleitman, L., January, D., Nappa, R., & Trueswell, J. (in press). On the give-and-take between event apprehension and utterance formulation*. Journal of Memory and Language.*

Glenberg, A.M. & Kaschak, M.P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558-565.

Gries, S.T. (2005). Syntactic priming: a corpus-based approach. *Journal of Psycholinguistic Research, 34,* 365–399.

Grieve, R., & Wales, R. (1973). Passives and topicalization. *British Journal of Psychology, 64,* 173-182.

Griffin, Z.M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*, B1–B14.

Griffin, Z.M. (2004a). Why look? Reasons for eye movements related to language production. In J. Henderson & F. Ferreira, (Eds.) *The Integration of Language, Vision, And Action: Eye Movements And the Visual World* (pp. 192-223). New York: Psychology Press.

Griffin, Z.M. (2004b). The eyes are right when the mouth is wrong. *Psychological Science, 15,* 814-821.

Griffin, Z. M. & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38,* 313–338.

Griffin, Z.M. & Bock, K. (2000). What the eyes say about speaking. *Psychological Research. 11(4),* 274-279.

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.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439–484.

Hahne, A. & Frederci, A.D. (1999). Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11(2),* 194-205.

Halliday, M.A.K. (1967/1968). Notes in transitivity and theme in English: Parts 1, 2 and 3. *Journal of Linguistics, 3(1), 3(2),* and *4(2).*

Halliday, M.A.K. (1985). *An Introduction to Functional Grammar*. London: Arnold.

Harley, T.A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes, 8,* 291-309.

Hartsuiker, R.J. & Kolk, H.H.J. (1998). Syntactic persistence in Dutch. *Language and Speech, 41,* 143– 184.

Hartsuiker, R. J., Kolk. H. H. J., & Huiskamp, P. (1999). Priming Word Order in Sentence Production. *Quarterly Journal of Experimental Psychology*, *52A,* 129-147.

Hartsuiker, R.J., Pickering, M.J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science, 15,* 409–414.

Hartsuiker, R.J. & Westenberg, C. (2000). Persistence of word order in written and spoken sentence production. *Cognition, 75,* B27–B39.

Haviland, S.B.E. & Clark, H.H. (1974). What' s new? Acquir ing new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13,* 512-521.

Hayashi, A., Tomlin, R.S. & Yokota, T. (2002). *Attention detection and Japanese*. Unpublished manuscript.

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language, 50,* 182-195.

Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.) *Language and space* (pp. 1-30). Cambridge, MA: MIT Press.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. New York, NY: Oxford University Press.

Jackendoff, R. (in press). Linguistics in cognitive science: The state of the art. *The Linguistic Review.*

Jakobson, R. (1936/1971). *Beitrag zur allgemeinen Kasuslehre. Gesamtbedeutungen der russischen Kassus.* The Hague: Mouton.

James, W. (1890) *The Principles of Psychology*. *Vol 1.* New York: Dover.

Kaiser, E. (2002). The syntax-pragmatics interface and Finnish ditransitive verbs. In M. van Koppen, E. Thrift, E. J. van der Torre, and M. Zimmermann (Eds.), *Proceedings of ConSOLE IX.*

Kaiser, E. & Trueswell, J.C. (2004). The role of discourse context in the processing of a flexible word-order language, *Cognition, 94,* 113-147.

Kako, E. (2006). Thematic role properties of subjects and objects. *Cognition, 101,* 1-42.

Kalita, J. & Shastri, L. (1994). A connectionist approach to generation of simple sentences and word choice. In G. Adriaens & U. Hahn (Eds.), *Parallel natural language processing* (pp. 395–420). Norwood, NJ: Ablex Publishing.

Kanwisher, N., McDermott, J. & Chun, M.M. (1997) The fusiform face area: a module in extrastriate visual cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302-4311.

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

Kelly, M., Bock, J.K., & Keil, F. (1986). Prototypicality in a linguistic context: Effects on sentence structure. *Journal of Memory and Language*, *25,* 59–74.

Kempen, G. & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11,* 201-258.

Kempen, G. & Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing: a cognitive architecture based on activation decay and simulation annealing. *Connection Science, 1,* 275-292.

King, T.H. (1995). *Configuring Topic and Focus in Russian*. Stanford, CA: CSLI Publications.

Knoeferele, P. & Crocker, M.W. (2006). The coordinated interplay of scene, utterance, and world knowledge: evidence from eye tracking. *Cognitive Science, 30* 481-529.

Knoeferle, P., Crocker, M.W., Scheepers, C., and Pickering, M.J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition, 95,* 95–127.

Kovtunova, I. (1976). *Porjadok slov i aktual'noe chlenenie predlozhenija,* Moscow: Prosveschenie.

Krylova, O. & Khavronina, S. (1988). *Word Order in Russian*. Moscow: Russky Yazyk.

Kuno, S. (1972). Functional sentence perspective. *Linguistic Inquiry, 3,* 269-320 .

Kutas, M. & Van Petten, C. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 83-143). New York: Academic Press.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.

Landau, B. & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences, 16,* 217-265.

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

Levelt, W.J.M. & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology, 14(1),* 78-106.

Levelt, W.J.M. & Meyer, A.S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology, 12(4),* 433-452.

Loebel, H. & Bock, J.K. (2003). Structural priming across languages. *Linguistics, 41(5),* 791-824.

MacWhinney, B. (1977). Starting points. *Language*, *53(1),* 152-168.

MacWhinney, B., & Bates, E. (1978). Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behavior, 17,* 539–558.

Mandler, J.M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review, 99,* 587–604.

Mathesisus, V. (1929). Functional Linguistics. In J. Vachek (Ed.). *Praguiana* (pp. 121-142). Amsterdam: John Benjamins.

Mathesius, V. (1947). *Čeština a obecný jazykozpyt*. Praha.

McCandliss, B.D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7,* 293-299.

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

McLean, J.F., Pickering, M.J., & Branigan, H.P. (2005). Lexical repetition and syntactic priming in dialogue. In J.C. Trueswell & M.K. Tanenhaus (Eds.), *Approaches to studying world-situated language use* (pp. 193-208). Cambridge, MA: The MIT Press.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88,* 375-407.

Melinger, A., & Dobel, C. (2005). Lexically-driven syntactic priming. *Cognition, 98,* B11-B20.

Meyer, A.S. (2004). The use of eye tracking in studies of sentence generation. In J. Henderson & F.Ferreira (Eds.), *The Interface of Language, Vision, and Action* (pp. 191-211). New York: Psychology Press.

Meyer, A.S., Roelofs, A, & Levelt, W.J.M. (2003). Word length effect in object naming: the role of a response criterion. *Journal of Memory and Language, 48*, 131-147.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66,* B25-B33.

Miller, N.A. & Kroll, J.F. (2002). Stroop effects in bilingual translation. *Memory & Cognition, 30,* 614-628.

Myachykov, A., Garrod, S.C., & Coquillon, S. (2007). *The interplay between perceptual and lexical priming during the generation of syntax.* Poster presented at 20th CUNY Conference on Human Sentence Processing, San Diego, USA.

Myachykov, A. & Posner, M.I. (2005). Attention in language. In L. Itti, G. Rees, & J. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 324-329). Academic Press/Elsevier.

Myachykov, A., Tomlin, R.S., and Posner, M.I. (2005). Attention and empirical studies of grammar. *The Linguistic Review, 22,* 347-364.

Myachykov, A., Posner, M.I. & Tomlin, R.S., (in press). A parallel interface for language and cognition in sentence production: Theory, method, and experimental evidence. *The Linguistic Review.*

Navarette, E., Basagni. B., Alario, F.-X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production? *The Quarterly Journal of Experimental Psychology, 59(10),* 1681-1690.

Newman, A.J., Pancheva, R., Ozawa, K., Neville, H.J., and Ullman, M.T. (2001). An event-related fMRI study of syntactic and semantic violations. *Journal of Psycholinguist Resources, 30(3),* 339-364.

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

Osgood, C. E. & Bock, J. K. (1977). Salience and sentencing: Some production principles. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory* (pp. 89-140). Hillsdale, N.J.: Erlbaum.

Osterhout, L. & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34,* 739-773.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76,* 241-263.

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

Pickering, M.J. & Branigan, H.P. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences, 3(4),* 136-141.

Pickering, M.J., Branigan, H.P., Cleland, A.A., & Stewart, A.J. (2000). Activation of syntactic information during language production. *Journal of Psycholinguistic Research, 29(2),* 205-216.

Pickering, M.J. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27,* 169-225.

Pickering, M.J. & Garrod, S. (2006). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11(3),* 105-110.

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

Posner, M.I. & Fan, J. (2007). Attention as an organ system. In J. Pomerantz (Ed.), *Neurobiology of Perception and Communication: From Synapse to Society. The IVth De Lange Conference* (pp.). Cambridge,UK: Cambridge University Press.

Posner, M.I., Nissen, M.J., & Ogden, W.C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H.L. Pick & E. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–157). Hillsdale: Lawrence Erlbaum Associates.

Posner M.I. & Petersen S.E. (1990). The attention system of the human brain. *Annual Review Neuroscience, 13,* 25-42.

Posner, M.I. & Raichle M.E. (1994). *Images of Mind.* Scientific American Library: New York.

Posner, M.I. & Rothbart, M.K. (2007a). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology, 58,* 1-23.

Posner, M.I. & Rothbart, M.K. (2007b). *Educating the Human Brain*. Washington, DC: APA Books.

Posner, M.I., Rothbart, M.K., & Sheese, B.E. (2007). Attention genes. *Developmental Science, 10,* 24-29.

Posner, M.I. & Petersen, S.E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25-42.

Potter, M.C. & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language, 29,* 633–654.

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.

Prince, A. & Smolensky, P. (1997). Optimality: From neural networks to universal grammar. *Science, 275,* 1604.1620.

Richardson, D.C., Spivey, M.J., McRae, K., & Barsalou, L.W. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, *27,* 767-780.

Raichle, M.E., Fiez, J.A., Videen, T.O., McCleod, A.M.K., Pardo, J.V., Fox,P.T., & Petersen, S.E. (1994). Practice-related changes in the human brain: functional anatomy during non-motor learning. *Cerebral Cortex, 4,* 8-26.

Roelofs, A. (1992). A spreading activation theory of lemma retrieval in speaking. *Cognition, 42,* 107-142.

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

Roelofs, A. (2003). Goal-referenced selection of verbal action: modelling attentional control in the Stroop task. *Psychological Review, 110,* 88-125.

Roelofs, A. (2006). Context effects of pictures and words in naming objects, reading words, and generating simple phrases. *The Quarterly Journal of Experimental Psychology, 59(10),* 1764-1784.

Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory and Language, 57,* 232-251.

Rohde, D.L.T. (2002*). A connectionist model of sentence comprehension and production*. Unpublished doctoral dissertation, Carnegie Mellon University, Department of Computer Science, Pittsburgh, PA.

Rohde, D.L.T. & Plaut, D.C. (2003). Connectionist models of sentence processing. *Cognitive Studies, Japan, 10(1),* 1-28.

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.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistics representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science, 6,* 557–567.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year olds. *Language Learning and Development, 2(1),* 27-49.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production. *Cognition, 89,* 179-205.

Scheepers, C. & Crocker, M. W. (2004) Constituent order priming from reading to listening: A visual-world study In M. Carreiras & C. Clifton, Jr. (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERP, and Beyond* (pp. 167-185). New York: Psychology Press.

Scheepers, C. & Myachykov, A. (2006). Implicit learning and syntactic priming of Russian word order in speakers unfamiliar with Russian. *Paper presented at 12th AMLAP Conference, Nijmegen, The Netherlands.*

Scheepers, C. & Myachykov, A. (in progress). *Lexical Overlap and Syntactic Priming in Sentence Production,* ongoing research.

Schenkein, J. (1980). A taxonomy for repeating action sequences in natural conversation. In B. Butterworth, (Ed.), *Language production Vol. 1* (pp. 21–47), Academic Press, London.

Shulman, G.L., Remington, R.W., McLean, J.P. (1979). Moving attention through visual space. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 522-526.

Seidenberg, M.S. (1997). Language acquisition and use: learning and applying probabilistic constraints, *Science, 275,* 1599–1603.

Slobin, D. (1996). From "thought" and "language" to thinking for speaking. In J. Gumperz & S. C. Levinson, (Eds) *Rethinking Linguistic Relativity* (pp. 70-96). Chicago: University of Chicago Press.

Slobin, D. (2003). Language and thought online: cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.) *Language In Mind* (pp. 157-192). Cambridge, MA: MIT Press.

Smith, M. & Wheeldon, L. (2001). Syntactic priming in spoken sentence production – an online study. *Cognition, 78,* 123-164.

Spivey, M.J. & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, *65,* 235-241.

Spivey, M.J., Richardson, D.C., & Fitneva, S.A. (2004). Thinking outside the brain: Spatial indices to visual and linguistic information. . In J. Henderson & F. Ferreira, (Eds.) *The Integration of Language, Vision, And Action: Eye Movements And the Visual World* (pp. 161-190). New York: Psychology Press.

Sridhar, S.N. (1988). *Cognition and Sentence Production: A Cross-linguistic study*. New York: Springer-Verlag.

Stemberger, J.P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language, Vol. 1,* (pp. 143–186). Hillsdale, NJ: Erlbaum.

Stemmer, B. (1999). Discourse studies in neurologically impaired populations: A quest for action. *Brain and Language, 68(3),* 402-418.

Szmrecsanyi, B. (2005). Creatures of habit: a corpus linguistic analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory, 1,* 113–149.

Tabor, W. & Tanenhaus, M. (1999). Dynamic sentence processing. *Cognitive Science, 23,* 491-515.

Tabossi, P., Spivey-Knowlton, M.J., McRae, K., & Tanenhaus, M.K. (1994). Semantic effects on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. In C. Umilta and M. Moscovitch (Eds.), *Attention and Performance XV* (pp. 589-615). Cambrideg, MA: The MIT Press.

Talmy, L. (2000). *Toward a cognitive semantics*. *Vol 1*. Cambridge, MA: MIT press.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M. & Sedivy, J.E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268,* 1632-1634.

Tanenhaus, M.K., & Trueswell, J.C. (1995). Sentence comprehension. In J.Miller & P Elimas (Eds.), *Handbook of perception and cognition: Speech, language, and cognition* (pp. 217-262). San Diego, CA: Academic Press.

Tannenbaum, P.H. & Williams, F. (1968). Generation of active and passive sentences as a function of subject or object focus. *Journal of Verbal Learning and Verbal Behavior, 7(1),* 246-250.

Titchener, E.B. (1908). *Lectures on the elementary psychology of feeling and attention.* New York: Macmillan.

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. (1997). Mapping Conceptual Representations into Linguistic Representations: The Role of Attention in Grammar. In J. Nuyts & E. Pederson (Eds.) *Language and Conceptualization* (pp. 162-189). Cambridge: Cambridge University Press.

Trueswell, J. (2005). Minding your manners in event perception: What eye movements can tell us about the relation between language and thought. *Paper presented at 11th Annual Conference on Architectures and Mechanisms for Language Processing.* Ghent, Belgium.

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.

van der Meulen, F.F. (2001). Moving eyes and naming objects. Unpublished doctoral dissertation, Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands.

van der Meulen, F.F. (2003). Stability of speech-to-gaze alignment in multiple object naming: Beautiful or boring? *Paper presented at the 12th Annual European Conference on Eye Movements,* Dundee, Scotland.

van der Meulen, F.F., Meyer, A.S., & Levelt, W.J.M. (2001). Eye movements during the production of nouns and pronouns. *Memory & Cognition, 29(3),* 512-521.

van der Velde, F. & de Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences, 29,* 37-70.

Vigliocco, G. & Hartsuiker, R.J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin, 128,* 442-472.

Vigliocco, G. & Nicol, J . (1998). Separating hierarchical relations and word order in language production. Is proximity concord syntactic or linear? *Cognition*, *68,* 13-29.

Vilkuna, M. (1995). Discourse configurationality in Finnish. In Kiss, K.É. (Ed.) *Discourse Configurational Languages* (pp. 244-268). New York: Oxford University Press.

Vinogradov, V.V. & Istrina, E.S. (1960). *Grammatika Russkogo Jazyka. Tom II: Sintaksis*, *ch. I.* Moskva: Izdatel'stvo Akademii Nauk SSSR.

Vygotsky, L. (1962). *Thought and Language.* Cambridge, MA:MIT Press.

Ward, N. (1991). *A flexible, parallel model of natural language generation.* Unpublished doctoral dissertation, Computer Science Division, University of California, Berkeley, CA.

Weiner, E.J., & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics, 19,* 29-58.

Whorf, B.L., (1965), Language, Thought and Reality. In J. B. Caroll (Ed.), *Selected Writings of B. L. Whorf* (pp. 207-219), Cambridge, MA: MIT Press.

Yang, S. & Lust, B. (2005). Effects of Bilingualism on the Attention Networks Test: Its significance and implications. *Web proceedings of Boston University Conference on Language Development*. Boston, MA.

Yantis, S., Schwarzbach, J., Serences, J.T., Carlson, R.L., Steinmetz, M.A., Pekar, J.J., & Courtney, S.M. (2002). Transient Neural Activity In Human Parietal Cortex During Spatial Attention Shifts. *Nature Neuroscience, 5(10),* 995-1002.

Yamashita, H. & Chang, F. (2006). Sentence production in Japanese. In M. Nakayama, R. Mazuka, & Y. Shirai (Eds.), *Handbook of East Asian Psycholinguistics*, Cambridge, UK: Cambridge University Press.

Yokoyama, O. (1986). *Discourse and Word Order.* Amsterdam/Philadelphia: Benjamins.

Appendices

Appendix I. Figures

Figure 1

Transitive event

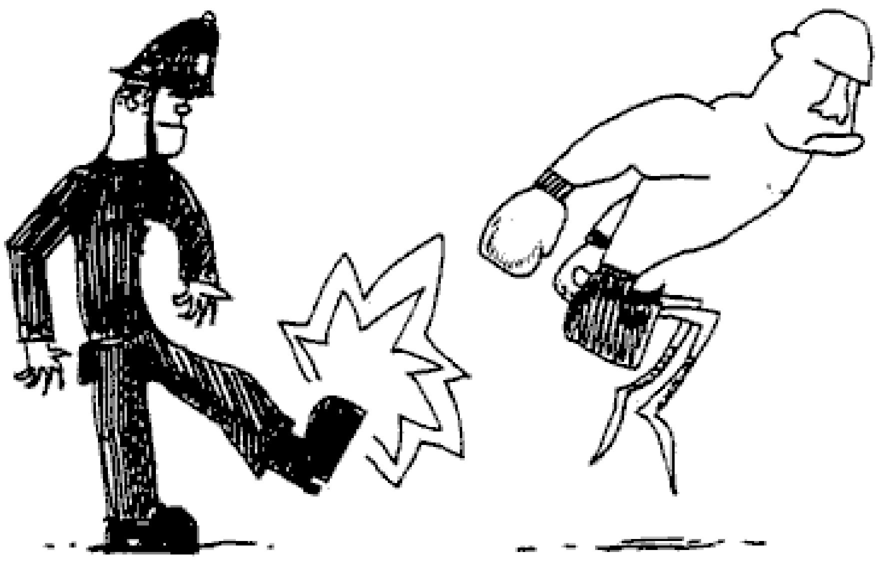


Figure 2

Fish Film (Tomlin, 1995)

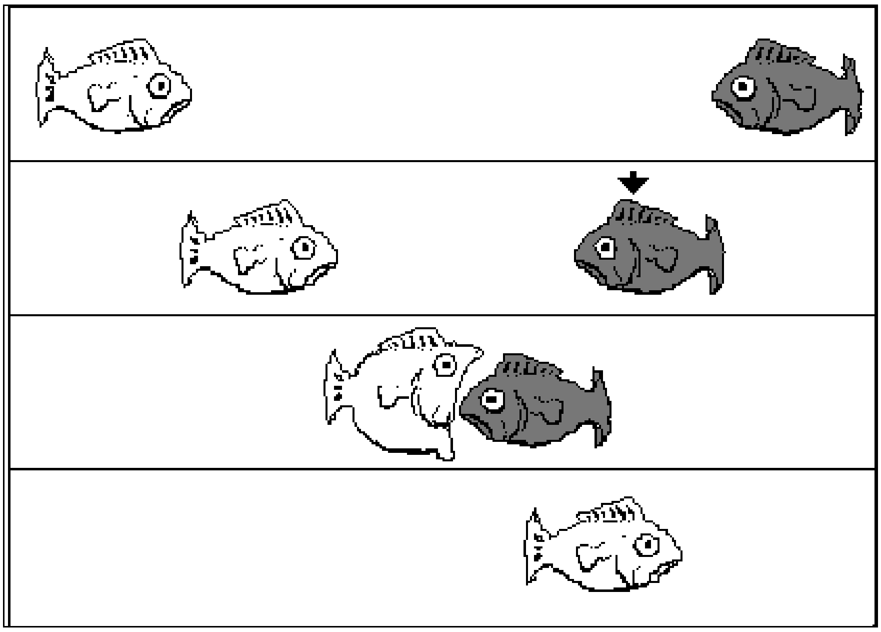


Figure 3

Model of sentence production (Bock & Levelt, 1994)

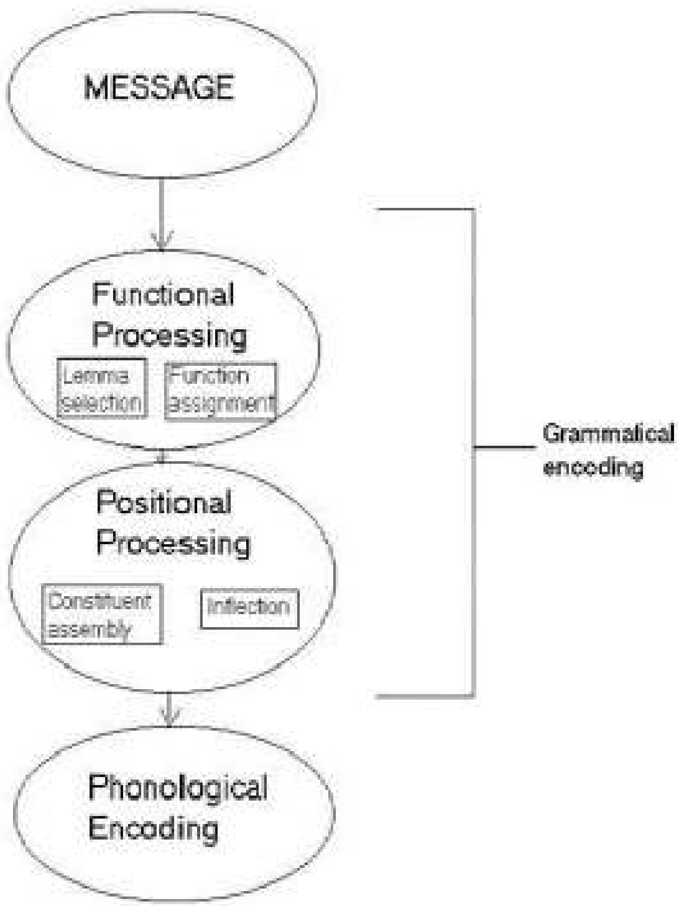


Figure 4.

Model of sentence production (Tomlin, 1997)

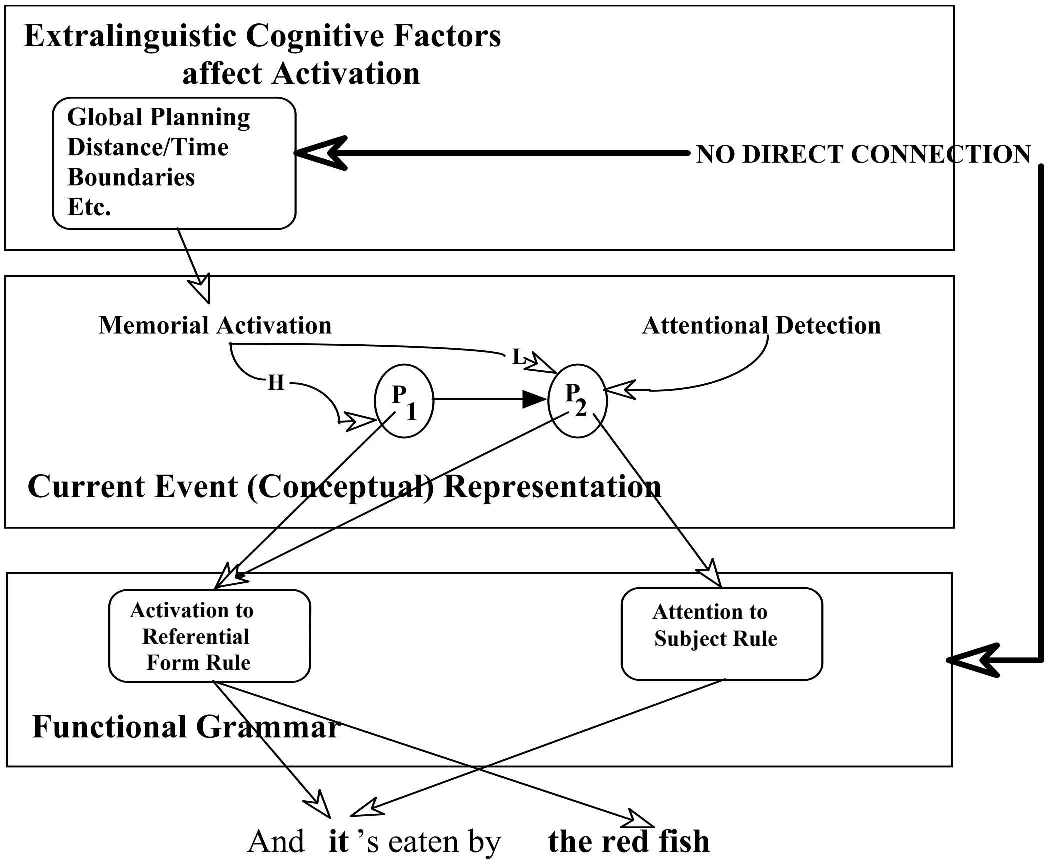


Figure 5

Global Workspace Model (Dehaene, Kerszberg, and Changeux, 1998)

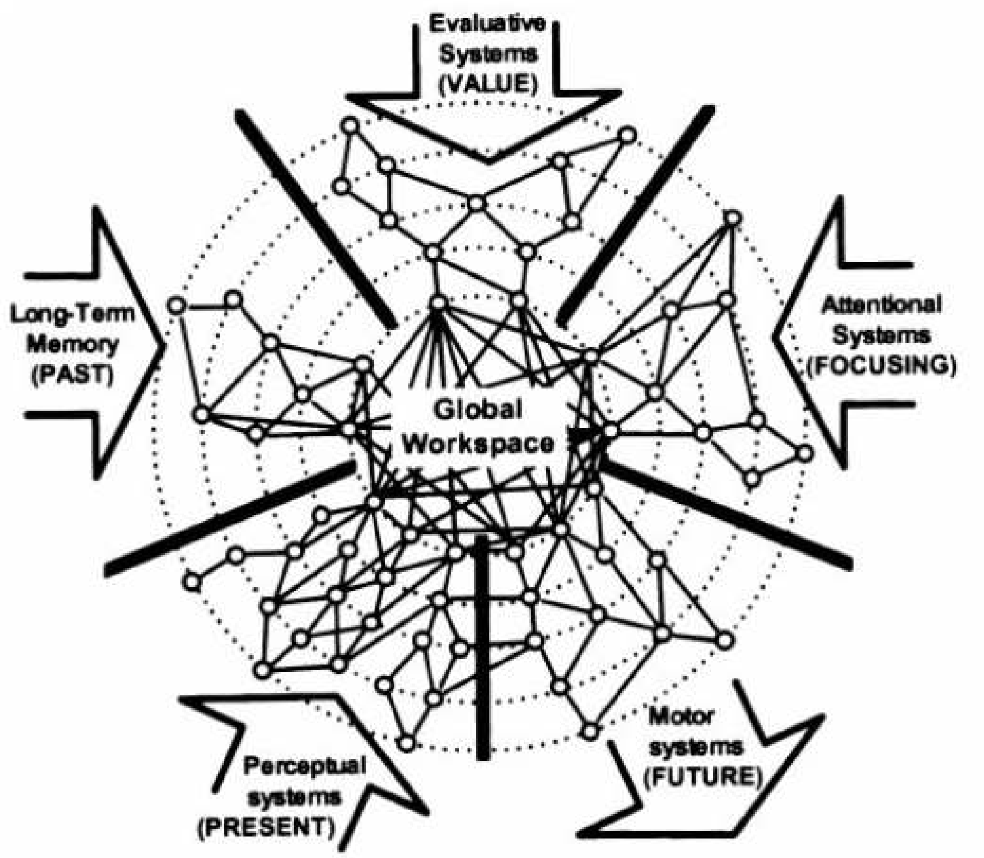
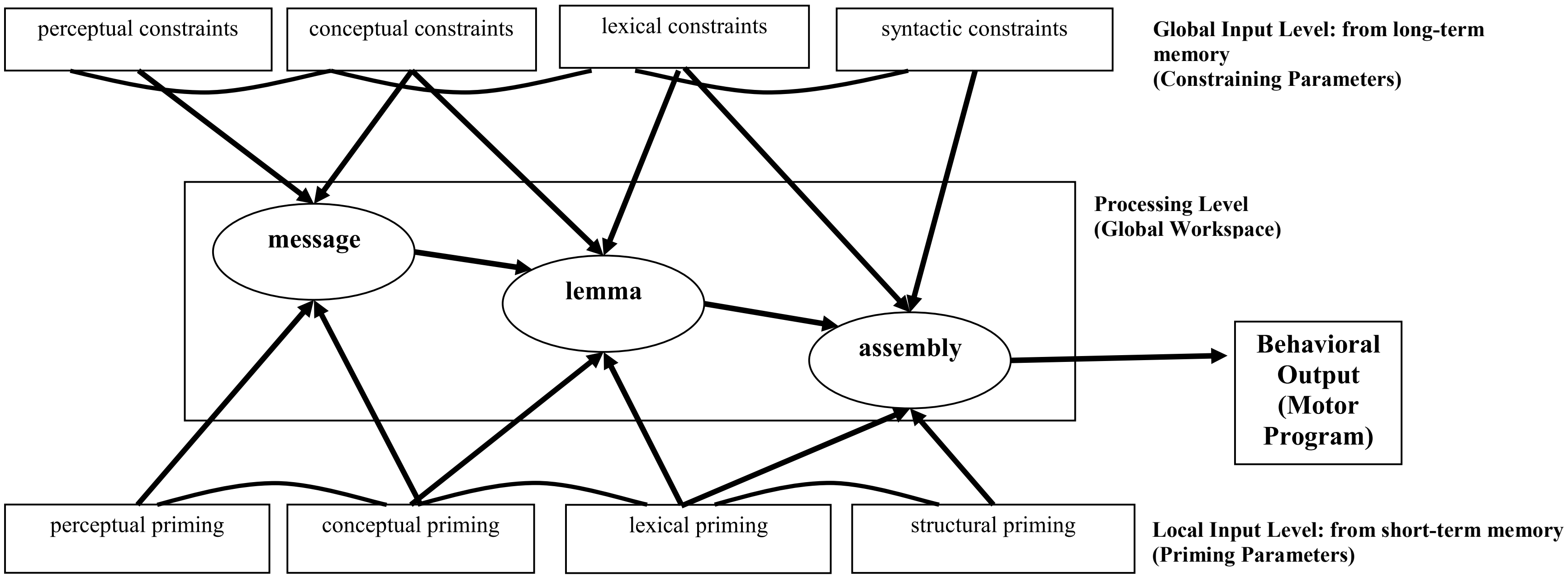


Figure 6.

Parallel Interface Model (Example of Voice Alternation)



Interfacing principles:

- **Competition:** Priming will sometimes compete with the linguistic constraints
- **Connection:** Processing within a stage closer to the output will have a greater chance to influence the output's parameters
- **Weight:** The larger the weight of the connection, the greater the chances its influences will survive
- **Interaction:** Only the boxes adjacent to each other can possibly interact

Figure 7

Production stages and measurements.

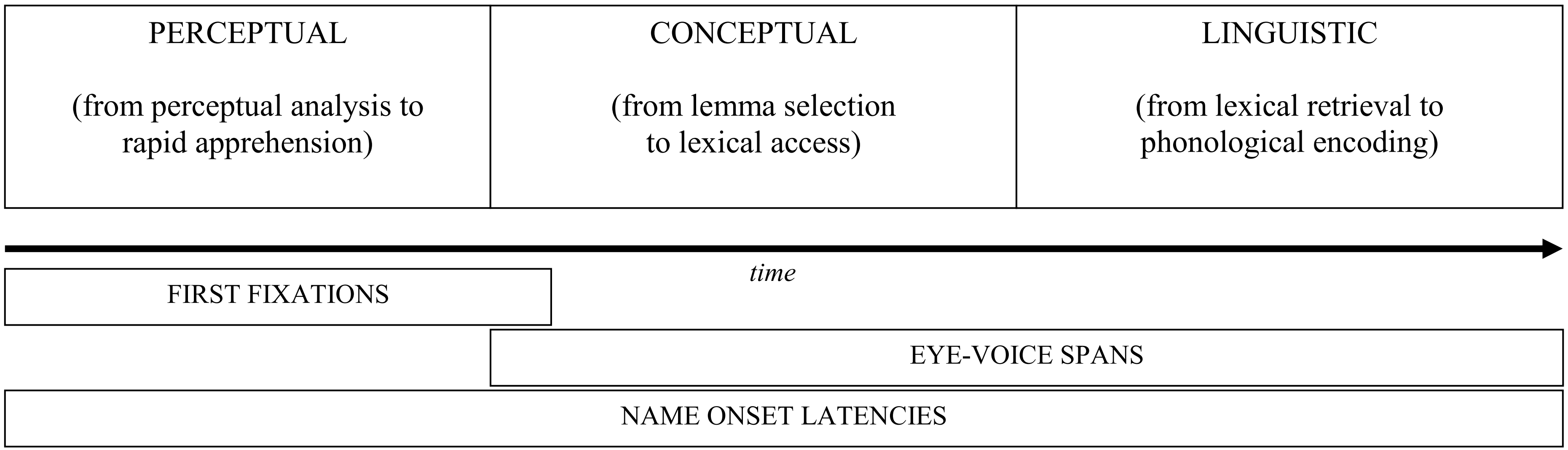


Figure 8a.

Experiments 2 and 3. Naming task.

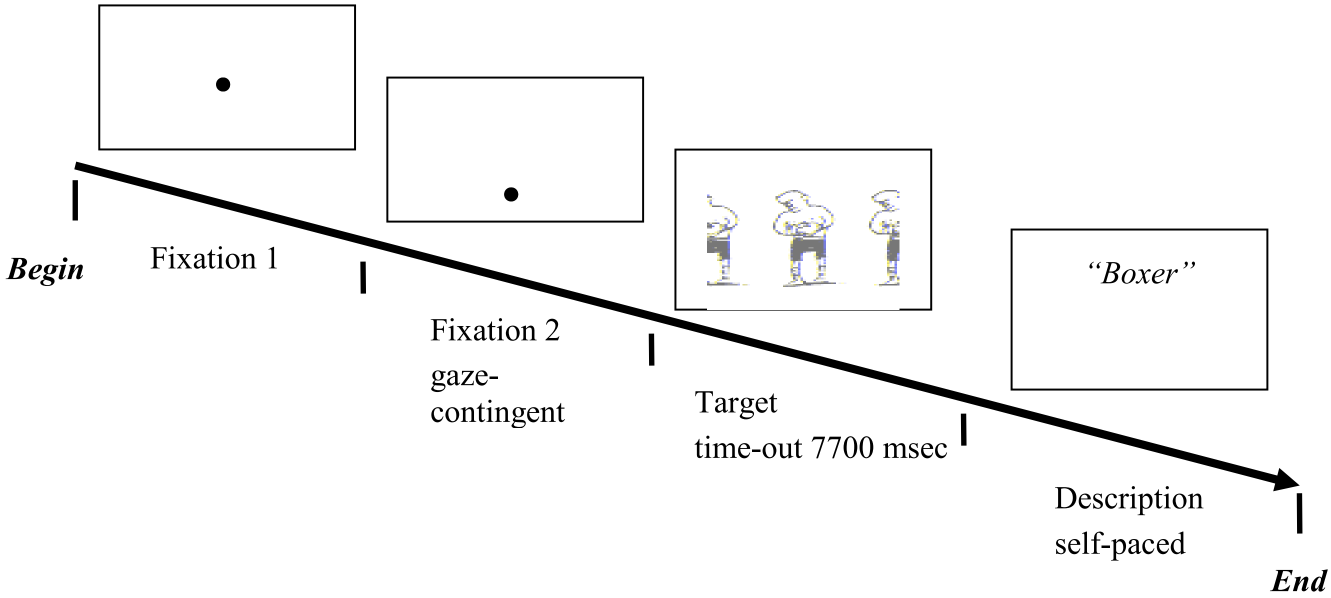


Figure 8b.

Experiments 2 and 3. Event description task.

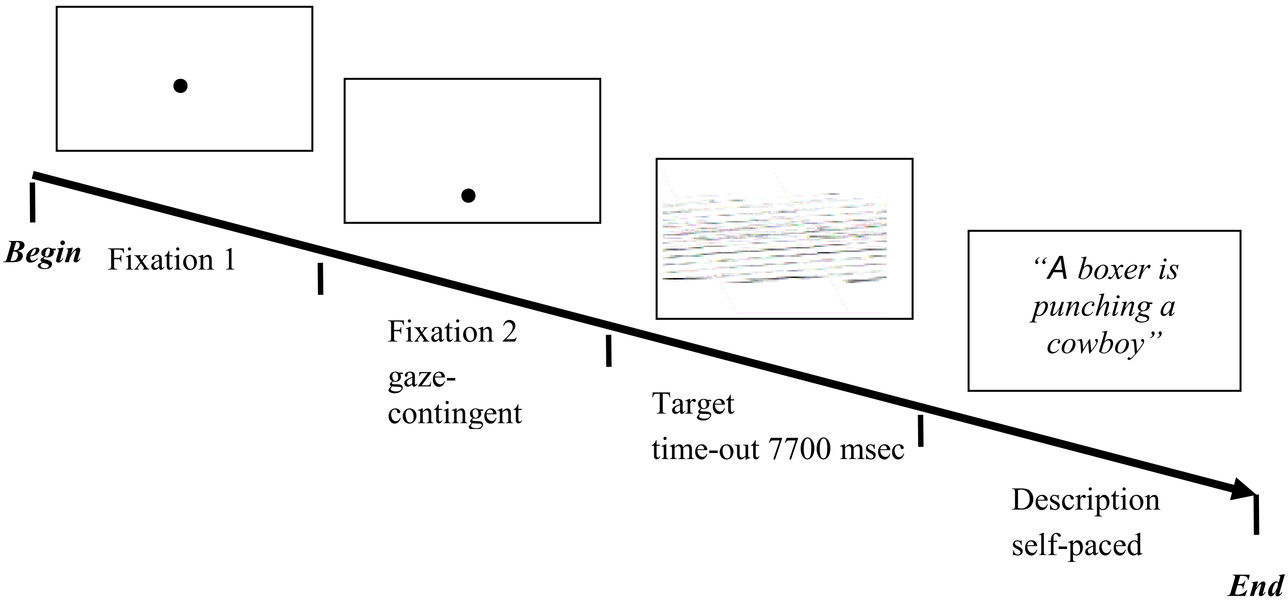


Figure 9.

Experiments 2 and 3. Main results from Naming task.

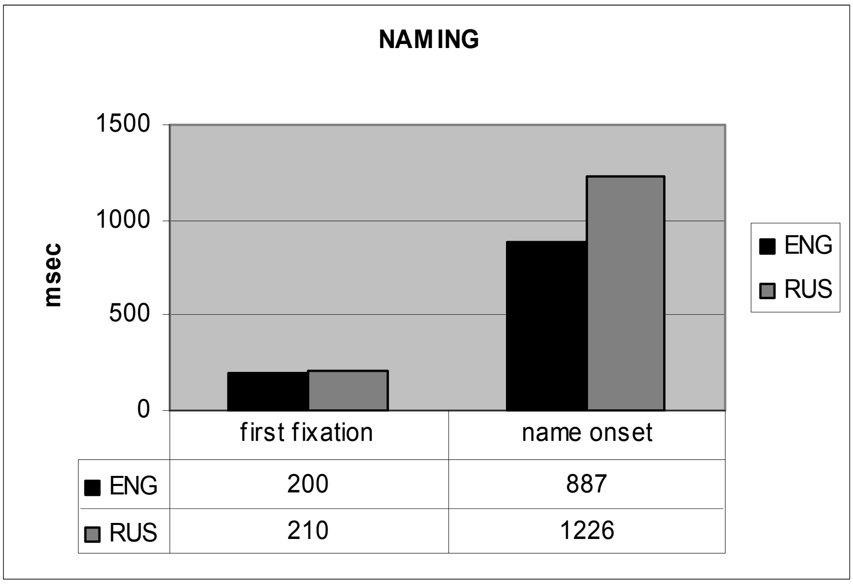


Figure 10.

Experiments 2 and 3. First Fixations: Event description task.

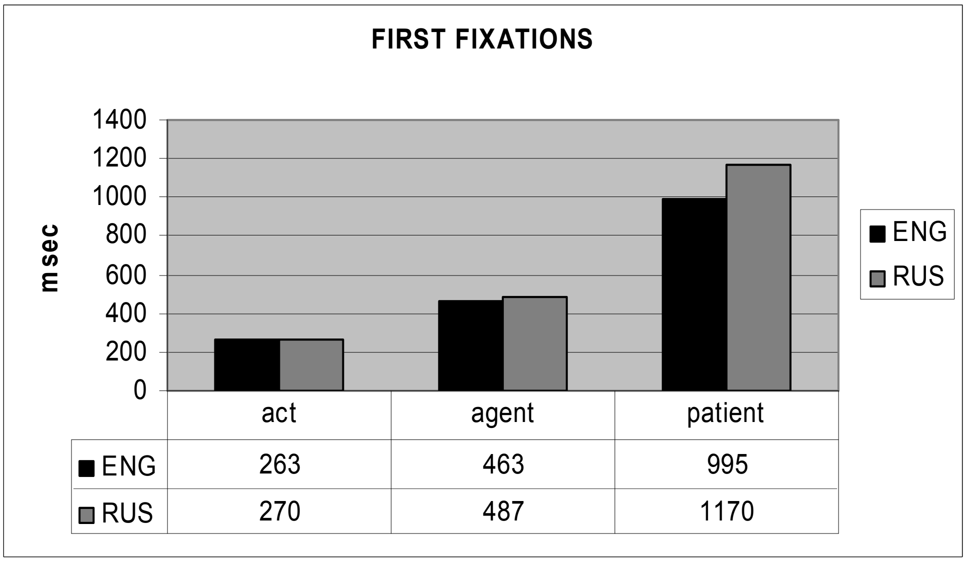


Figure 11.

Experiments 2 and 3. Eye-Voice Spans: Event description task.

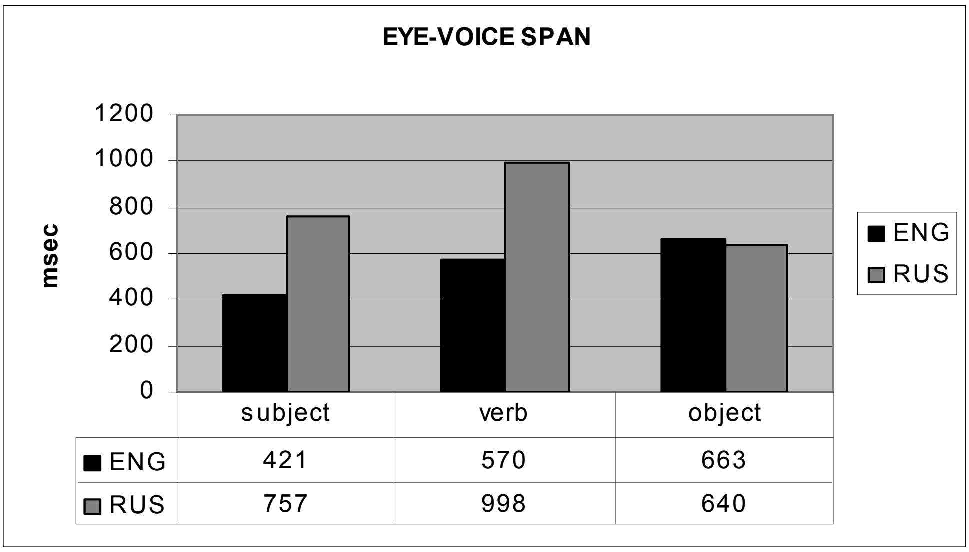


Figure 12.

Experiments 2 and 3. Name Onset Latencies: Event description task.

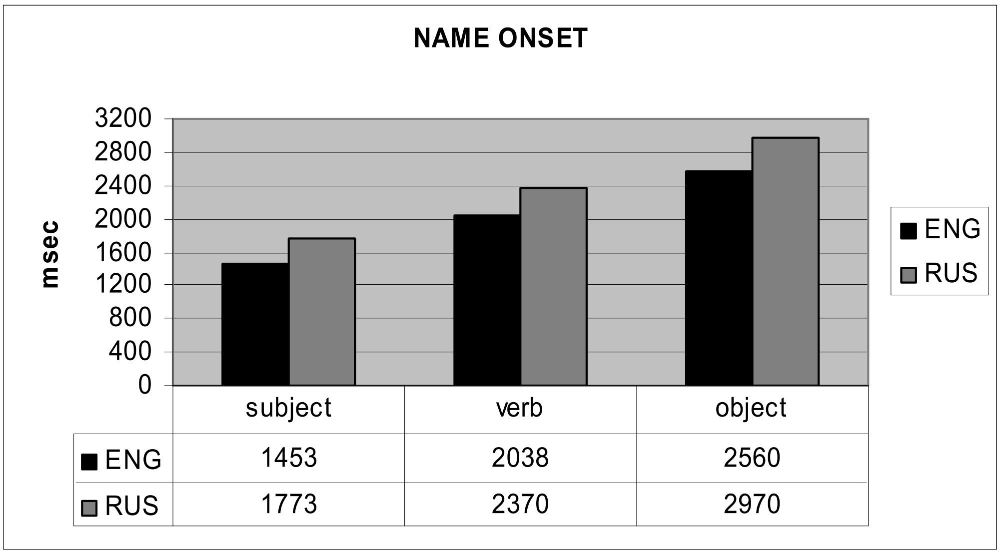


Figure 13

Experiment 4. Experimental sequence.

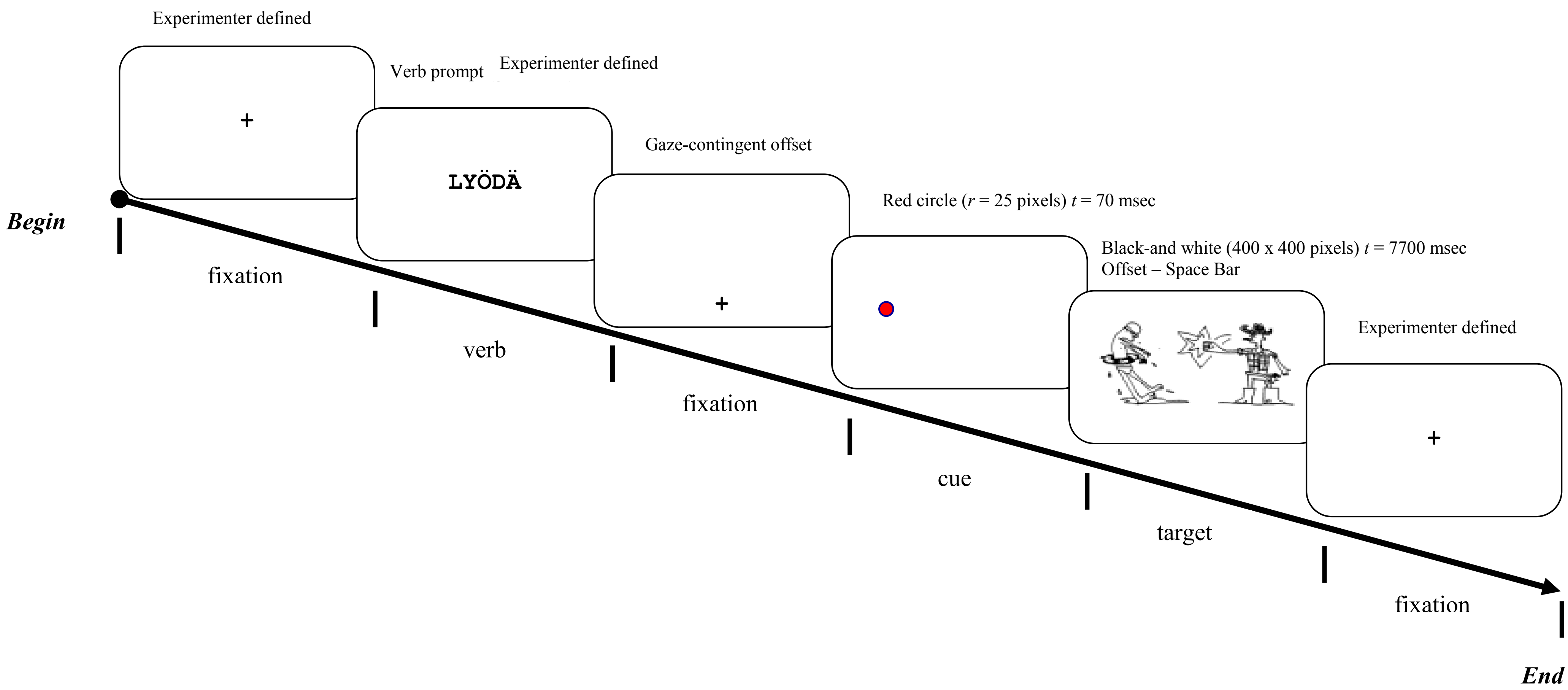


Figure 14

Experiment 5. Experimental sequence.

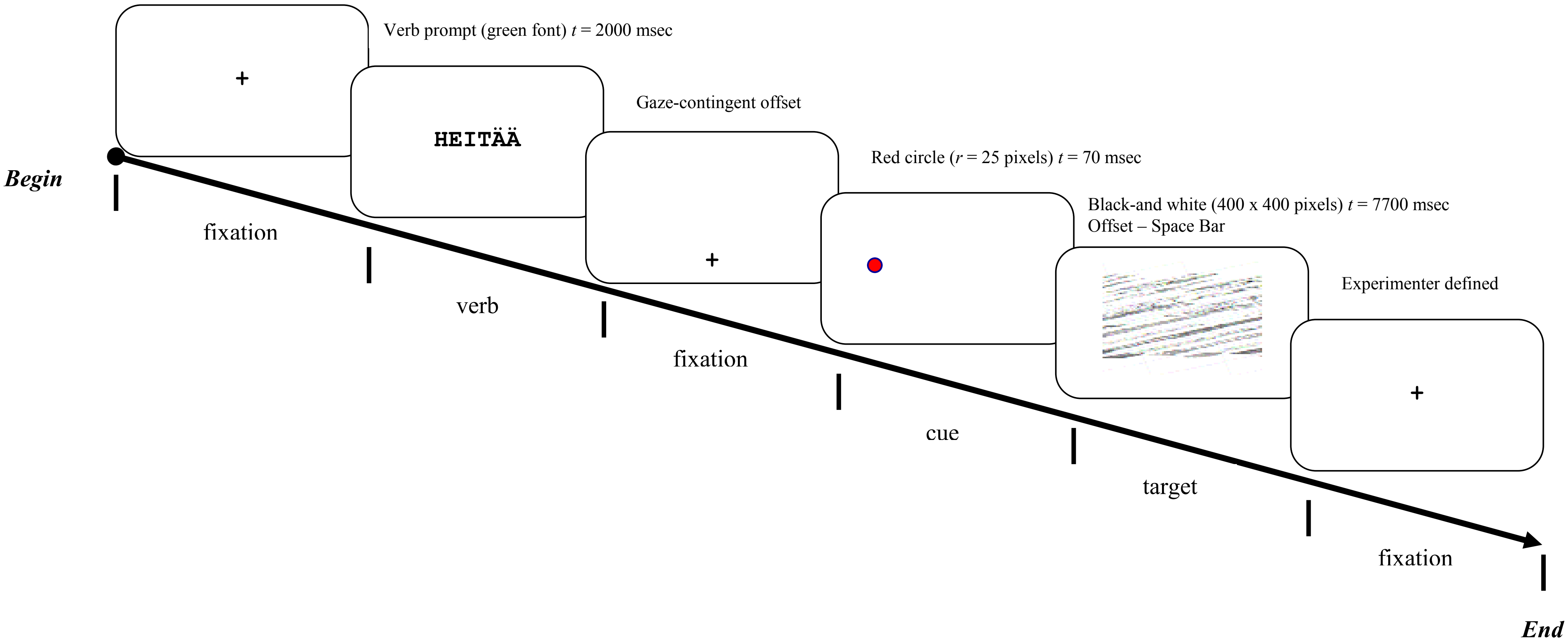


Figure 15.

Experiment 5. Example of experimental materials.

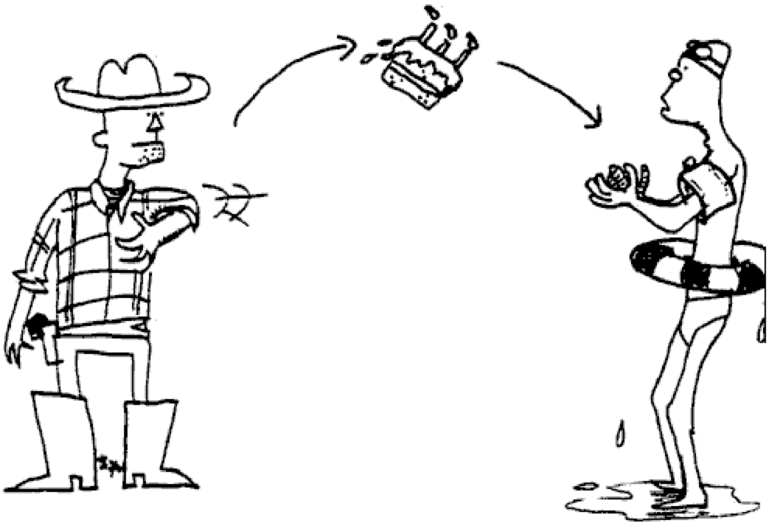


Figure 16.

Experiment 6. Example of filler materials.

Figure 17

Experiment 6. Experimental sequence.

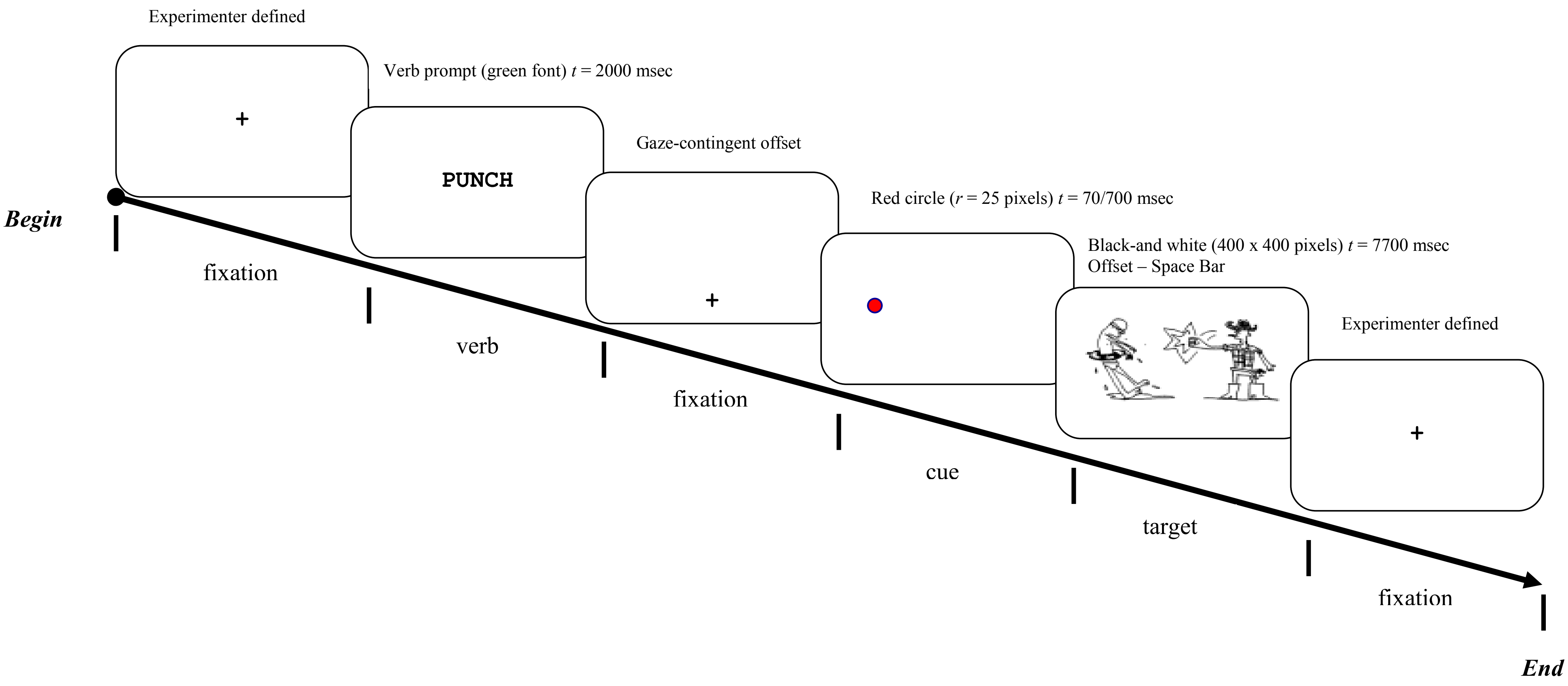


Figure 18

Experiment 7. Experimental sequence.

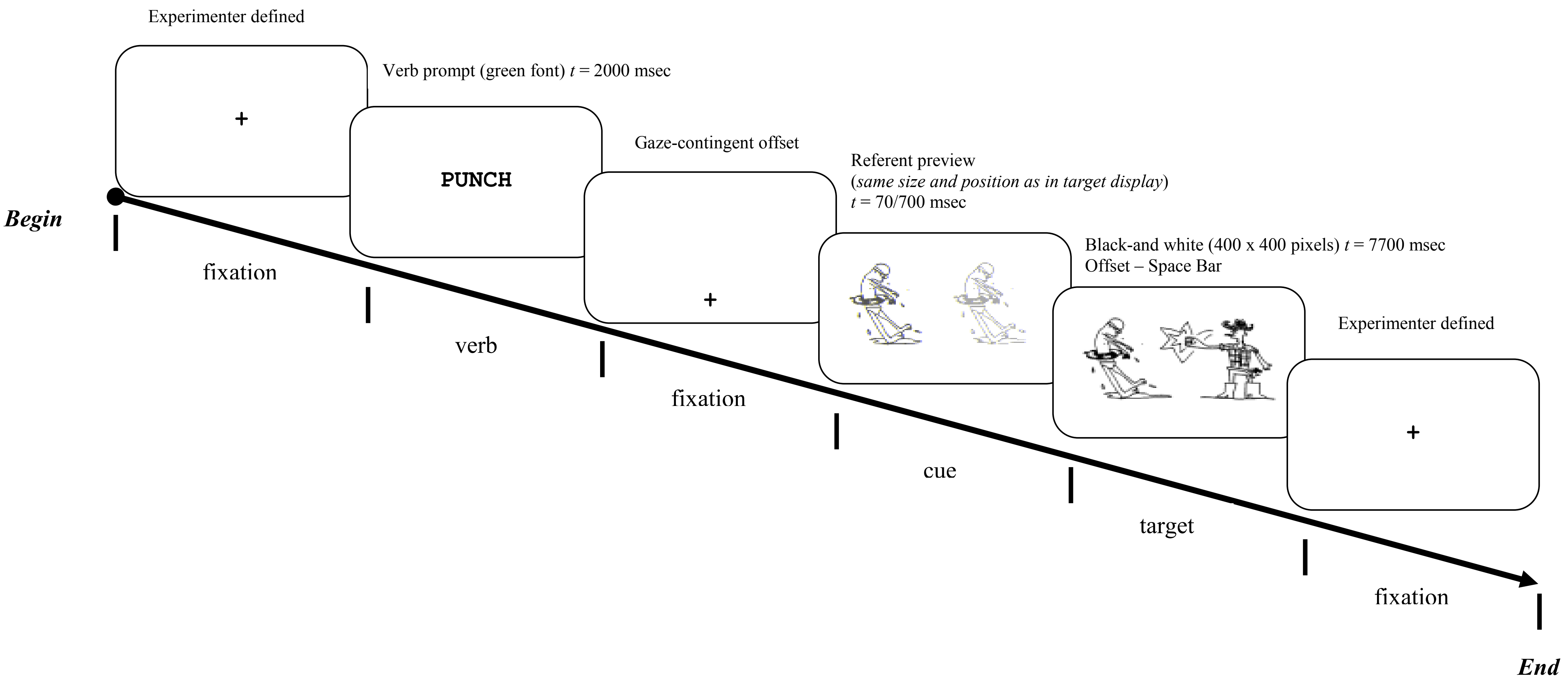


Figure 19

Experiment 6. Word order. Cue location x Cue SOA

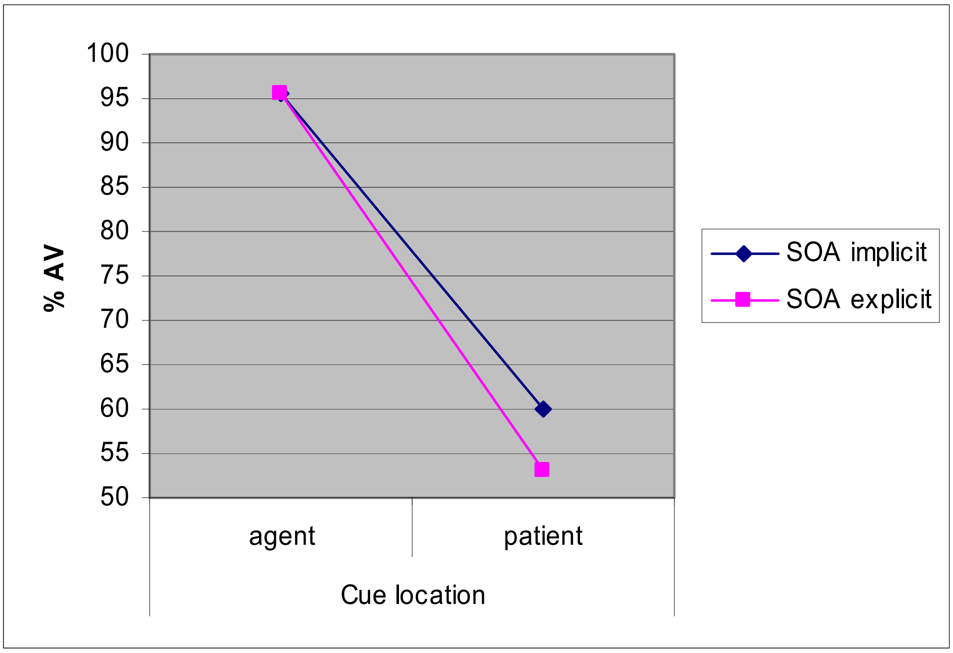


Figure 20.

Experiment 6. Word order. Cue location x Instruction

Figure 21.

Experiment 6. First fixation to the non-cued referent.

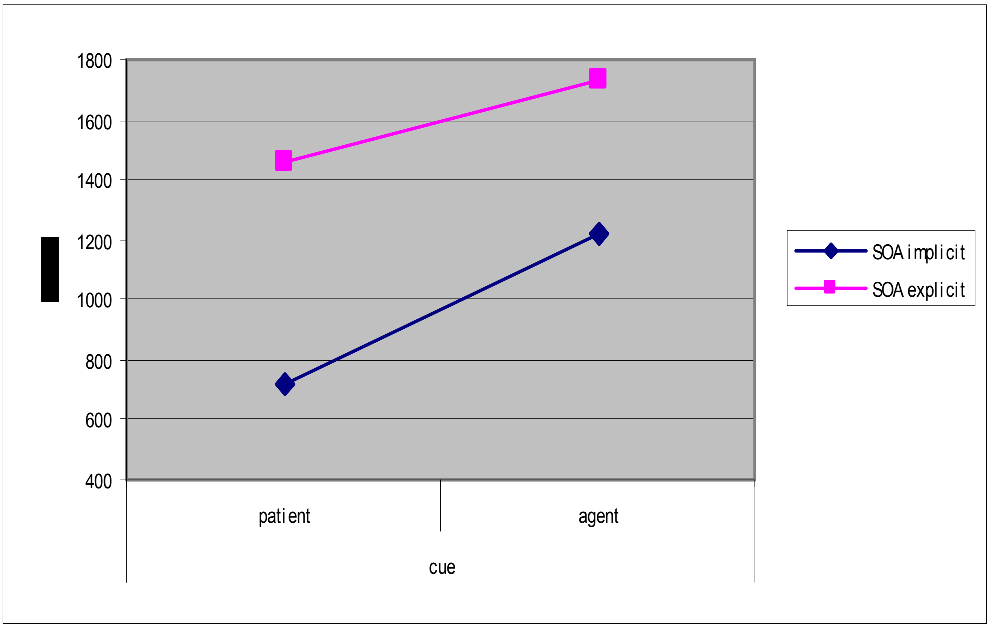


Figure 22.

Experiment 6. Name onset latencies (Subject). Cue Location x SOA

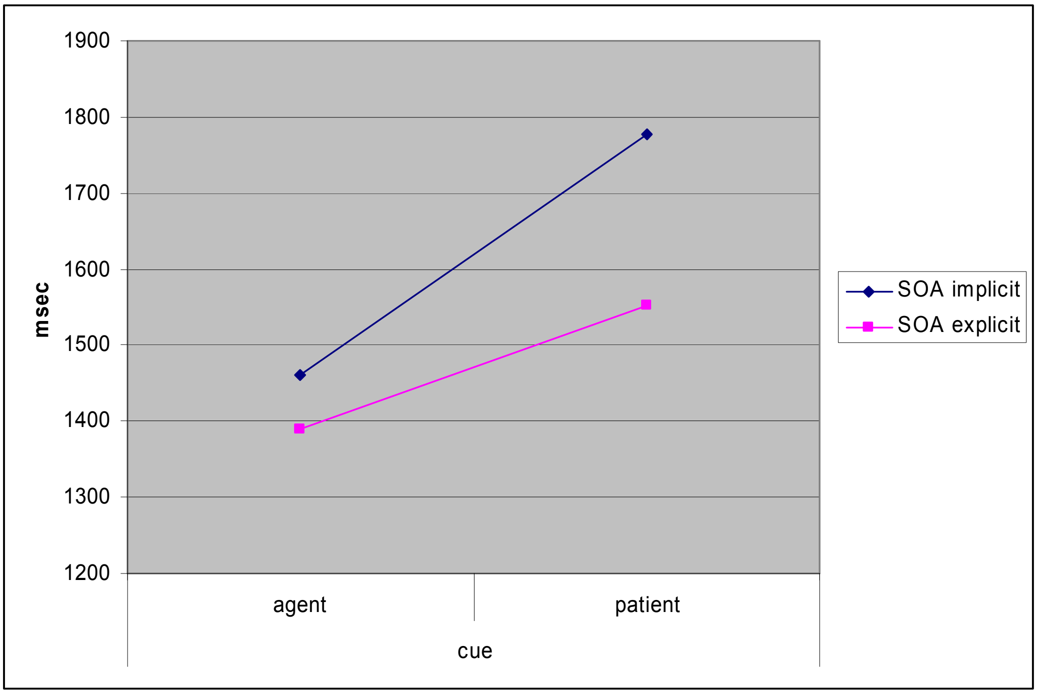


Figure 23.

Experiment 6. Name onset latencies (Verb). Cue Location x Instruction

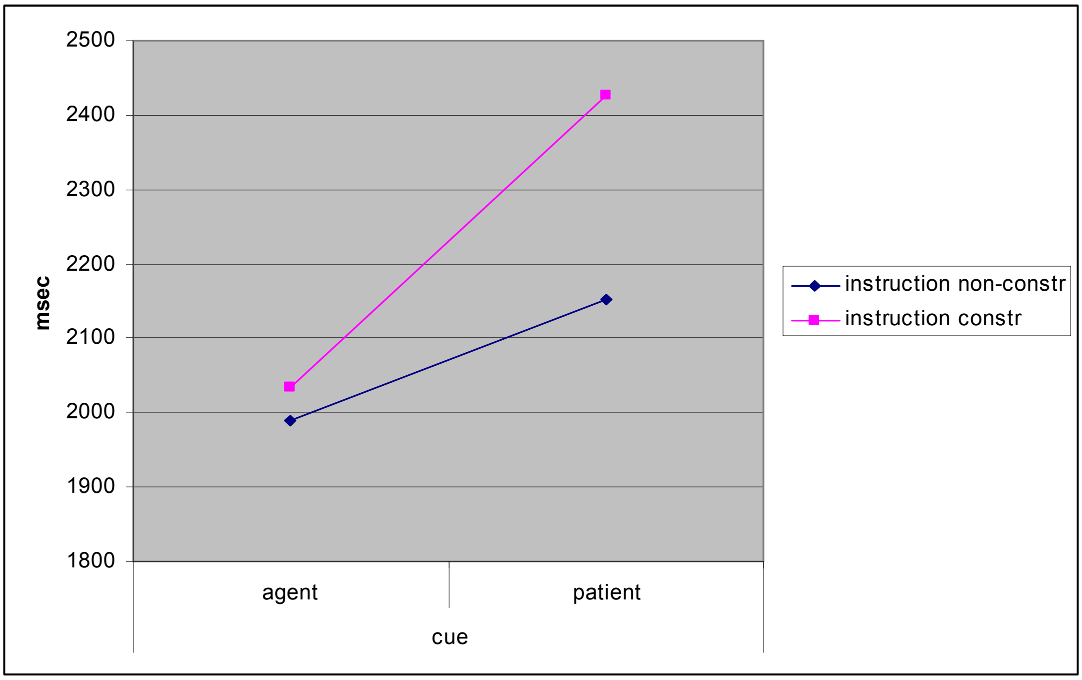
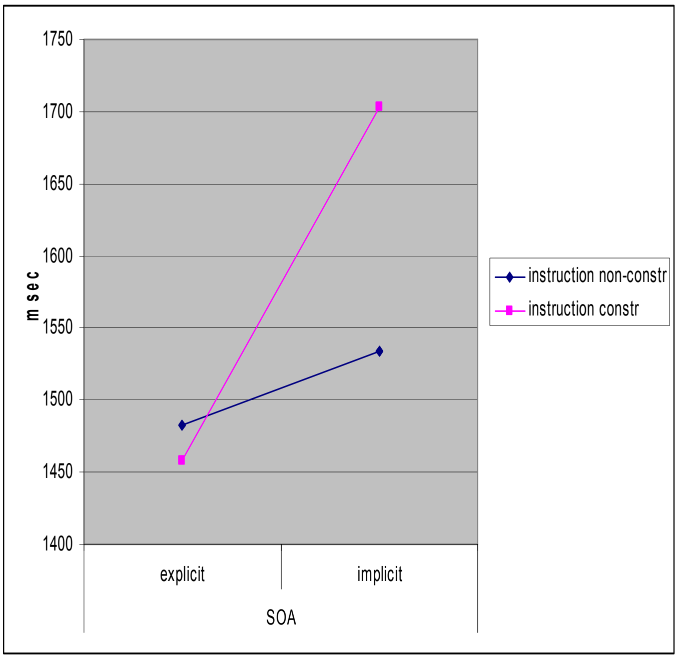


Figure 24 (a,b).

Experiments 6 and 7. Name onset latencies (Subject). SOA x Instruction

Experiment 6



Experiment 7

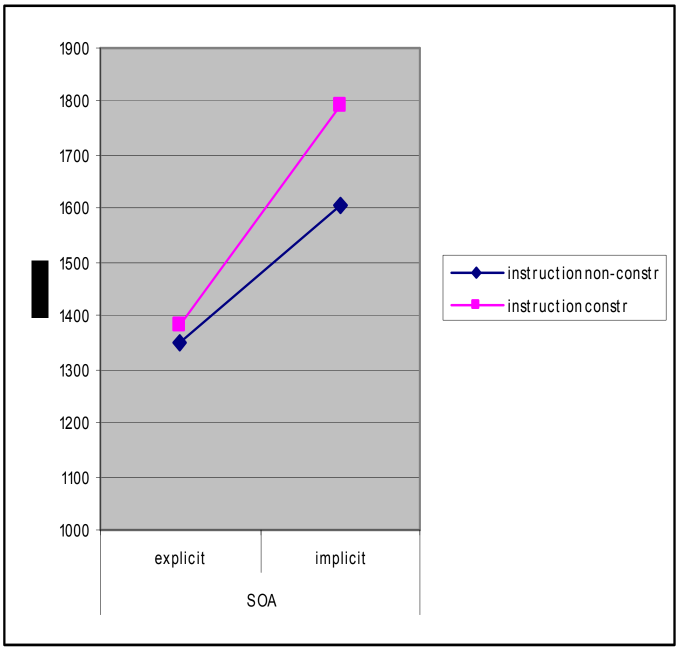


Figure 25

Experiments 6 and 7. First fixation onsets to the non-cued referent. Experiment x Cue Location interaction.

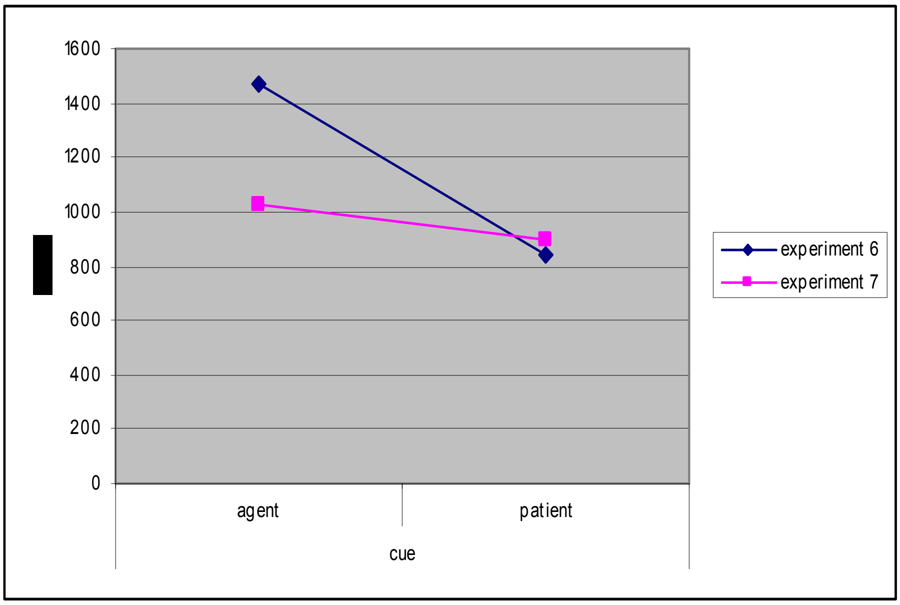


Figure 26

Experiments 6 and 7. Eye-voice spans. Experiment x SOA interaction.

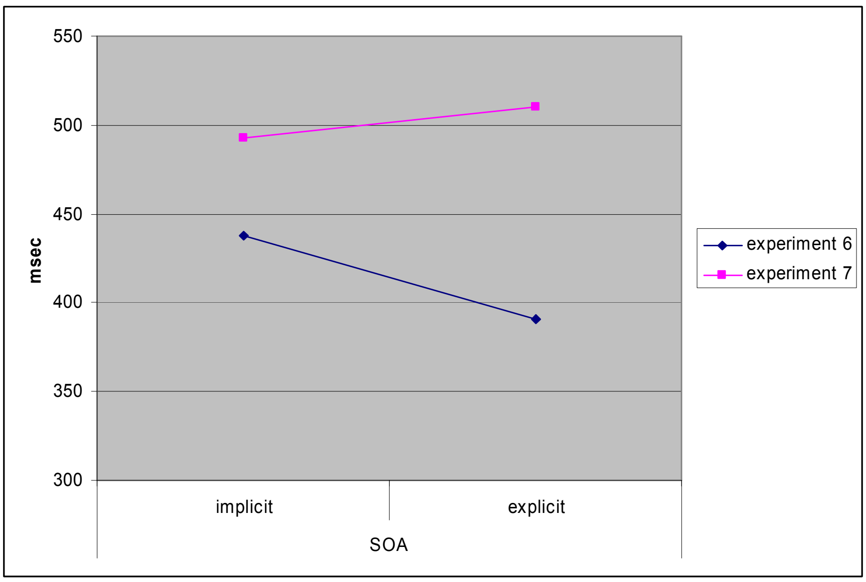


Figure 27

Experiments 6 and 7. Eye-voice spans. Experiment x Constituent interaction.

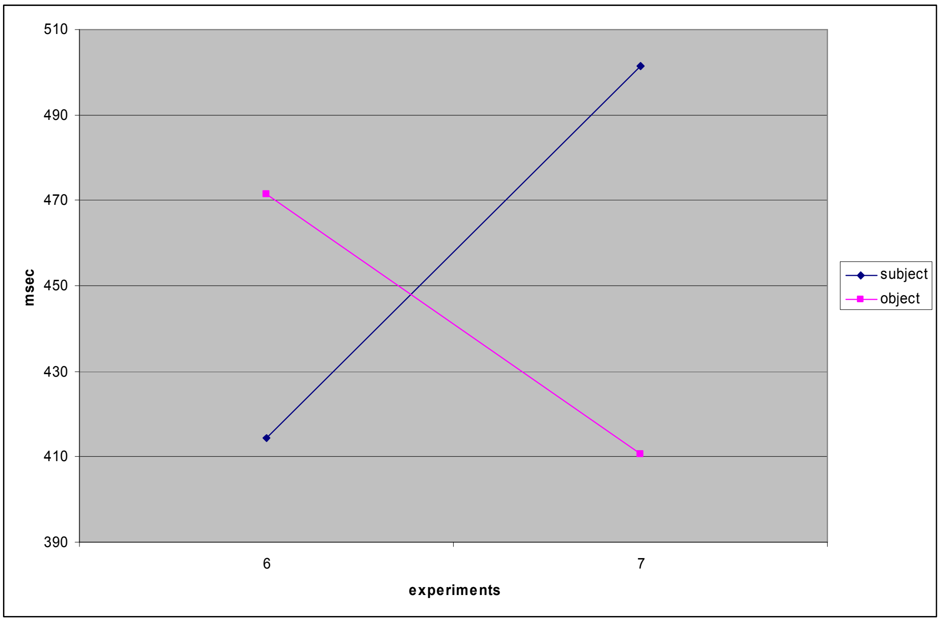


Figure 28

Experiments 6 and 7. Name onset latencies. Experiment x SOA interaction.

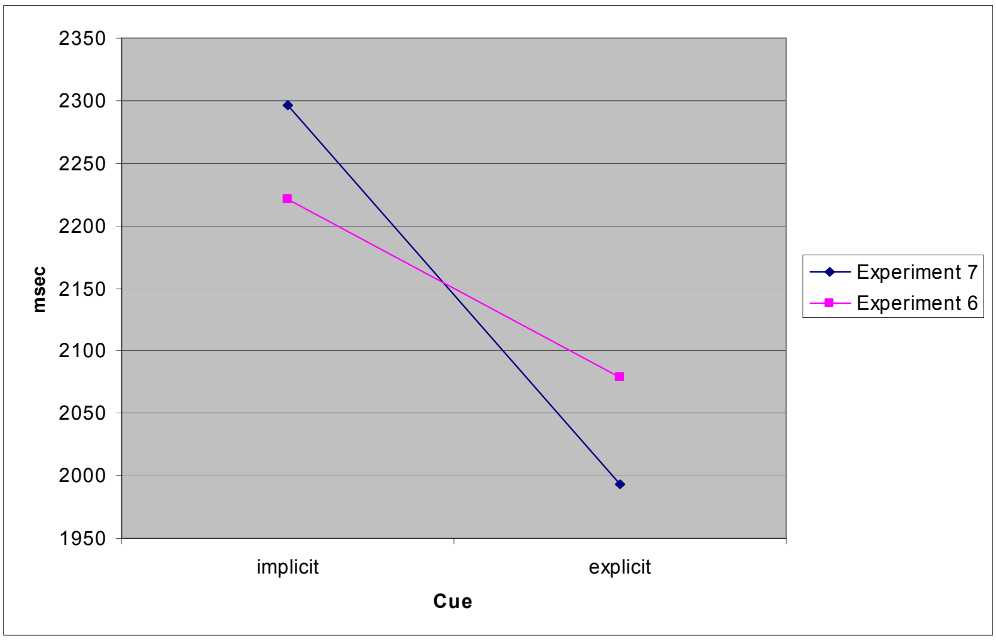


Figure 29

Experiment 8. Experimental sequence.

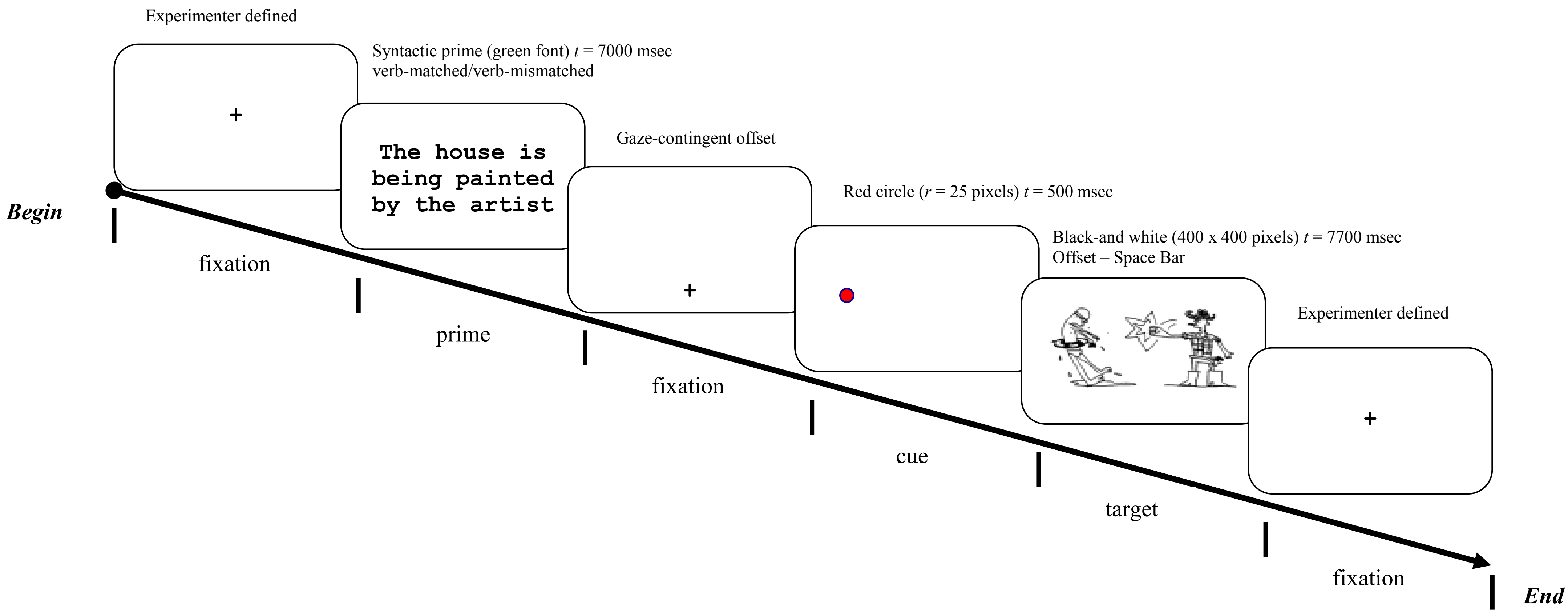


Figure 30

Experiment 9.Experimental sequence.

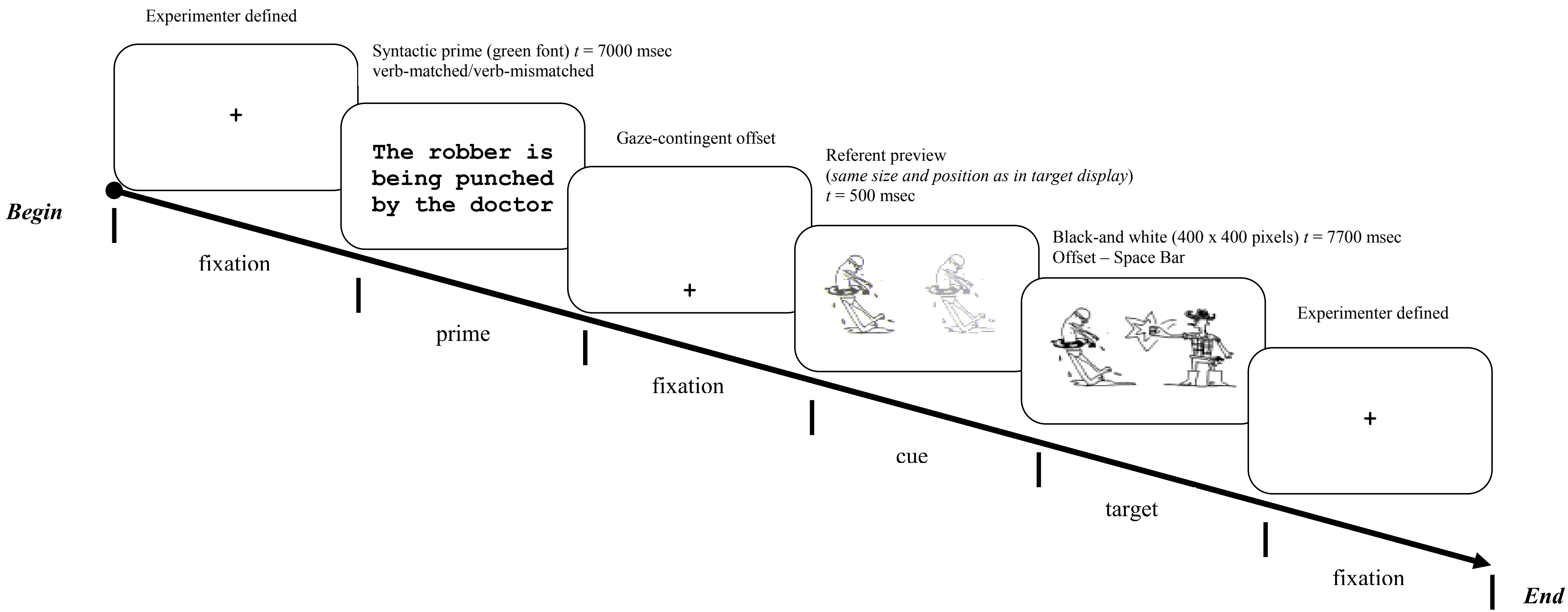


Figure 31

Experiment 8. Word order. Prime x Verb Match interaction.

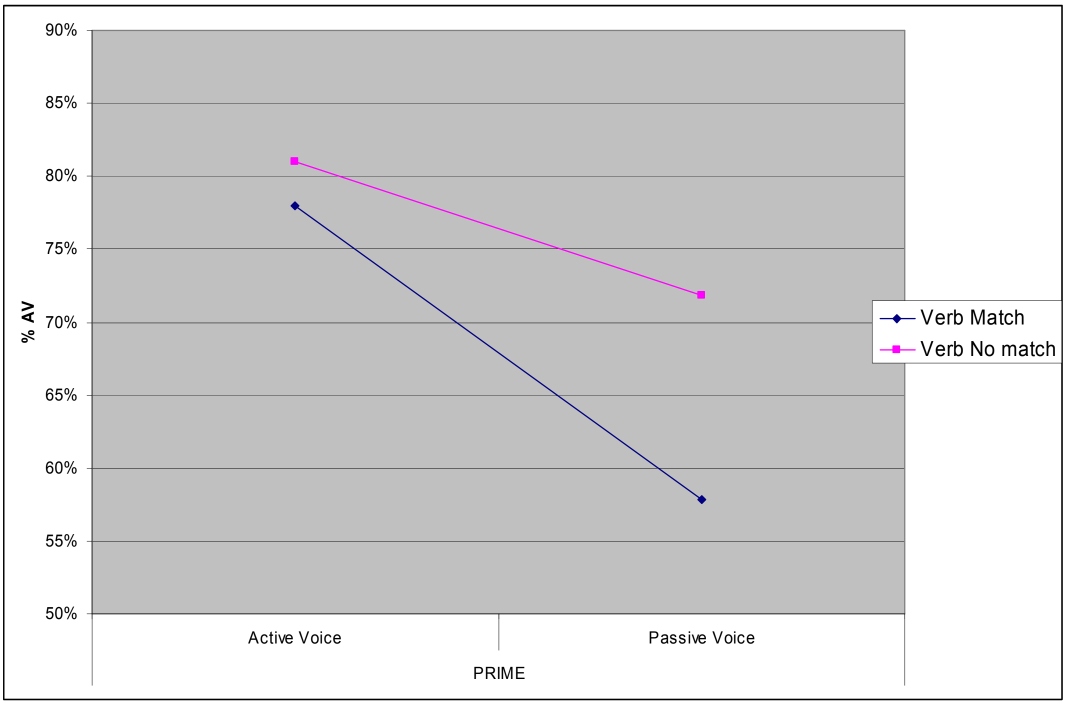


Figure 32

Experiment 8. Word order. Cue x Verb Match interaction

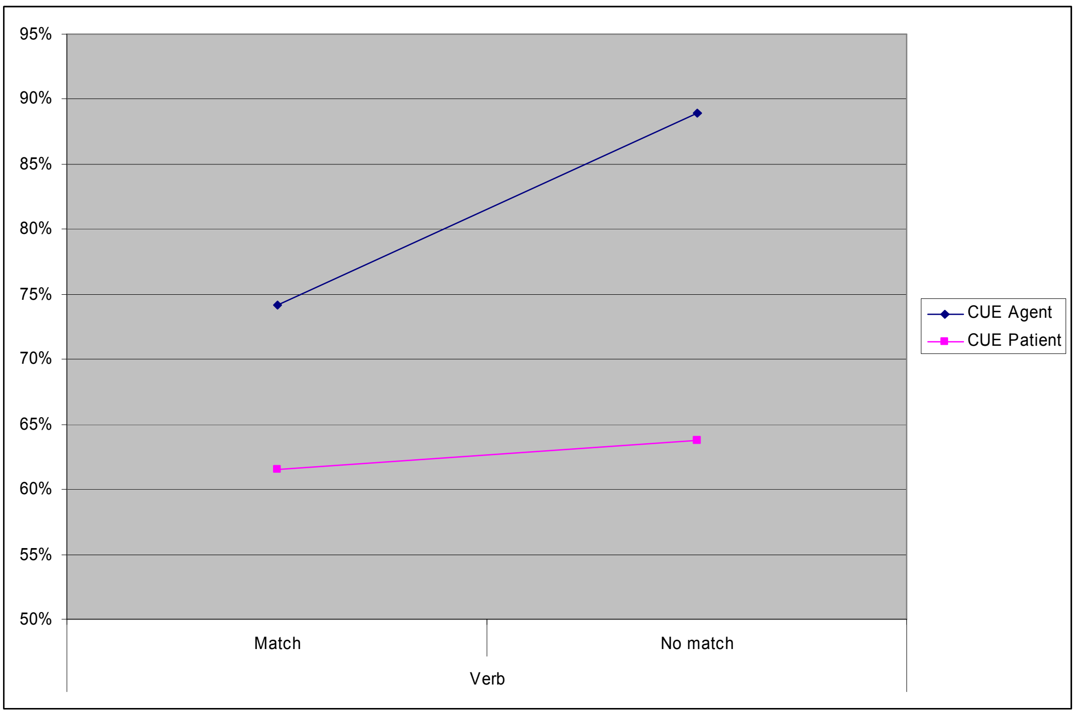


Figure 33

Experiment 8. Eye-voice spans. Prime x Cue interaction. Subject

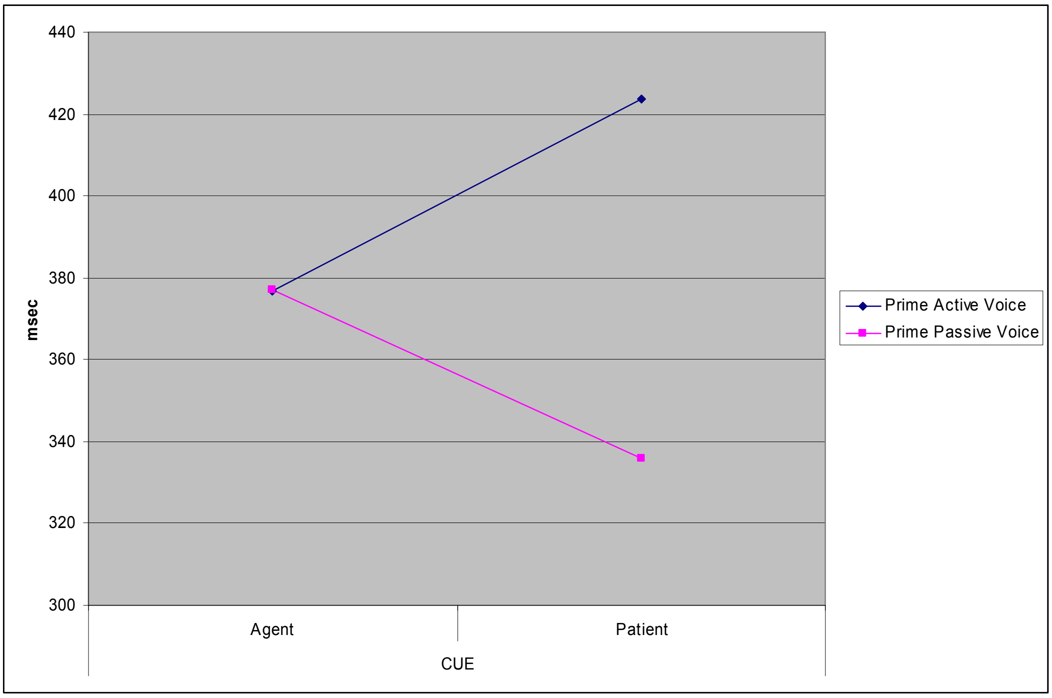


Figure 34

Experiment 8. Eye-voice spans. Prime x Verb Match interaction. Subject.

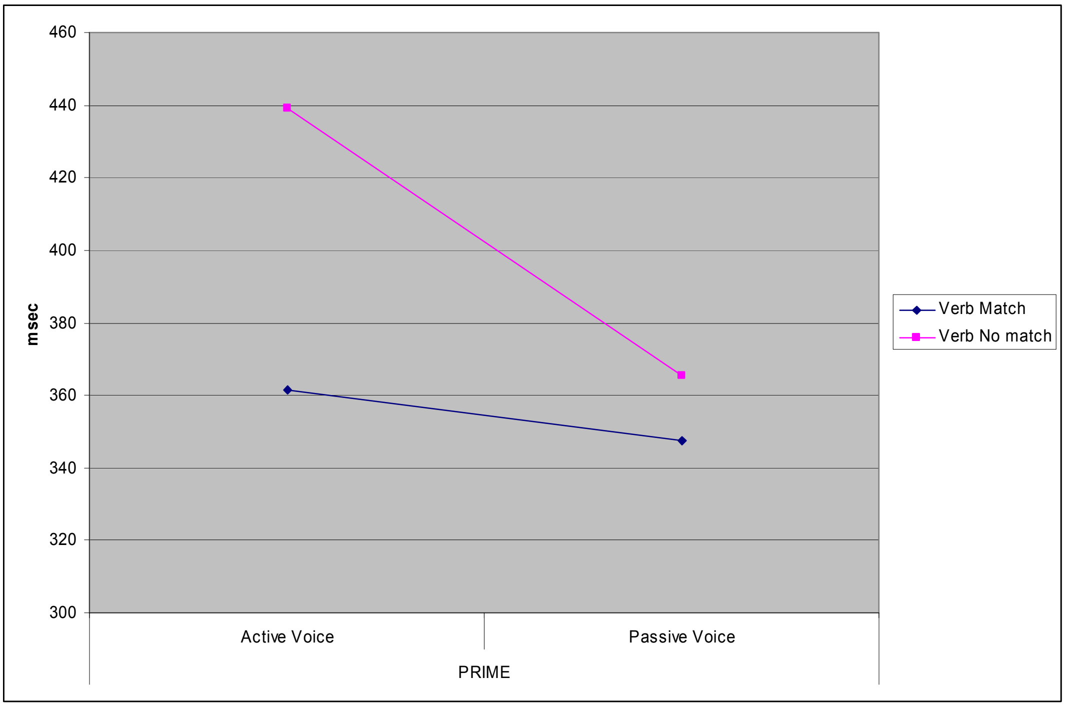


Figure 35

Experiments 8 and 9. Word order. Experiment x Prime interaction.

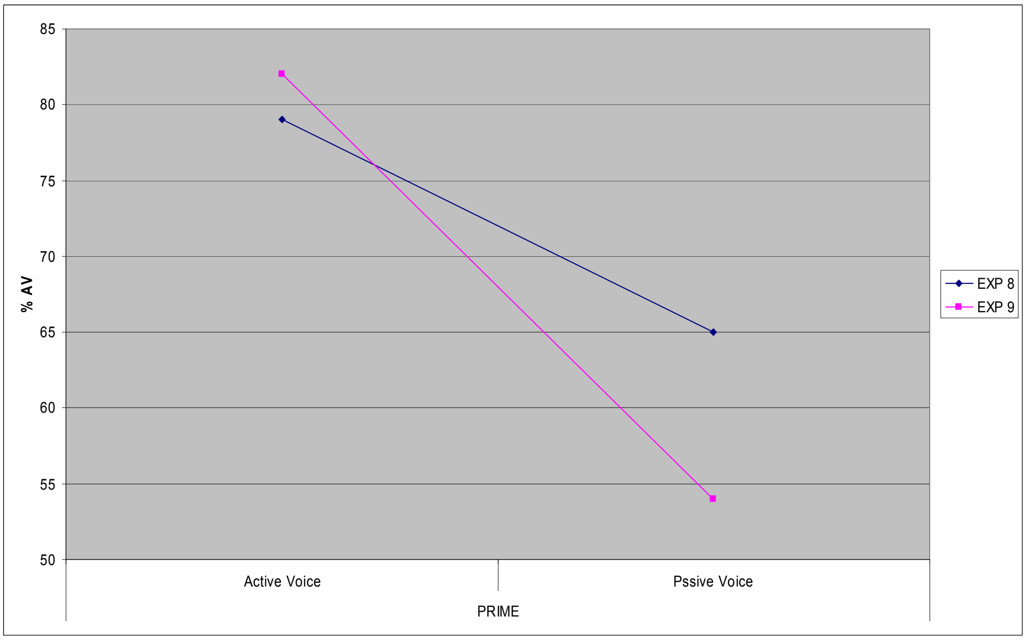


Figure 36

Experiments 8 and 9. First fixation latencies to the non-cued referent. Experiment x Verb Match interaction.

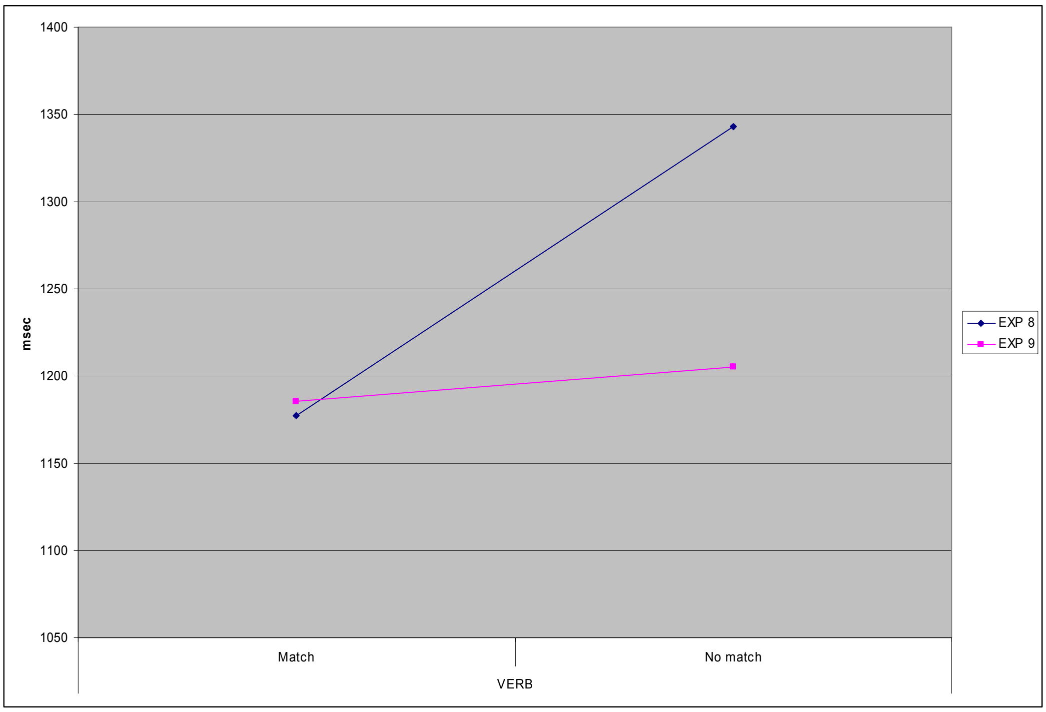


Figure 37

Experiment 9. Eye-voice spans. Cue x Prime interaction. Object

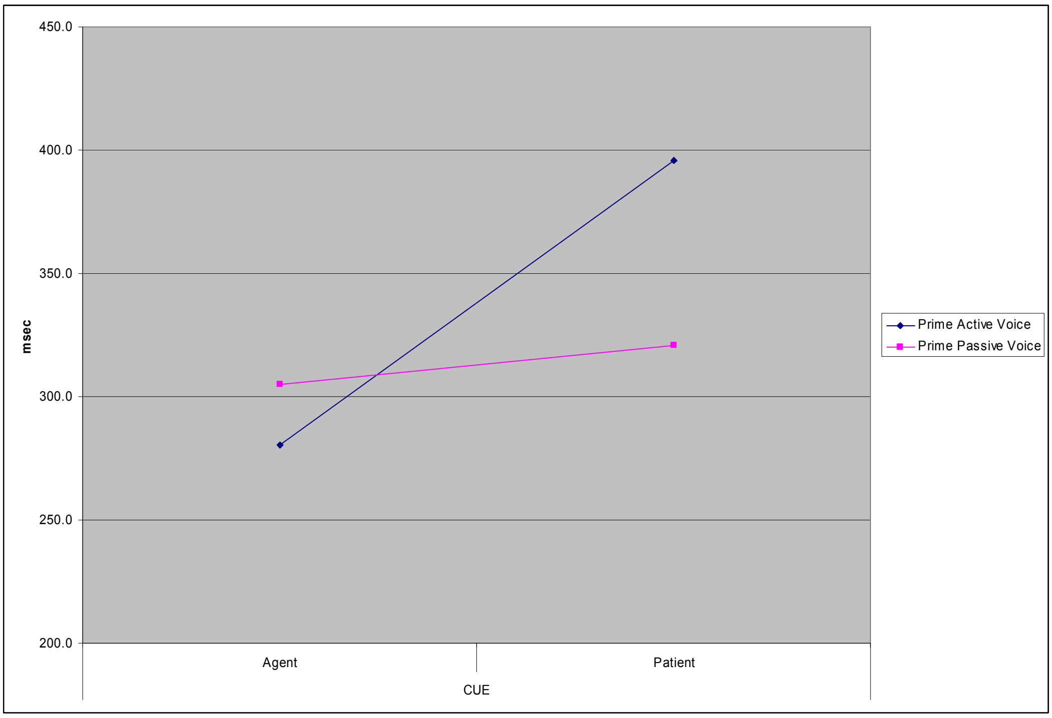
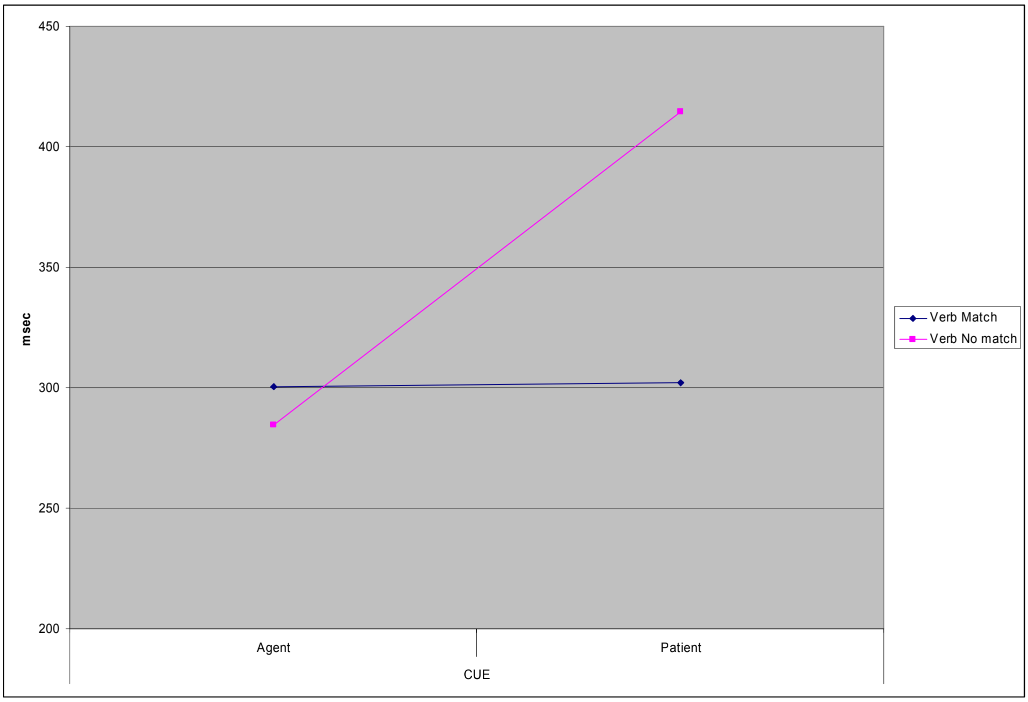


Figure 38

Experiment 9. Eye-voice spans. Cue x Verb Match interaction. Object



Appendix II. Tables

Table 1
Examples of materials from Turner & Rommetveit (1968)

non-reversible actives	*The bunny was eating the carrot*
reversible actives	*The mommy was kissing the daddy*
non-reversible passives	*The carrot was eaten by the bunny*
reversible passives	*The daddy was kissed by the mommy*

Table 2
Experiment 1. Word order distribution

Word order	Construction	Total		The agent-cueing condition		The patient-cueing condition	
		#	%	#	%	#	%
SVO							
S[N]-V-O[N]	*Sinja-ja (ryb-a)s'ela krasnu-ju (ryb-u)*						
	Blue fish [N(Nom)] ate red fish [N(Acc)]	231	48.1	114	47.5	117	48.8
S[N]-V-O[Pr]	*Sinja-ja (ryb-a)s'ela ee*						
	Blue fish [N(Nom)] ate her [Pr(Acc)]	2	.4	0	0	2	.8
S[Pr]-V-O[N]	*Ona s'ela krasnu-ju (ryb-u)*						
	She [Pr(Nom)] ate red fish [N(Acc)]	74	15.4	74	30.9	0	0
S[Pr]-V-O[Pr]	*Ona s'ela ee*						
	She [Pr(Nom)] ate red fish [Pr(Acc)]	31	6.5	31	12.9	0	0
SOV							
S[N]- O[Pr]-V	*Sinja-ja (ryb-a) ee s'ela*						
	Blue fish [N(Nom)] her [Pr(Acc)] ate	37	7.7	5	2.1	32	13.3
S[Pr]- O[Pr]-V	*Ona ee s'ela*						
	She [N(Nom)] her [Pr(Acc)] ate	12	2.5	11	4.6	1	.4
OVS							
O[N]-V- S[N]	*Krasnu-ju (ryb-u) s'ela sinja-ja (ryb-a)*						
	Red fish [N(Acc)] ate blue fish [N(Nom)]	32	6.7	4	1.7	28	11.7
O[Pr]-V- S[N]	*Ee s'ela sinja-ja (ryb-a)*						
	Her [Pr(Acc)] ate blue fish [N(Nom)]	40	8.3	0	0	40	16.7
Passive Voice							
SVO	*Krasna-ja (ryb-a) byla s'edena sin-ej (ryb-oj)*	18	3.8	0	0	18	7.5
	Red fish [N(Nom)] was eaten by blue fish [N(Inst)]						

Table 3
Experiment 1. Word order by the specified word order groups.

Word order		agent-cueing trials		patient-cueing trials	
Agent-first patterns	Observed	235	49.4	152	31.9
	Expected	240	50	0	0
Patient-first patterns	Observed	4	.6	86	18.1
	Expected	0	0	240	50

Table 4
Experiment 1. Error rates.

	Agent-initial sentences		Patient-initial sentences	
	#	%	#	%
Total	239	100	239	100
The agent-cueing trials	1	.418	0	0
the patient-cueing trials	15	6.276	1	.418

Table 5
Experiment 1. Speech onset latencies

	Agent-initial sentences (msec.)	Patient-initial sentences (msec.)
The agent-cueing trials	297	–
the patient-cueing trials	479	378

Table 6.
Experiments 2 and 3. Average number of syllables: Naming task.

	Name
English	1.9
Russian	2.3

Table 7.
Experiments 2 and 3. Average number of syllables: Event Description task.

	Subject	Verb	Object
English	1.88	2.0	1.88
Russian	2.5	2.4	2.9

Table 8.
Experiments 2 and 3. Word order.

	Active Voice/SVO (%)	Passive Voice (%)	OVS (%)	Other (%)
English	97.75	1.25	-	1.0
Russian	96.4	0	1.5	2.1

Table 9.
Experiment 4. Word order.

	Word Order (%) Cue	
Word order	Agent	Patient
SVO	97	96
Other	3	4

Table 10.
Experiment 4. First fixation onset latencies

	First fixation onset latencies (msec) Cue	
Referent	Agent	Patient
Agent	151	402
Patient	533	148

Table 11.
Experiment 4. Eye-voice spans.

	Eye-Voice spans (msec) Cue		
Constituent	Agent	Patient	**MEAN**
Subject	439	459	**449**
Object	333	345	**339**
MEAN	**386**	**402**	

Table 12.
Experiment 4. Name onset latencies.

	Name onset latencies (msec) Cue	
	agent	patient
Constituent		
Subject	1335	1364
Verb	1769	1931
Object	2346	2467

Table 13.
Experiment 5. Word order

	Word Oder (%) Cue		
	Agent	Theme	Patient
Word Order			
SVTO	86	86	87
SVOT	14	14	13

Table 14.
Experiment 5. First fixation onset latencies

	First Fixation latencies (msec) Cue		
	Agent	Theme	Patient
Referent			
Agent	153	394	442
Theme	414	158	395
Patient	837	465	150

Table 15.
Experiment 5. Eye-voice spans.

	Eye-Voice spans (msec) Cue			
Constituent	Agent	Theme	Patient	**MEAN**
Subject	477	516	510	**501**
Verb	441	560	456	**486**
Theme	749	742	668	**720**
Object	403	333	345	**360**
MEAN	**518**	**538**	**495**	

Table 16.
Experiment 5. Eye-voice spans for Theme and Object between structures.

			Eye-Voice spans (msec) Cue		
		Constituent	Agent	Theme	Patient
Structure	SVTO	Theme	743	745	660
		Object	351	394	356
	SVOT	Theme	772	716	719
		Object	833	765	735

Table 17.
Experiment 5. Name onset latencies.

	Name onset latencies (msec) Cue		
Constituent	Agent	Theme	Patient
Subject	1458	1488	1458
Verb	1978	2008	1987
Object 1	2522	2551	2478
Object 2	3107	3144	3075

Table 18.
Word order choice. Meta-analysis.

	Tomlin (1995) English, explicit cue	Experiment 1 (this thesis) Russian, explicit cue	Gleitman, et al. (in press) English, implicit cue	Experiment 4 (this thesis) Finnish, implicit cue
Agent-driven structures/agent-cueing condition	~100%	~100%	~85%	~100%
Agent-driven structures/patient-cueing condition	~0%	~65%	~75%	~100%

Table 19
Experiment 6. Cueing efficiency.

	Cueing efficiency (%)			
	Cue SOA			
	Implicit		Explicit	
	Instruction			
Cue location	Non-constraining	Constraining	Non-constraining	Constraining
Agent	87	88	84	91
Patient	79	83	81	94

Table 20
Experiment 6. Word order choice.

	Active Voice (%)			
	Cue SOA			
	Implicit		Explicit	
	Instruction			
Cue location	Non-constraining	Constraining	Non-constraining	Constraining
Agent	93	98	94	97
Patient	77	43	70	36

Table 21.
Experiment 6. ANOVA. Word order choice.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	35.140	.000	1, 60	50.842	.000
Cue SOA (B)	1, 23	11.385	.003	1, 60	3.712	.059
Instruction (C)	1, 23	16.731	.000	1, 60	67.471	.000
Interaction AxB	1, 23	5.948	.023	1, 60	4.305	.042
Interaction AxC	1, 23	22.399	.000	1, 60	106.882	.000
Interaction BxC	1, 23	ns	ns	1, 60	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 22
Experiment 6. First fixation onsets.

	First fixation onsets (msec)							
	Cue SOA							
	Implicit				Explicit			
	Instruction							
	Non-constraining		Constraining		Non-constraining		Constraining	
	Cue Location							
Referent	agent	patient	agent	patient	agent	patient	agent	patient
Agent	208	722	254	-	215	954	240	-
Patient	1221	215	-	248	1727	217	-	243

Table 23.
Experiment 6. ANOVA First fixation onsets to the cued referent.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	ns	ns	1, 60	ns	ns
Cue SOA (B)	1, 23	ns	ns	1, 60	ns	ns
Instruction (C)	1, 23	17.500	.000	1, 60	78.784	.000
Interaction AxB	1, 23	ns	ns	1, 60	ns	ns
Interaction AxC	1, 23	ns	ns	1, 60	ns	ns
Interaction BxC	1, 23	ns	ns	1, 60	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 24.
Experiment 6. ANOVA First fixation onsets to the non-cued referent in non-constraining condition.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	39.317	.000	1, 63	38.599	.000
Cue SOA (B)	1, 23	152.095	.000	1, 63	99.457	.000
Interaction AxB	1, 23	7.658	.011	1, 63	ns	ns

Table 25.
Experiment 6. Eye-voice spans.

	Eye-Voice spans (msec)							
	Cue SOA							
	Implicit				Explicit			
	Instruction							
	Non-constraining		Constraining		Non-constraining		Constraining	
	Cue Location							
Constituent	agent	patient	agent	patient	agent	patient	agent	patient
Subject	469	407	650	920	422	360	580	970
Object	496	457	2550	2049	473	460	2328	1998

Table 26
Experiment 6. Eye-voice spans in Passive Voice sentences.

	Eye-voice spans. Passive Voice (msec)	
Conditions	Subject	Object
Non-constrain/AC/implicit 6 participants	500	212
Non-constrain/AC/explicit 5 participants	390	482
Non-constrain/PC/implicit 15 participants	388	569
Non-constrain/PC/explicit 14 participants	448	537
MEAN	**432**	**450**

Table 27

Experiment 6. ANOVA Eye-voice spans. Subject

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	9.936	.004	1, 63	5.529	.022
Cue SOA (B)	1, 23	ns	ns	1, 63	ns	ns
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns

Table 28

Experiment 6. Name onset latencies.

	Name onset latencies (msec)							
	Cue SOA							
	Implicit				Explicit			
	Instruction							
	Non-constraining		Constraining		Non-constraining		Constraining	
	Cue Location							
Constituent	agent	patient	agent	patient	agent	patient	agent	patient
Subject	1423	1645	1498	1908	1427	1538	1352	1565
Verb	1981	2220	2100	2586	1997	2085	1966	2268
Object	2628	2845	2948	3311	2615	2649	2870	3085

Table 29

Experiment 6. Name onset latencies in Passive Voice sentences.

	Name onset latencies. Passive Voice (msec)		
Conditions	Subject	Verb	Object
Non-constrain/AC/implicit 6 participants	1425	1967	2464
Non-constrain/AC/explicit 5 participants	1752	2374	3135
Non-constrain/PC/implicit 15 participants	1578	2264	2968
Non-constrain/PC/explicit 14 participants	1656	2273	2869
Constrain/AC/implicit 1 participant	1533	1997	2744
Constrain/AC/explicit 2 participants	2592	3289	3945
Constrain/PC/implicit 20 participants	1780	2496	3267
Constrain/PC/explicit 20 participants	1543	2236	3214

Table 30
Experiment 6. ANOVA Name onset latencies to produce Subject.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	42.768	.000	1, 60	15.593	.000
Cue SOA (B)	1, 23	42.640	.000	1, 60	5.995	.017
Instruction (C)	1, 23	ns	ns	1, 60	5.944	.018
Interaction AxB	1, 23	10.754	.003	1, 60	ns	ns
Interaction AxC	1, 23	5.381	.030	1, 60	5.957	.018
Interaction BxC	1, 23	7.358	.012	1, 60	10.573	.002
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 31
Experiment 6. ANOVA Name onset latencies to produce Verb.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	56.993	.000	1, 60	16.017	.000
Cue SOA (B)	1, 23	31.636	.000	1, 60	4.196	.045
Instruction (C)	1, 23	9.244	.006	1, 60	23.602	.000
Interaction AxB	1, 23	6.761	.016	1, 60	ns	ns
Interaction AxC	1, 23	10.107	.004	1, 60	12.387	.001
Interaction BxC	1, 23	4.302	.049	1, 60	6.381	.014
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 32
Experiment 6. ANOVA Name onset latencies to produce Object.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	45.129	.000	1, 60	6.323	.015
Cue SOA (B)	1, 23	13.726	.000	1, 60	ns	ns
Instruction (C)	1, 23	28.664	.000	1, 60	62.783	.000
Interaction AxB	1, 23	4.106	.054	1, 60	ns	ns
Interaction AxC	1, 23	ns	ns	1, 60	ns	ns
Interaction BxC	1, 23	ns	ns	1, 60	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 33
Experiment 7. Cueing efficiency.

	Cueing efficiency (%)			
	Cue SOA			
	Implicit		Explicit	
	Instruction			
Cue location	Non-constraining	Constraining	Non-constraining	Constraining
Agent	89	91	97	96
Patient	77	87	95	94

Table 34
Experiment 7. Word order choice.

	Active Voice (%)			
	Cue SOA			
	Implicit		Explicit	
	Instruction			
Cue location	Non-constraining	Constraining	Non-constraining	Constraining
Agent	91	90	93	96
Patient	82	46	72	31

Table 35.
Experiment 7. ANOVA. Word order.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	87.650	.000	1, 60	180.642	.000
Cue SOA (B)	1, 23	4.944	.036	1, 60	2.897	.094
Instruction (C)	1, 23	15.136	.001	1, 60	178.144	.000
Interaction AxB	1, 23	14.301	.001	1, 60	10.415	.002
Interaction AxC	1, 23	20.763	.000	1, 60	183.957	.000
Interaction BxC	1, 23	ns	ns	1, 60	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 36
Experiment 7. First fixation onsets.

	First fixation onsets (msec)							
	Cue SOA							
	Implicit				Explicit			
	Instruction							
	Non-constraining		Constraining		Non-constraining		Constraining	
	Cue Location							
Referent	agent	patient	agent	patient	agent	patient	agent	patient
Agent	223	599	266	-	207	1204	236	-
Patient	690	220	-	238	1368	234	-	245

Table 37
Experiment 7. ANOVA First fixation onsets to the cued referent.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	ns	ns	1, 60	ns	ns
Cue SOA (B)	1, 23	ns	ns	1, 60	ns	ns
Instruction (C)	1, 23	20.858	.000	1, 60	52.095	.000
Interaction AxB	1, 23	ns	ns	1, 60	ns	ns
Interaction AxC	1, 23	ns	ns	1, 60	ns	ns
Interaction BxC	1, 23	ns	ns	1, 60	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 38
Experiment 7. ANOVA First fixation onsets to the non-cued referent in non-constraining condition.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	4.291	.05	1, 63	5.787	.019
Cue SOA (B)	1, 23	166.382	.000	1, 63	147.480	.000
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns

Table 39
Experiment 7. Eye-voice spans.

	Eye-Voice spans (msec)							
	Cue SOA							
	Implicit				Explicit			
	Instruction							
	Non-constraining		Constraining		Non-constraining		Constraining	
	Cue Location							
Constituent	agent	patient	agent	patient	agent	patient	agent	patient
Subject	528	457	680	841	537	484	709	932
Object	440	429	2328	2227	396	377	2027	2016

Table 40
Experiment 7. Eye-voice spans in Passive Voice sentences.

	Eye-voice spans. Passive Voice	
Conditions	Subject	Object
Non-constrain/AC/implicit 9 participants	595	456
Non-constrain/AC/explicit 8 participants	431	414
Non-constrain/PC/implicit 12 participants	425	685
Non-constrain/PC/explicit 15 participants	600	322
MEAN	**513**	**469**

Table 41
Experiment 7. ANOVA Eye-voice spans. Subject.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	5.723	.025	1, 60	4.202	.045
Cue SOA (B)	1, 23	ns	ns	1, 60	ns	ns
Interaction AxB	1, 23	ns	ns	1, 60	ns	ns

Table 42
Experiment 7. Name onset latencies.

	Name onset latencies (msec)							
	Cue SOA							
	Implicit				Explicit			
	Instruction							
	Non-constraining		Constraining		Non-constraining		Constraining	
	Cue Location							
	agent	patient	agent	patient	agent	patient	agent	patient
Constituent								
Subject	1503	1707	1666	1922	1311	1390	1296	1468
Verb	2047	2275	2267	2595	1895	1960	1938	2180
Object	2677	2947	3086	3374	2501	2609	2687	2999

Table 43
Experiment 7. Name onset latencies in Passive Voice sentences.

	Name onset latencies. Passive Voice		
Conditions	Subject	Verb	Object
Non-constrain/AC/implicit 9 participants	1709	2377	2960
Non-constrain/AC/explicit 8 participants	1463	2255	2905
Non-constrain/PC/implicit 12 participants	1850	2485	3198
Non-constrain/PC/explicit 15 participants	1439	2164	2943
Constrain/AC/implicit 10 participant	2028	2750	3490
Constrain/AC/explicit 7 participants	1453	2275	2944
Constrain/PC/implicit 22 participants	1933	2757	3525
Constrain/PC/explicit 23 participants	1388	2217	3155

Table 44
Experiment 7. ANOVA Name onset latencies to produce Subject.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	27.655	.000	1, 60	12.636	.001
Cue SOA (B)	1, 23	88.798	.000	1, 60	44.256	.000
Instruction (C)	1, 23	8.872	.007	1, 60	14.076	.000
Interaction AxB	1, 23	ns	ns	1, 60	ns	ns
Interaction AxC	1, 23	ns	ns	1, 60	ns	ns
Interaction BxC	1, 23	6.953	.015	1, 60	7.122	.010
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 45
Experiment 7. ANOVA Name onset latencies to produce Verb.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	37.812	.000	1, 60	15.046	.000
Cue SOA (B)	1, 23	64.486	.000	1, 60	29.695	.000
Instruction (C)	1, 23	22.572	.000	1, 60	36.221	.000
Interaction AxB	1, 23	ns	ns	1, 60	ns	ns
Interaction AxC	1, 23	ns	ns	1, 60	ns	ns
Interaction BxC	1, 23	3.575	.071	1, 60	4.344	.041
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 46
Experiment 7. ANOVA Name onset latencies to produce Object.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Cue location (A)	1, 23	49.681	.000	1, 60	11.368	.001
Cue SOA (B)	1, 23	54.876	.000	1, 60	19.735	.000
Instruction (C)	1, 23	38.264	.000	1, 60	27.138	.000
Interaction AxB	1, 23	ns	ns	1, 60	ns	ns
Interaction AxC	1, 23	ns	ns	1, 60	ns	ns
Interaction BxC	1, 23	ns	ns	1, 60	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 60	ns	ns

Table 47.
Experiments 6 and 7. Name onset latencies. Experiment x Cue SOA interaction

	df_1	F_1	p_1	df_2	F_2	p_2
Subject	1, 47	5.398	.037	1, 31	22.947	.002
Verb	1, 47	7.614	.010	1, 31	34.037	.001
Object	1, 47	9.983	.009	1, 31	39.952	.001

Table 48.
Experiments 8 and 9. Priming Materials.

Verb match

AC/AV/VM		
CHASE	1	The mouse is chasing the octopus
KICK	2	The stewardess is kicking the pilot
PUNCH	3	The student is punching the driver
PUSH	4	The elephant is pushing the cow
PULL	5	Two fishermen are pulling the shark
SHOOT	6	The turtle is shooting the dog
SCOLD	7	Parents are scolding the child
TOUCH	8	The baby is touching the book
AC/PV/VM		
CHASE	1	Then elephant is being chased by the cow
KICK	2	The boy is being kicked by the girl
PUNCH	3	The nurse is being punched by the patient
PUSH	4	The worker is being pushed by the foreman
PULL	5	The dog is being pulled by the old lady
SHOOT	6	The doctor is being shot by the fireguard
SCOLD	7	Then astronaut is being scolded by his mother
TOUCH	8	The donkey is being touched by Shrek
PC/AV/VM		
CHASE	1	Cinderella is chasing her prince
KICK	2	Then ostrich is kicking the tourist
PUNCH	3	The wrestler is punching the gymnast
PUSH	4	Builders are pushing the cart
PULL	5	Then acrobat is pulling the manager
SHOOT	6	The terrorist is shooting the captain
SCOLD	7	The teacher is scolding the principle
TOUCH	8	The professor is touching the shop-assistant
PC/PV/VM		
CHASE	1	The turtle is being chased by the cat
KICK	2	The singer is being kicked by the pianist

PUNCH	3	The librarian is being punched by the policeman
PUSH	4	The child is being pushed by the teenager
PULL	5	The cow is being pulled by the milkmaid
SHOOT	6	The dog is being shot by the kangaroo
SCOLD	7	The shopper is being scolded by the shop manager
TOUCH	8	Then engineer is being touched by the surgeon

Verb no match

AC/AV/VN

CHASE	1	The boy is building the house
KICK	2	The rat is tickling the cat
PUNCH	3	The drummer is buying the piano
PUSH	4	Then orphan is robbing the bank
PULL	5	The thief is reading the book
SHOOT	6	The squirrel is pulling the elephant
SCOLD	7	The bartender is breaking the table
TOUCH	8	The student is stealing the candy

AC/PV/VN

CHASE	1	The bank is being robbed by schoolchildren
KICK	2	The car is being driven by the dog
PUNCH	3	The boat is being built by the professor
PUSH	4	The donkey is being killed by the savage
PULL	5	The tree is being cut by the beaver
SHOOT	6	The pigeon is being eaten by the chef
SCOLD	7	The man is being beaten by the athlete
TOUCH	8	The dog is being buried by the magician

PC/AV/VN

CHASE	1	The witness is saving the robber
KICK	2	The secretary is washing the car
PUNCH	3	The manager is stealing the diamond ring
PUSH	4	The soldier is killing the officer
PULL	5	The peasant is riding the elephant
SHOOT	6	The janitor is writing the book
SCOLD	7	The nurse is flying the helicopter
TOUCH	8	The buffalo is catching the kangaroo

PC/PV/VN

CHASE	1	The whale is being killed by the Eskimo
KICK	2	The boat is being sunk by the submarine
PUNCH	3	The eagle is being shot by the librarian
PUSH	4	The stone is being crushed by the hammer
PULL	5	The house is being torn down by builders
SHOOT	6	Pinocchio is being robbed by hooligans
SCOLD	7	Alice is being educated by her mother
TOUCH	8	The computer is being bought by the immigrant

Table 49.
Experiments 8 and 9. Filler Materials.

1.	The moon is over the sun
2.	The mouse is under the computer
3.	The house is below of the convenience store
4.	The fish is right of the vase
5.	The picture is under the table
6.	The rabbit is under the blanket
7.	The book is on the shelf
8.	The desk is left of the window
9.	The octopus is over the sun
10.	The match is right of the box
11.	The box is left of the shoes
12.	The portrait is above the bed
13.	The acrobat is over the roof
14.	The rat is under the floor
15.	The hat is right of the coat
16.	The stereo is left of the fireplace
17.	The flowers are under the magazine
18.	The ash-tray is on the table
19.	The room is left of the entrance
20.	The athlete is over the nurse
21.	The eagle is right of the mountain top
22.	The shark is far from the house
23.	The head is right of the body
24.	The ice-cream is far from the fridge
25.	The folder is right of the pencil
26.	The window is left of the door
27.	The gun is under the pillow
28.	The elevator is right of the reception
29.	The star is left of the satellite
30.	The car is near the garage
31.	The fly is above the jam
32.	The patient is under the blanket
33.	The prince is right of the queen
34.	The thief is left of the nurse
35.	The giraffe is far from the zoo
36.	The sun is over the moon
37.	The monitor is left of the keyboard
38.	The cup is right of the sink
39.	The boy is under the airplane
40.	The spy is undercover
41.	The space-shuttle is far from the moon
42.	The door is right of the letter-box
43.	The suitcase is next to the backpack
44.	The photograph is above the calendar
45.	The toaster is right of the salt shaker
46.	The pier is near the marina
47.	The drawing is left of the text box

48. The piano is right of the drum-set
49. The leg is right of the wrecked bike
50. The boat is under water
51. The bus is left of the parking lot
52. The apple is under the orange
53. The bag is right of the lamp
54. The monkey is under the deck
55. The ship is far from the shore
56. The goat is right of the donkey
57. The actor is under the director
58. The mother is right of the father
59. The climber is over the ridge
60. The slug is under the leaves
61. The tree is left of the rock
62. The boat is right of the lighthouse
63. The stapler is on the desk
64. The wall is left of the door
65. The hanger is under the coat
66. The radio is left of the table clock
67. The brewery is near the hotel
68. The dog is right of the kennel
69. The drugstore is under the bridge
70. The wolf is on top of the hill
71. The newspaper is on the shelf
72. The bottle is left of the fridge
73. The typewriter is right of the computer
74. The photograph is above the bed
75. The gloves are under the magazine
76. The cigarette pack is on the table
77. The clown is on the stage
78. The mouse is under the floor
79. The helicopter is right of the mountain
80. The train is at the station
81. The closet is left of the wardrobe
82. The manager is behind the office assistant
83. The leg is left of the arm
84. The fudge cake is on the counter
85. The mug is right of the glass
86. The locker is left of the door
87. The letter is under the book
88. The bucket is right of the broom
89. The kiosk is left of the intersection
90. The truck is near the gas station
91. The ant is above the fly
92. The body is under the cover
93. The army chief is right of the princess
94. The thug is left of the policeman
95. The lion is far from the zoo
96. The planet is over the star
97. The scanner is left of the power box

98. The sauce pan is right of the sink
99. The comedian is under the projector
100. The sister is right of the brother
101. The bus is far from the city
102. The monkey is right of the banana
103. The platypus is over the log
104. The note is right of the pencil
105. The bag is next to the shoe-box
106. The drawing is above the photograph
107. The microscope is under the shelf
108. The pliers are next to the screwdriver
109. The speakers are right of the comb
110. The sea port is near the bus station
111. The ball is left of the baseball bat
112. The guitar is right of the cello
113. The eyeliner is right of the powder box
114. The scooter is under the rain
115. The toy is under the trousers
116. The article is on the notebook
117. The motorcycle is left of the parking lot
118. The kiwi is under the grapefruits
119. The backpack is right of the hiking shoes
120. The pigeon is under the roof
121. The ocean liner is far from the harbour
122. The cow is right of the sheep
123. The rescuer is over the mountain top
124. The bug is under the leaves
125. The bush is left of the gate
126. The bay is right of the lighthouse
127. The cigars are under the coat
128. The highlighter is next to the pencil sharpener
129. The cup is right of the plate
130. The novel is near the end

Table 50
Experiment 8. Cueing efficiency.

	Cueing efficiency (%)			
	Syntactic Prime			
	Active voice		Passive voice	
	Verb Overlap			
	no match	match	no match	match
Cue location				
Agent	97	90	97	91
Patient	92	91	94	94

Table 51
Experiment 8. Word order.

	Active Voice (%)			
	Syntactic Prime			
	Active voice		Passive voice	
	Verb Overlap			
	No match	Match	No match	Match
Cue location				
Agent	94	84	84	64
Patient	67	71	60	51

Table 52
Experiment 8. ANOVA. Word order.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	5.916	.023	1, 63	47.232	.000
Cue location (B)	1, 23	9.611	.005	1, 63	78.451	.000
Verb match (C)	1, 23	16.393	.000	1, 63	16.139	.000
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns
Interaction AxC	1, 23	3.226	.086	1, 63	6.670	.012
Interaction BxC	1, 23	11.939	.002	1, 63	7.978	.007
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 53
Experiment 8. First fixation onsets.

	First fixation onsets (msec)							
	Syntactic Prime							
	Active voice				Passive Voice			
	Cue Location							
	Agent		Patient		Agent		Patient	
	Verb Overlap							
	no match	match	no match	match	no match	match	no match	match
Referent								
Agent	206	212	1217	1055	216	205	1213	1181
Patient	1405	1127	207	206	1410	1229	204	208

Table 54
Experiment 8. ANOVA. First fixation onsets to the non-cued referents.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	ns	ns	1, 63	ns	ns
Cue location (B)	1, 23	11.268	.003	1, 63	3.518	.066
Verb match (C)	1, 23	25.698	.000	1, 63	5.862	.019
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns
Interaction AxC	1, 23	ns	ns	1, 63	ns	ns
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 55
Experiment 8. Eye-voice spans.

	Eye-voice spans (msec)							
	Syntactic Prime							
	Active voice				Passive Voice			
	Cue Location							
	Agent		Patient		Agent		Patient	
	Verb Overlap							
Constituent	no match	match	no match	match	no match	match	no match	match
Subject	395	359	483	364	386	368	345	327
Verb	1059	892	918	771	746	667	943	972
Object	301	341	593	289	325	342	321	304

Table 56
Experiment 8. Eye-voice spans in Passive Voice sentences.

	Eye-voice spans. Passive Voice. msec	
Conditions	Subject	Object
AV/AC/VN	356	364
AV/AC/VM	285	209
AV/PC/VN	461	501
AV/PC/VM	296	283
Mean AV	*350*	*339*
PV/AC/VN	378	272
PV/AC/VM	320	303
PV/PC/VN	317	310
PV/PC/VM	326	284
Mean PV	*335*	*292*

Table 57
Experiment 8. ANOVA Eye-voice spans. Subject.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	ns	ns	1, 63	ns	ns
Cue location (B)	1, 23	ns	ns	1, 63	ns	ns
Verb match (C)	1, 23	12.196	.002	1, 63	3.831	.05
Interaction AxB	1, 23	9.842	.004	1, 63	ns	ns
Interaction AxC	1, 23	8.093	.008	1, 63	ns	ns
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 58
Experiment 8. ANOVA Eye-voice spans. Object.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	ns	ns	1, 63	ns	ns
Cue location (B)	1, 23	ns	ns	1, 63	ns	ns
Verb match (C)	1, 23	15.378	.001	1, 63	ns	ns
Interaction AxB	1, 23	18.368	.000	1, 63	ns	ns
Interaction AxC	1, 23	19.704	.000	1, 63	4.526	.038
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 59
Experiment 8. Name onset latencies.

	Name onset latencies (msec)							
	Syntactic Prime							
	Active voice				Passive Voice			
	Cue Location							
	Agent		Patient		Agent		Patient	
	Verb Overlap							
Constituent	no match	match	no match	match	no match	match	no match	match
Subject	1469	1338	1406	1420	1421	1323	1491	1423
Verb	2060	1907	1976	1974	2001	1882	2124	2017
Object	2665	2513	2599	2621	2620	2488	2759	2647

Table 60
Experiment 8. Name onset latencies in Passive Voice sentences.

	Name onset latencies. Passive Voice		
	Subject	Verb	Object
Conditions			
AV/AC/VN	1698	2299	3140
AV/AC/VM	1746	2294	2945
AV/PC/VN	1610	2213	2918
AV/PC/VM	1588	2180	2894
MEAN AV	*1661*	*2247*	*2974*
PV/AC/VN	1711	2381	3119
PV/AC/VM	1447	2029	2699
PV/PC/VN	1555	2182	2910
PV/PC/VM	1462	2058	2696
MEAN PV	*1544*	*2162*	*2856*

Table 61
Experiment 8. ANOVA Name onset latencies to produce Subject.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	ns	ns	1, 63	ns	ns
Cue location (B)	1, 23	11.471	.003	1, 63	ns	ns
Verb match (C)	1, 23	5.228	.032	1, 63	ns	ns
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns
Interaction AxC	1, 23	ns	ns	1, 63	ns	ns
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 62
Experiment 8. ANOVA Name onset latencies to produce Verb

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	4.858	.038	1, 63	ns	ns
Cue location (B)	1, 23	14.423	.001	1, 63	ns	ns
Verb match (C)	1, 23	8.669	.007	1, 63	ns	ns
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns
Interaction AxC	1, 23	ns	ns	1, 63	ns	ns
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 63
Experiment 8. ANOVA Name onset latencies to produce Object.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	ns	ns	1, 63	ns	ns
Cue location (B)	1, 23	12.208	.002	1, 63	ns	ns
Verb match (C)	1, 23	9.341	.006	1, 63	ns	ns
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns
Interaction AxC	1, 23	ns	ns	1, 63	ns	ns
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 64
Experiment 9. Cueing efficiency.

	Cueing efficiency (%)			
	Syntactic Prime			
	Active voice		Passive voice	
	Verb Overlap			
Cue location	no match	match	no match	match
Agent	93	93	95	95.5
Patient	94	90	89	91

Table 65
Experiment 9. Word order.

	Active Voice (%)			
	Syntactic Prime			
	Active voice		Passive voice	
	Verb Overlap			
Cue location	No match	Match	No match	Match
Agent	93	94	77	50
Patient	68	75	53	36

Table 66
Experiment 9. ANOVA. Word order.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	42.543	.000	1, 63	186.196	.000
Cue location (B)	1, 23	9.420	.003	1, 63	99.388	.000
Verb match (C)	1, 23	24.053	.000	1, 63	18.652	.000
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns
Interaction AxC	1, 23	29.884	.000	1, 63	38.520	.000
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 67
Experiment 9. First fixation onsets.

	First fixation onsets (msec)							
	Syntactic Prime							
	Active voice				Passive Voice			
	Cue Location							
	Agent		Patient		Agent		Patient	
	Verb Overlap							
Referent	no match	match	no match	match	no match	match	no match	match
Agent	202	203	1101	1107	214	213	1125	1141
Patient	1218	1246	206	200	1209	1013	203	207

Table 68
Experiment 9. Eye-voice spans.

	Eye-voice spans (msec)							
	Syntactic Prime							
	Active voice				Passive Voice			
	Cue Location							
	Agent		Patient		Agent		Patient	
	Verb Overlap							
Constituent	no match	match	no match	match	no match	match	no match	match
Subject	349	402	447	363	360	344	344	285
Verb	940	1068	694	599	815	733	720	723
Object	257	304	487	305	313	297	342	299

Table 69
Experiment 9. Eye-voice spans in Passive Voice sentences.

Conditions	Eye-voice spans. Passive Voice. msec: Subject	Object
AV/AC/VN	301	253
AV/AC/VM	219	196
AV/PC/VN	446	567
AV/PC/VM	380	381
Mean	*337*	*349*
PV/AC/VN	265	321
PV/AC/VM	301	291
PV/PC/VN	339	407
PV/PC/VM	277	299
Mean	*296*	*330*

Table 70
Experiment 9. ANOVA Eye-voice spans. Object.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	ns	ns	1, 63	ns	ns
Cue location (B)	1, 23	17.585	.000	1, 63	ns	ns
Verb match (C)	1, 23	7.123	.014	1, 63	ns	ns
Interaction AxB	1, 23	17.390	.000	1, 63	ns	ns
Interaction AxC	1, 23	ns	ns	1, 63	ns	ns
Interaction BxC	1, 23	14.421	.001	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 71
Experiment 9. Name onset latencies.

	Name onset latencies (msec)							
	Syntactic Prime							
	Active voice				Passive Voice			
	Cue Location							
	Agent		Patient		Agent		Patient	
	Verb Overlap							
Constituent	no match	match	no match	match	no match	match	no match	match
Subject	1273	1178	1366	1259	1354	1246	1416	1272
Verb	1933	1758	1982	1876	1970	1845	2095	1908
Object	2604	2459	2673	2572	2662	2538	2836	2586

Table 72
Experiment 9. Name onset latencies in Passive Voice sentences.

	Name onset latencies. Passive Voice		
	Subject	Verb	Object
Conditions			
AV/AC/VN	1622	2398	3206
AV/AC/VM	1730	2390	3511
AV/PC/VN	1544	2189	2883
AV/PC/VM	1373	2087	2929
MEAN AV	*1567*	*2266*	*3132*
PV/AC/VN	1541	2251	2990
PV/AC/VM	1409	1996	2728
PV/PC/VN	1591	2369	3174
PV/PC/VM	1399	2050	2777
MEAN PV	*1485*	*2166*	*2917*

Table 73
Experiment 9. ANOVA Name onset latencies to produce Subject.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	10.496	.004	1, 63	ns	ns
Cue location (B)	1, 23	7.127	.014	1, 63	ns	ns
Verb match (C)	1, 23	17.848	.000	1, 63	3.867	.054
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns
Interaction AxC	1, 23	ns	ns	1, 63	ns	ns
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 74
Experiment 9. ANOVA Name onset latencies to produce Verb.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	15.075	.000	1, 63	ns	ns
Cue location (B)	1, 23	13.393	.001	1, 63	ns	ns
Verb match (C)	1, 23	22.677	.000	1, 63	4.023	.050
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns
Interaction AxC	1, 23	ns	ns	1, 63	ns	ns
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 75
Experiment 9. ANOVA Name onset latencies to produce Object.

Variance	df_1	F_1	p_1	df_2	F_2	p_2
Syntactic prime (A)	1, 23	11.548	.002	1, 63	ns	ns
Cue location (B)	1, 23	7.236	.013	1, 63	ns	ns
Verb match (C)	1, 23	19.054	.000	1, 63	6.567	.013
Interaction AxB	1, 23	ns	ns	1, 63	ns	ns
Interaction AxC	1, 23	ns	ns	1, 63	ns	ns
Interaction BxC	1, 23	ns	ns	1, 63	ns	ns
Interaction AxBxC	1, 23	ns	ns	1, 63	ns	ns

Table 76
Comparison of EVS data across experiments.

	Eye-voice spans (msec) Experiments							
	Experiment 2	Experiment 3	Experiment 4	Experiment 5	Experiment 6	Experiment 7	Experiment 8	Experiment 9
Language	English	Russian	Finnish	Finnish	English	English	English	English
Perceptual priming	no	no	weak	weak	weak/strong	weak/strong	strong	strong
Verb preview	no	no	yes	yes	yes	yes	yes/no	yes/no
Referential priming	no	no	no	no	no	yes	no	yes
Syntactic priming	no	no	no	no	no	no	yes	yes
Constituent								
Subject	421	757	449	501	415	502	379	362
Verb	570	998	-	486	-	-	-	-
Theme	-	-	-	720	-	-	-	-
Object	663	640	339	360	472	411	352	326