A parallel interface for language and cognition in sentence production: Theory, method, and experimental evidence

ANDRIY MYACHYKOV, MICHAEL I. POSNER AND
RUSSELL S. TOMLIN

Abstract

The debate about the place of linguistic theory in cognitive science encouraged by The Linguistic Review *is a good example of communication between different research communities. In this follow-up paper we (1) clarify our theoretical and methodological positions, (2) propose a theoretical model for language production similar to Jackendoff's* Parallel Architecture, *and (2) discuss emerging empirical evidence for this model. Our data suggest that perceptual, semantic, and syntactic information becomes available to the speaker in parallel providing competing production cues. Main architectural parameters of the proposed model are similar to Parallel Architecture, but we put a greater focus on the interface between language-specific and general cognitive domains. We view such interface as a regular mapping mechanism between the grammatical constraints imposed by the language system and the perceptual, semantic, and grammatical priming parameters available in the communicative environment.*

1. Theory and method in language research

In his seminal book, *Foundations of Language* (FL, 2002), Jackendoff proposed a very powerful and ambitious model of language organization. Jackendoff's *Parallel Architecture* has a great theoretical promise as it suggests a solid and empirically informed account of the language faculty that goes beyond the focus on competence common among generativists. As such, it incorporates and is consistent with, a great deal of experimental evidence about neural and psychological reality behind language accumulated in the fields of cognitive neuroscience and psycholinguistics. It also makes specific predictions about the chronometrical and functional properties of language-cognition

The Linguistic Review 24 (2007), 457–474
DOI 10.1515/TLR.2007.020
0167–6318/07/024-0457

interactions making it possible to model linguistic phenomena in their relation to the general domains of attention and memory.

We agree with Jackendoff that many contributors to the Special Issue of *The Linguistic Review* (TLR), including ourselves, chose an easy way to justify their critique of the syntactic bias of generative linguistics. It is true that most critics largely overlooked the motivation of Jackendoff's parallel architecture as a theory that addresses the variety of phonological, semantic, and pragmatic parameters of language as well as its relation to other cognitive domains in favor of criticizing classical Chomskian generativism. An attentive reader should not only notice that Jackendoff's proposal retains very little of classical Chomskian separation between competence and performance, but also that Jackendoff's model possesses important properties that help to reduce the level of encapsulation within the levels of language organization as well as to relate linguistic processes to less specialized cognitive domains. As such, *Parallel Architecture* has a much greater explanatory power and functional efficiency.

The publication of FL encouraged a discussion of the place of linguistic theory in our understanding of cognitive mechanisms of language organization, its evolution, and development in human mind. It may be useful to invert this question and ask what role psychological and neuroscientific explorations can play in developing a better theoretical account of language? Answering these questions will require both linguists who are ready to study psychology and statistical methods and psychologists who are not shy to take a course in theoretical linguistics. The special issue (TLR) was an important attempt to bring together researchers from fields usually separated by theoretical and methodological barriers. Hopefully, these barriers are not impenetrable. The discussion continues, and Jackendoff's reply to this special issue of TLR can indeed "play the role of a simulated conversation among its authors", both virtual and actual (Jackendoff 2007: 397). One prerequisite for the success of such discussion is proper understanding of the relationship between theory and experimental evidence. Is it all right for the same empirical fact to have multiple theoretical partners or is there just one truthful candidate? Should these two be always happily married, or do they have a right to partial separation, or divorce at any time? The first question is relatively easy. It is quite typical to have multiple theoretical accounts explaining the same psychological phenomenon. However, while some try to convince the scientific community by providing the support of experimental evidence, others abandon experimentation in favor of eloquent theorizing. The emergence of a dozen elaborate linguistic theories, largely introspective, normative, or observational in nature is exactly what led to alienation of the formal linguistics from the experimental psychology. This is probably true for both generative and cognitive linguistic traditions. While Ray Jackendoff's colleagues "refused to consider linguistics a part of psychology or even of science, because linguists don't do experiments" (Jackendoff

2007: 354), a famous linguist once replied: "You are a psychologist, go ahead and test my theory!" Both attitudes are very problematic as they reduce the basis for communication.

This leads us to the second issue. There is no clear-cut dichotomy between empirically informed and empirically uninformed theories. Rather, there is a continuum between empirically ignorant and empirically obsessed approaches at the extremes. Some theorists do not read experimental papers; some experimenters could not be bothered to understand formal theories. But there are at least two more types: those who attempt to incorporate experimental evidence provided by others without generating experimental programs of their own and those who follow a theory-experiment-theory dialectics and, in addition, are willing to discuss their ideas with theoretical formalists in *The Linguistic Review*. In our view, both positions are quite adequate in their interdisciplinary adequacy, but the former one still has a slight problem.

Another important question is the methodology used to relate theory to experimental research. One can use grammaticality judgments ("quick and dirty experiments" that is), employ quasi-experimental techniques like those used by Chafe (1980) or Slobin (Berman and Slobin 1994),[1] or conduct full-fledged experiments with naïve participants. These are all legitimate instruments of investigation, and a good theory requires the whole variety to increase its validity and explanatory power. However, the researcher has to recognize the limitations of the technique being used and its appropriateness for testing the theory.

From this point of view, Jackendoff's criticisms of TLR authors are not always correct. For example, he observes that Ferreira (2005) and Myachykov, Tomlin, and Posner (2005) "bemoan" alienation of psycholinguists from formal linguistics (Jackendoff 2007: 351). A simple answer is: "No, we don't". The relative encapsulation of these fields is the result of actions in both camps. As Ferreira (2005: 368-369) notes, psycholinguistic tradition derived in part from generative pursuit; the alienation came later, when mainstream generativism stopped to treat performance as "an important problem" for "a richer and more comprehensive theory of mind" (Chomsky 1965: 207) and because mainstream theory chose to ignore massive counterevidence against its claims (Jackendoff 2007: 357). Like Ferreira, we are more concerned with what could

1. Jackendoff incorrectly interprets our position with respect to methods used by Chafe and others. It is true that their work is "based on intuitive judgments and not on experimental evidence." (Jackendoff 2007: 354). As a matter of fact, we did not speak favorably or unfavorably of their work. In our review of methodology used in language research, we merely mentioned that their work has stimulated others to explore cognitive foundations of language. The same misunderstanding led Jackendoff to claim that our "animus (was) more toward mainstream generative grammar per se than toward its methodology" (354). Wrong again. We have no animus toward any methodology *per se*, but toward the lack of appreciation of the inferential limitations imposed by respective methodologies.

be done to bring the fields closer again, and what the necessary conditions of reunification should be.

The relationship between linguistics and cognitive neuroscience has a shorter history, but the methods employed by neuroscientists can nevertheless be very useful in improving our understanding of language. For example, the methods of neuroimaging (Posner and Raichle 1994) and sequencing the human genome (Venter et al. 2001) have been the driving forces in forging the alliance between cognition and neuroscience. These methods have made it possible to understand how complex tasks, some of which involve language, are implemented in the brain.

Probably the closest example of studying natural language in neuroimaging is studies of reading. Among other things, these studies 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, Cohen, and Dehaene 2003). In the right hemisphere, similar areas are involved in the perception and individuation of faces (Kanwisher, McDermott, and Chun 1997). While these areas were first thought to be word and face specific, more recent conceptualizations argue that they are more related to processes 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. Jackendoff recognizes that his crucial step 3b (2007: 393) may involve a more general working memory system rather than being specifically language oriented.

All of these findings support the idea that the brain localizes *processes* or *mental operations* not particular types of representation either linguistic or nonlinguistic.[2]

These findings have important implications for Jackendoff's effort to understand the psychological operations involved in language in terms of brain activity.

An understanding of a network of operations that produce language can also enhance Jackendoff's goals of finding specific genes and an evolutionary his-

2. Since the theoretical account that follows is more concerned with the issue of the interface between language and cognition, it is less important for us to join the argument about the nature of mental representations or the psychological reality behind them. It is noteworthy however that Jackendoff himself uses parallel terminology when he claims that "it has been important here to abandon the idea that linguistic entities in the brain are *symbols* or *representations.* We have instead been able to treat them simply as *structures* built of discrete combinatorial units." (Jackendoff 2002: 423)

tory for the development of language. Brain areas are associated with specific kinds of neurotransmitters and modulators can provide clues as to the genes involved. For example, searches based on these principles have yielded specific genes that influence attention (Posner, Rothbart, and Sheese 2007). Because of widespread genetic conservation and the extensive databases of gene functions that exist for various animal model systems, the location of specific genes can enhance the goal of an evolution of the mental operations involved in language.

The next Section will discuss recent data that provide important empirical support for the idea of Parallel Architecture with a larger focus on the interplay between linguistic and general cognitive domains during preparation and production of sentences across languages with different grammatical systems.

2. Interfacing language and cognition.

It is important to note that the remaining part of this paper will discuss real-time language processing, rather that the general organization of language faculty. Psycholinguistics of speech production and comprehension attempts to inform us about psychological principles governing real-time linguistic communication. Successful linguistic communication depends on the speakers' ability to make rapid syntactic choices and maintain cohesive organization of the discourse. If the discourse is about visually perceived events, it is likely that the cognitive operations otherwise limited to our linguistic behavior (e.g., syntactic processing) might somehow relate to basic cognitive processes involved in scene perception (e.g., memorial activation and attentional control). So far, the empirical evidence for such interplay is limited and controversial. Some studies suggest that speakers' syntactic choices directly depend on the perceptual properties of the described world (Olson and Filby 1972; Tomlin 1995; Forrest 1997; Myachykov and Posner 2005; Myachykov, Tomlin, and Posner 2005; Gleitman et al. in press). Others maintain that the way we perceive the world has little to do with the subsequent syntactic choices. For example, Myachykov (2007) and Myachykov and Garrod (in preparation) provide evidence for a stronger dependence of such choices on a speaker's linguistic knowledge, while Griffin and Bock (2000) argue for a structure-dependent tendency of speakers to visually interrogate the scene as they describe it.

It is important to draw a cautious demarcation line between the visually-cued production studies that elicit alternating syntactic structures (e.g., Tomlin 1995; Myachykov, Tomlin, and Posner 2005) from those interested in the speakers' choice of the sentential starting point (e.g., Forrest 1997; Gleitman, et al. in press). The reason for such separation is primarily theoretical: While both types of studies make claims about the syntactic choices, in fact only the tasks used in the former group require changing the sentential frame (e.g., Voice alternation).

The tasks used in the latter group of studies elicit structures with conjoined noun phrases, verbs of perspective (give/receive), or locatives (above/below). The problem with interpreting the results in these studies derives from the fact that while a speaker's perspective may be influenced by the perceptual cue, the syntax of the sentence does not require alternation. This, in effect, limits the inferences about possible reliance of the syntactic processing on the attentional focus in visually mediated speech.

Whether speakers' linguistic choices depend on attentional control or not, there most researchers seem to agree about the existence of some form of a regular interface between the systematic processes associated originated in the linguistic code and the perceptual and dynamic information encoded in the conceptual structure. While some argue that language is for the most part an autonomous and encapsulated module with a very limited influence from other cognitive systems, others try to derive virtually all linguistic phenomena from general cognition. These two proposals are not mutually exclusive.

Many models of sentence processing allow parallel simultaneous contributions from perception, local semantics, and general linguistic knowledge (e.g., McClelland and Rummelhart 1981; Dell 1986; Kempen and Vosse 1989; Vigliocco and Hartsuiker 2002, for production; Tanenhaus and Trueswell 1995; Tabor and Tanenhaus 1999; Knoeferle and Crocker 2006, for comprehension). Often, authors invoking parallel architecture describe organization of the processor with the help of metaphors, such as "workspace" (Ferreira 2005) or "blackboard" (Jackendoff 2002; van der Velde and de Kamps 2006). A good example of such approach is found in Dehaene, Kerszberg, and Changeux (1998), Dehaene and Naccache (2001), and Dehaene and Changeux (2004) (see Figure 1).

Two important features make Dehaene's global workspace model and Jackendoff's parallel architecture very similar: Both models try to avoid strict directionality of input-output relations between the processing stages; both permit 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 2004: 1147). This type of architecture allows modelling of real-time language processing as a set of cascading parallel interactions between different types of

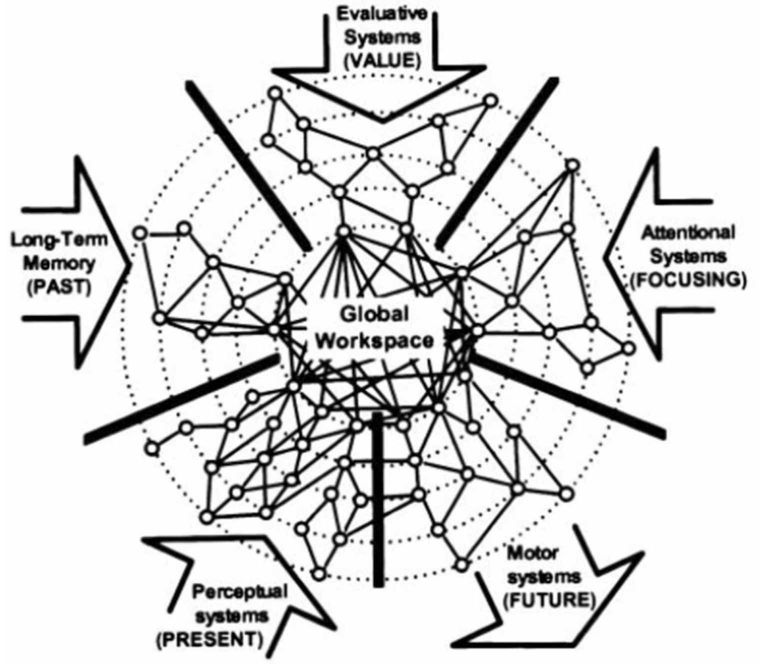
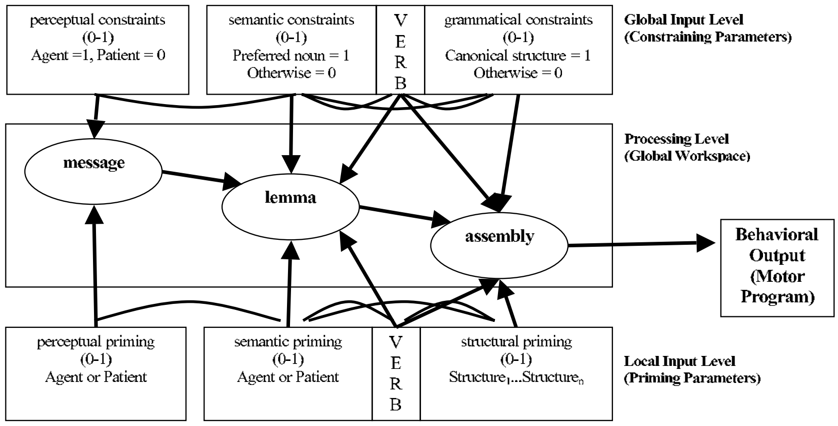


Figure 1. Global Workspace Model (Dehaene, Kerszberg, and Changeux 1998)

information, each of which tries to bias the course of preparation, production, and comprehension of speech.

The experimental data we will briefly discuss here happen to fit Jackendoff's theoretical platform very well. However, due to specific requirements of online processing, parallel architecture is accommodated to reflect extemporaneously produced speech. While Jackendoff explores parallelisms both within language layers and the processes between language and other cognitive domains, our primary goal was to analyze 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 relevant processes within other cognitive domains, especially the distribution of attention in the described scene. Figure 2 illustrates parameters of the resulting model.

The proposed model is organized around a number of principles. Some of these principles help generate specific hypotheses about the model's behavior in different language systems. The processing level (workspace) includes traditional Levelt's (1989) nodes – message, lemma, and assembly. Operations



Interfacing principles:
- Competition: Priming will sometimes compete with the linguistic constraints
- Connection: The further the connection reaches, the greater the chances its influences will be carried on
- 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 2. Production model (example of voice alternation)

at the processing level occur in parallel cascade fashion. Therefore, operations within neighboring nodes can temporarily overlap. *Hypothesis 1:* there may be effects of early syntactic processes in speech production which are not predicted by strictly encapsulated feed-forward models.

Processing nodes receive inputs from two input levels – global and local (cf. Prat-Sala and Branigan 2000). Informational inputs at both input levels are *roughly* divided into three types – perceptual, semantic, and grammatical. Inputs from the global level are organized into a set of *constraining parameters.* Inputs from the local level comprise a set of *priming parameters.* Constraining parameters exist as interlocutors' communicative 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 a 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 cues in the described array – a subset of the language grammatical defaults. An example of a grammatical constraint is a speaker's tendency to utilize canonical grammar regardless of the distribution of referential prominence.

While perceptual and semantic preferences might be very similar between speakers of different languages, the power and the nature of grammatical constraints may differ across grammatical systems. In other words, some grammars may be more constraining than others. For example, an additional regulator for the assignment of word order, case marking, might lead to higher dependence of the speakers of case marking languages to use canonical structures. *Hypothesis 2*: 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, within the communicative situation. They 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 visual 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. *Hypothesis 3:* in the absence of competing cues, speakers of SVO languages will tend to fixate agents first, and produce active voice SVO sentences 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. Processing directly biasing linguistic output occurs within "language-related" nodes (lemma and assembly). In other words, these nodes are "busier" than the message node during processing. Therefore, an increase in the semantic or syntactic complexity of the produced structure will decrease the likelihood of perceptual effects to promote its biases to the output sentence. *Hypothesis 4:* 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 in order to decrease the cognitive load, 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. *Hypothesis 5:* the endogenous restriction of attentional focus toward the patient will increase the relative weight of this priming parameter and help promote accommodation of its bias at the lemma level, leading to higher probability of inclusion of the corresponding name at the frontal position of an upcoming transitive 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 provide weaker biases of the resulting syntactic structure because its effects are realized within the message level. *Hypothesis 6*: when a perceptually primed patient competes with a lexically primed agent, there is a larger probability that the agent will be included at the frontal position of an upcoming transitive sentence. *Interaction 1*: Only the processing nodes adjacent to each other will interact. For example, there is no direct connection between the message and the assembly nodes. *Hypothesis 7*: Priming at the message node will have to be corroborated by priming at the lemma node for its effect to be promoted to the assembly level. *Interaction 2*: Only the input nodes adjacent to each other will interact. *Hypothesis 8*: in the case of multiple sources of priming, one should expect interactions between perceptual and lexical, but not between perceptual and syntactic priming components.

Recent experiments provide evidence for the outlined model (Myachykov 2007). It will be impossible to discuss each study in detail here. An interested reader will be able to get familiarized with detailed accounts of the discussed research by reading separate data papers. In all the experiments, we used visual world paradigm to elicit extemporaneous descriptions of visually presented events while the participants' choice of syntactic structure was manipulated with the help of perceptual, semantic, and syntactic cues. In each experiment, we analyzed the same categories of data: structural choices made by speakers, speech onsets at each stage of sentence formulation, and eye-voice spans for each sentence constituent. The latter measurement represents the temporal lag between the last gaze to the referent relative to the onset of the corresponding name and the name onset itself. Although eye-voice span is usually interpreted as a signature of the lexical access during incremental formulation of sentences (Griffin and Bock 2000), we propose that it rather represents the general (not exceptionally lexical) processing load associated with early planning of the argument structure.

The first study asked speakers of English (Experiment 1) and Russian (Experiment 2) to provide descriptions of visually perceived events under no cueing manipulations (a free description task). This type of task allowed for the analysis of the language-perception interface during regular sentence production without biases resulting from perceptual, semantic, or syntactic priming asymmetries. Our results revealed both similarities and differences between such interfaces in the two languages. Speakers of both languages tended to analyze the scene in two processing stages. A sequence of quick fixations (rapid apprehension) allowed participants to extract the gist of the scene. In the absence of specific cues, speakers of both languages behaved similarly during this stage. i.e., initially fixating the event area, proceeding to the agent, and finally to the patient. This suggests (1) the importance of the event for rapid extraction of the referential information (cf. Knoeferle et al. 2005 for a similar effect in comprehension) and (2) early syntactic biases toward canonical sentence scheme during regular sentence production (see Hypothesis 1). During the second processing stage, participants' scanning behavior closely resembled the sequence observed in Griffin and Bock (2000): They tended to fixate the interest areas shortly before producing corresponding names in a sequence mirroring the ordering of the constituents in the sentence. These gazes, therefore, are linguistically driven, and they are likely to represent the succession of access to the referents' lexical forms reflecting the incremental nature of the sentence formulation. The same sequential properties at both processing stages in English and Russian are not surprising as both languages strongly favor SVO word order. In the absence of conflicting cues, the preference to use the default grammatical structure in virtually 100 % of cases was observed for both languages (see Hypothesis 3). However, production of Russian sentences was delayed at each constituent as compared to English; also *eye-voice span* revealed a higher processing load for Russian sentence formulation. The latter results seem to provide some initial insights into the organization of language-specific interfaces between grammatical and perceptual processes. All things equal, speakers of both languages behaved similarly in making their structural choices, but the necessity in Russian to (1) entertain various word orders and (2) properly choose between case markers resulted in slower reaction times and larger eye-voice spans. Therefore, it matters what language you speak when you try to map what you see onto what you say (cf. Slobin 1996, 2003) (see Hypothesis 2).

The next set of experiments, or Experiments 3 and 4, accordingly, revealed that in languages that use case marking to denote sentential roles, this additional grammatical operation complicates regular mapping from perceptually prominent referents to the prominent sentential roles. The first study was discussed in Myachykov, Tomlin, and Posner (2005). It demonstrated that Russian speakers are more likely to rely upon canonical grammar than their English counterparts. Unfortunately, the cueing manipulation used in that study was achieved via the presentation of quite a brutal perceptual cue (Posner 1980): the cue was clearly visible to the participants during each trial, and the instruction explicitly commanded participants to attend to the cued referent. In the second set of experiments we used a much subtler methodology and a different language – Finnish. Similar to Russian, Finnish is a case marking language with a flexible organization of word order and a strong SVO canonicality. Participants were instructed to describe pictures of transitive (Experiment 1) or ditransitive (Experiment 2) events while their attention to one of the referents was captured by presenting a 65 msec. flash. A much subtler attentional manipulation was enough to successfully capture the speakers' attention but it did not lead to any noticeable alternation of the word order in the produced sentences. A similar cueing manipulation typically leads to a much stronger alternation of syntactic structure in studies using English (i.e., Tomlin 1997; Gleitman, et al. in press). This result suggests that the grammatical systems have different constraining effects across languages. As a result, the amount of perceptual information filtered through to bias the resulting structure will depend on the constraining power of the grammatical parameter with the language system (see Hypothesis 2 again).

Similar to the behavior of Russian participants, the reaction time and eye-voice penalties associated with the use of canonical structure in patient-cued condition were observed in transitive sentences. However, there was no such effect in ditransitive descriptions: although participants tended to produce the same structure (Subject-Verb-Theme-Object) in the majority of their descriptions, there was no time difference associated with the saccadic delay in the patient or theme-cued condition. Perceptual effects in sentence production seem to be observable (either in structural alternations or in reaction time data) only if the corresponding structure is relatively simple (transitive vs. ditransitive sentences). The more processing is necessary to organize the sentence, the less likely the processor will take heed of the perceptual asymmetries (see Hypothesis 4).

In the next set of experiments using English, the relative power of the perceptual cue was manipulated along both exogenous and endogenous dimensions. The exogenous power of the cue was varied as the time the cue was presented to the participants; the endogenous power of the cue was manipulated as the absence or presence of the instruction to restrict visual focus to the cued element. Additionally, in Experiment 5 the cue was semantically informative, and in Experiment 6, it was semantically non-informative to the participants. This was achieved via presenting a red square in the area where either the agent or the patient of the event will subsequently appear (non-informative cue in Experiment 5) or previewing one of the characters (informative cue in Experiment 6). This component added another priming dimension to our investigation – referential identity or semantic availability of the referents, therefore crossing perceptual priming with semantic priming.

The analysis of the data revealed that (1) although perceptual manipulation was quite successful in attracting attention to the cued referent, its effect was short-lived and quite weak in its ability to bias speakers' syntactic choices. In order for perceptual effects to extend their influence to the processing stage where the syntactic frame is assembled, they needed to be both highly powerful and constraining (i.e., presented for longer time combined with the instruction to restrict focus to the cued referent only, or used in combination with other factors, such as referent identity priming (see Hypothesis 5). However, even a combination of explicit and constraining cue with the referent preview resulted in up to 36 % of canonical active voice sentences. Surprisingly, although participants often could not properly identify the non-cued referent due to visual focus restriction, they produced sentences like "*Someone* is punching the cowboy". Once again, these results demonstrate that very little influence is filtered through from the perceptual asymmetries inherent in the described scene unless these asymmetries are supported by the constraining power of default preferences.

In Experiments 7 and 8, three factors were manipulated: (1) a speaker's prior exposure to a particular syntactic structure (syntactic priming), (2) the level of the conceptual correspondence between the verb used in the prime and the target sentences (lexical overlap), and (3) the attentional focus to the event's agent or patient (perceptual cueing). The participants' syntactic choices were again more dependent on linguistic factors, such as syntactic priming, rather than perceptual and/or semantic factors. Moreover, perceptual salience did not influence the word order choice directly but by affecting the activation of units at the lemma level. This was revealed by the significant interaction between the perceptual cueing and lexical overlap but not between perceptual cueing and syntactic priming (see Hypothesis 7 and 8). In general, linguistic effects (canonical or primed grammar) quite easily overrode bottom-up perceptual manipulations even when the latter are strongly biasing toward the usage of the non-preferred forms.

The experimental evidence discussed above provides important support for models of parallel processing similar to Jackendoff's (2002). Moreover, it suggests how parallel architecture should be accommodated to describe the linguistic behavior during real-time speech processing. We attempted to amend the parallel architecture to investigate the language-cognition interface experimentally in languages whose grammatical systems are dramatically different from one another. The hypotheses generated by our model were supported by our experimental data. Our experimental analysis supported the existence of a parallel architecture with multiple interface modules that provide rapid integration of perceptual, semantic, and syntactic information during production of simple sentences. This architecture is largely dependent on the grammatical and lexical constraints stored as processing preferences in a particular language. There were important behavioral differences in the way production of English and Russian sentences proceeded. The flexibility of choice within the syntactic inventory of Russian comes with a price of an additional operation – case marking. The necessity to entertain various possible word order candidates with different case marking schemes complicated production of non-canonical sentences in Russian and resulted in larger dependence on the automated canonical grammar. On the other hand, non-linguistic properties of a described scene, such as perceptual salience, turned out to be much less relevant to a speaker's choices of what to say and where to say it in the sentence.

3. Conclusions

We tried to accomplish two parallel goals in this paper. One was to reply to some of the criticisms made by Jackendoff in his reply to the Special Issue of TLR. The other was to outline our own vision of how Parallel Architecture can be modified to derive a sentence production model with similar organization and functional parameters.

In order to play a better role in our understanding of cognitive operations behind language processes, it is not enough for a linguistic theory to have elegant parsimonious formalisms. An adequate theory of language also needs to be methodologically flexible, and, therefore, available for testing via experimentation. Finally, it needs to contain a comprehensive account of relevant evidence from other disciplines. Reciprocal exchange of knowledge may endanger existence of some established theoretical accounts and lead to creation of new subfields. At the same time, it will bring us closer to an adequate theory, linguistically solid, psychologically testable, and informed in interdisciplinary research.

It is important to use any methodological tool that helps advancement of such theory. However, the relation between theory and method is a two-way street. On one hand, it is important to try to provide experimental support for the observations made with the help of grammaticality judgments. On the other, the focus on methodology and powerful statistics should not distract from a greater picture. It is impossible to provide a big theoretical answer unless you have a big theoretical question that motivates experimental testing in the first place.

Jackendoff provides a good example of how to achieve the ambitious goal of unifying linguistic theory with cognitive psychology and neuroscience. His "logical decomposition" of language into a partial ordering of necessary achievements is an important step in fostering the connection between neuroscientific data and linguistic analysis. It starts to provide for linguistics the kind of componential analysis that has allowed cognitive neuroscience to associate networks of brain areas with specific cognitive processes, for example, in arithmetic, biographical memory, faces, fear, music, object perception, reading, reward, self reference etc. (Posner and Rothbart 2007a). The rapid improvement of methods to explore brain networks and the opportunities opened up by genomic sequencing make exploration of complex cognitive processes a timely venture. Jackendoff's Parallel Architecture is a valuable step to providing an increase in cooperative research designed to link linguistics more firmly to other cognitive disciplines.

Although cognitive operations involved in language processing seem to be quite specialized, linguistic communication is intimately interlinked with basic cognitive domains of attention and memory. Our own attempt to answer the question: "How much of language is cut from the same material?" resulted in a specific adaptation of the Parallel Architecture as a sentence production model. Our main interest was not to prove the independence of the language faculty or derive linguistic processing totally from general cognition. Jackendoff correctly suggests that "*all* higher mental capacities including language make use of some sort of basic machinery" (Jackendoff 2007: 389). Language acquisition is probably more related to mechanisms of joint attention (i.e., Baldwin 1995) than linguistic performance in adults. But even in acquisition there is probably a tight *linkage,* not a homomorphism, between linguistic and perceptual processing. As linguistic knowledge expands in a mature mind, language becomes more independent of situational use, and therefore the correspondences between linguistic units and real-world regularities become more violable. At the end of the day, our purpose is not to probe for degrees of separation between language and cognition but to test the nature of universal and language-specific properties of the language-cognition interface.

The proposed production model is largely integrative in nature. It permits simultaneous contributions to the processor from multiple sources of information. This information becomes available to the speaker in parallel: although processing within each node occurs within the allocated time window, simultaneous contributions to other nodes result in spreading activation from one node to another (cf. Dell 1986; Roelofs 1992, 1997; Levelt, Roelofs, and Meyer 1999; Chang 2002). Our experimental data suggest that regularities in language production are largely restricted by the constraints imposed by a speaker's linguistic knowledge. As linguistic processing becomes more automatic in adults, the constraints of automated grammar become reflected in regular patterns that reduce complexity of form and are aimed at minimization of processing load. Such minimization is achieved through development of regular interfaces between the aspects of linguistic processing and other parts of architecture the linguistic processes depend on, or are interlinked with.

As Jackendoff observes, some parts of conceptual structure, e.g., conceptualized visual context cannot directly influence the syntactic analysis (Jackendoff 2002: 204). Instead, there is little transparency between syntax and event semantics (Culicover and Jackendoff 2005). For perceptual factors to have effect on the choice of the syntactic structure, they either need to be dramatically powerful to carry on their influence to the lemma level or be combined with similar semantic and/or syntactic priming effects. On the other hand, constraints on linguistic processing are violable under specific conditions. A constraint-contradicting cueing scheme can at times override existing constraints and result in a specific bias to anchor a produced sentence to a locally prominent referent, name, or grammatical structure.

We also provide further support for cross-linguistic differences in the language-cognition interface. As predicted by Jackendoff (2007: 362), they "have to be localized *in the mapping* between LF and the surface". Linear mapping between asymmetries within conceptual representation and the referent positioning in the resulting syntactic structure is not equally possible across languages. Introduction of additional grammatical operations may lead to a larger inventory of syntactic choices, but at the same time restrict establishment of direct correspondence between conceptual prominence and distribution of sentential roles.

Overall, our experimental data help realize parallel architecture in a model of online sentence processing. We continue to explore the organization of the interface between "general and special", the rules of interplay and degree of codependence, and the extent of language-specific properties of the interface. On the other hand, we become interested in how the interface develops in children and how it deteriorates in linguistically impaired populations (i.e., Broca's aphasia).

University of Glasgow
University of Oregon

References

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

Berman, Ruth A. and Dan I. Slobin (1994). *Relating Events in Narrative: The Cross-Linguistic Study of Narrative Structure.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Chafe, Wallace (ed.) (1980). *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, New Jersey: Ablex.

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

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

Culicover, Peter W., and Ray Jackendoff (2005). *Simpler Syntax*. Oxford: Oxford University Press.

Dehaene, Stanislas and Jean-Pierre Changeux (2004). Neural mechanisms for access to consciousness. In *The Cognitive Neurosciences III*, M. Gazzaniga(ed.), 1145-1158. Cambridge, MA: MIT Press.

Dehaene, Stanislas, Michel Kerszberg, and Jean-Pierre Changeux (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences USA* 95: 14529–14534.

Dehaene, Stanislas and Lionel Naccache (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition* 79:1—37.

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

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

Forrest, Linda B. (1997). Discourse goals and attentional processes in sentence production: the dynamic construal of events. In *Conceptual Structure, Discourse and Language*, A.E. Goldberg (ed.), 149-162. Stanford, CA: CSLI Publications.

Gauthier, Isabel, Michael J. Tarr, Adam W. Anderson, Pawel Skudlarski, and John C. Gore (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing visual objects. *Nature Neuroscience* 2: 569-573

Gleitman, Lila R., David January, Rebecca Nappa, and John C. Trueswell (in press). On the *give* and *take* between event apprehension and utterance formulation. *Journal of Memory and Language.*

Griffin, Zenzi M. and Kathryn Bock (2000). What the eyes say about speaking. *Psychological Research* 11 (4): 274-279.

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

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

— (2007). Linguistics in Cognitive Science: The State of the Art. *The Linguistic Review,* this issue: 347–401.

Kanwisher, Nancy, Josh McDermott, and Marvin M. Chun (1997). The fusiform face area: A module in extrastriate visual cortex specialized for face perception. *Journal of Neuroscience* 17: 4302-4311.

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

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

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

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

Levelt, Willem J.M., Ardi Roelofs, and Antje S. Meyer (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22: 1-75.

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

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

Myachykov, Andriy (2007). Perceptual, semantic, and syntactic forces in sentence production. Manuscript, University of Glasgow.

Myachykov, Andriy and Simon Garrod (in preparation). Integrating perceptual, semantic, and syntactic information in sentence production across languages.

Myachykov, Andriy, Simon Garrod and Seppo Vainio (in preparation). Attentional focus and word order in Finnish.
Myachykov, Andriy and Michael I. Posner (2005). Attention in Language. In *Neurobiology of Attention*, L. Itti, G. Rees, and J. Tsotsos (eds.), 324-329. Academic Press/Elsevier.
Myachykov, Andriy, Russell S. Tomlin, and Michael I. Posner (2005). Attention and empirical studies of grammar. *The Linguistic Review* 22: 347-364.
Olson, David R. and Nikola Filby (1972). On the comprehension of active and passive sentences. *Cognitive Psychology,* 3:361-381.
Posner, Michael I. (1980). Orienting of attention, *Quarterly Journal of Experimental Psychology* 32: 3-25.
Posner, Michael I. and Marchus E. Raichle (1994). *Images of Mind.* Scientific American Library: New York.
Posner, Michael I. and Mary K. Rothbart (2007a). *Educating the Human Brain*. Washington, DC: APA Books.
Posner, Michael I. and Mary K. Rothbart (2007b) Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology* 58: 1-23.
Posner, Michael I., Mary K. Rothbart and B. E. Sheese (2007). Attention genes. *Developmental Science* 10: 24-29.
Prat-Sala, Mercè and Holly P. Branigan (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory & Language* 42: 168-182.
Roelofs, Ardi (1992). A spreading activation theory of lemma retrieval in speaking. *Cognition* 42: 107-142.
— (1997). The WEAVER model of word-form encoding in speech production. *Cognition* 64: 249-284.
Slobin, Dan (1996). From “thought” and “language” to thinking for speaking. In *Rethinking Linguistic Relativity,* John J. Gumperz and Stephen C. Levinson (eds.), 70-96. Chicago: University of Chicago Press.
— (2003). Language and thought online: cognitive consequences of linguistic relativity. In *Language in Mind,* Dedre Gentner and Susan Goldin-Meadow (eds.), 157-191. Cambridge, MA: MIT Press.
Tabor, Whit and Michael K. Tanenhaus (1999). Dynamic sentence processing. *Cognitive Science* 23: 491-515.
Tanenhaus, Michael K. and John C. Trueswell (1995). Sentence comprehension. In *Handbook of Perception and Cognition: Speech, Language, and Cognition,* Joanne L. Miller and Peter D. Eimas (eds.), 217-262. San Diego, CA: Academic Press.
Tomlin, Russell S. (1995). Focal attention, voice, and word order. In *Word Order in Discourse,* P. Downing and M. Noonan (eds.), 517-552. Amsterdam: John Benjamins Publishing.
— (1997). Mapping conceptual representations into linguistic representations: the role of attention in grammar. In *Language and Conceptualization,* J. Nuyts and E. Pederson (eds.), 162-189. Cambridge: Cambridge University Press.
van der Velde, Frank and Marc de Kamps (2006). Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences* 29: 37-70.
Venter, J.Craig, Mark D. Adams, Eugene W. Myers, Peter W. Li, Richard J. Mural, et al. (2001). The sequence of the human genome. *Science* 291: 1304-1335.
Vigliocco, Gabriella and Robert J. Hartsuiker (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin* 128: 442-472.