

Parser-grammar relations: We don't understand everything twice

Colin Phillips
University of Maryland

Tom Bever was one of the first people who I met in America, when I walked into his office in 1990 as a fresh-faced graduate in medieval German. After ignoring me for a while, as he characteristically shot off a few emails (at the time I didn't know what an email was), he went straight to the tough questions. "What are you doing here?" I wasn't exactly sure, but I agreed that this was probably an important issue. Tom then intimated that perhaps I wasn't really ready to be there. I'm sure that he was right, though I mumbled some weak arguments. This first exchange set the tone for much of our interaction since then. We have continued to disagree on many different things, and this chapter is no exception in that regard. But we do tend to agree on what the important issues are. This chapter is concerned with the place of a linguistic grammar in the psychology of language. I follow Tom in believing that this is an important issue. And I probably would not be writing this were it not for his vision of cross-disciplinary integration that got me excited about becoming a linguist.

A classic view, which has been standard in many areas of language science for the past 40 years, was articulated by Bever in 1970 as follows: "the relation between grammatical rules and perceptual operations is "abstract" rather than direct" (Bever 1970, p. 287). And it is echoed in his more recent work: "Grammatical rules may find behavioral instantiation, but only as a backup system slowly brought into play in the rare cases when the behavioral strategies fail" (Townsend & Bever 2001, p. 37). The basic idea is summarized in the slogan *We understand everything twice*. A more articulated version claims that in comprehension: "a quick-and-dirty parse is initially elicited. [...] This preliminary analysis uses a variety of surface schemata [...] to organize an initial hypothesis about meaning. From that, a complete parse may be accessed that computes a derivation and fills in the remaining syntactic details" (Townsend & Bever 2001, p. 163).

One might worry that Bever's view makes the linguistic grammar somewhat resistant to scrutiny, since it does not directly participate in immediate comprehension processes, and it operates on a unclear schedule. However, in much of linguistics and psycholinguistics it is assumed that a speaker's grammar is not directly recruited in speaking and understanding. In fact, Bever's view makes the grammar more amenable to scrutiny than do the views of many or even most linguists, for he does at least assume that grammatical theories are descriptions of real mental systems, and that the representations and the processes described in linguists' grammars should be understood as genuine mental processes. This view, which Shevaun Lewis and I have described as a *literalist* interpretation of grammars (Phillips & Lewis in press), contrasts with more widespread views that ascribe less reality to the components of grammatical theories. This includes the *formalist* view, according to which the representations defined by the grammar are constructed in comprehension, but the steps of grammatical derivations do not map onto any temporal sequence of operations. It also includes the *extensionalist* view, according to which grammars simply provide a recursive characterization of the grammatical and ungrammatical sentences of a language,

with little further significance attributed to the individual components of a grammatical theory.

The received wisdom, therefore, is that grammars are not directly implicated in real-time language processes, or at least that grammars provide a set of constraints that are stated in a process-neutral fashion, such that they can be used as a resource by language processing systems (e.g., Sag & Wasow, 2011). This view did not come out of nowhere. It is based on a series of arguments, many of which are connected in some way to Tom Bever. Summarizing, they are:

- 1 a. The rise and fall of the Derivational Theory of Complexity (DTC)
- b. The difficulty of reversing the generative operations of a transformational grammar
- c. Evidence for heuristics and strategies in comprehension
 - i. Garden path sentences
 - ii. Semantic heuristics
 - iii. Sensitivity of on-line analyses to probabilistic biases
 - iv. Grammatical illusions
- d. Delayed grammatical analyses
- e. Distinctness of parsing and production mechanisms

Note that the question of how literally the grammar should be taken as a model of on-line speaking and understanding is not the question of whether the grammar is 'psychologically real'. There is a tendency in psycholinguistics to focus only on things that are closely time-locked to speaking and understanding, but it is a perfectly reasonable hypothesis that there are mental processes whose operation is not tightly synchronized with external stimuli.

In what follows I comment on the current status of these arguments on the role of grammars in language processes.

1. Are grammatical representations more 'real' than grammatical processes?

The new linguistics of the 1960s led to a new psycholinguistics that explored the question of how to understand generative grammar as a model of the psychology of language. There is, of course, more to the psychology of language than the question of how we speak and understand, but that is a natural place to start. Generative grammars such as Chomsky's *Aspects* model (Chomsky 1965) provided accounts of how sentences could be generated via a combination of rewrite rules, e.g., $S \rightarrow NP VP$, and ordered transformations. Since the generative procedure was described as a sequence of steps, could it be taken as a model of ordered mental operations? Generative grammars also made interesting claims about deep structures and about the existence of syntactic material that are phonologically null. Would these constructs be corroborated by studies on comprehension difficulty?

The results of this period of research, mostly occurring in the late 1960s and early 1970s, are summarized nicely in a series of reviews (Fillenbaum 1971; Fodor, Bever, & Garrett 1974; Levelt 1974). The most commonly voiced summary of this work is that the experimental studies provided empirical support for the syntactic representations claimed by generative grammarians, but not the transformation derivations. The surface and deep structure representations were supported by findings on such phenomena as relatedness

judgments and memory recall (Johnson 1965; Suci, Ammon, & Gamlin 1967; Levelt 1970), perceptual dislocation of clicks to phrase boundaries in auditory comprehension (Bever, Lackner, & Kirk 1969), and the perceptual complexity induced by phonologically null arguments (Blumenthal 1967; Wanner 1968). On the other hand, it was widely claimed that transformational operations are not ‘psychologically real’, due to the failure of a linking hypothesis that came to be known as the Derivational Theory of Complexity (DTC). This conclusion is somewhat of an over-simplification of what was actually tested and found. It merits further discussion, not because it is important to specifically defend the DTC, which was a rather optimistic linking hypothesis, but because the simplistic summary of this work is widely cited as decisive evidence on the de-merits of transformational grammars (e.g., Bresnan 1978; Sag & Wasow 2011).

DTC is a linking hypothesis that relates the length of transformational derivations to the perceptual complexity of a sentence. It can be traced to a widely cited remark in an early collaboration between Chomsky and Miller:

2. “The psychological plausibility of a transformational model of the language user would be strengthened, of course, if it could be shown that our performance on tasks requiring an appreciation of the structure of transformed sentences is some function of the nature, number, and complexity of the grammatical transformations involved.” (Miller & Chomsky 1963, p. 481)

Note that this remark is nuanced, and it is not really a fleshed-out linking hypothesis. Its main claim is that mental computations should take some time and/or effort, and should impact the completion of tasks that involve those computations. This remains entirely standard in psycholinguistics, and is a core assumption of much work in the cognitive neuroscience of language, especially studies using fMRI (see Marantz 2005 for further discussion). Discussions of DTC focused on the narrower claim that the main predictor of the difficulty of comprehending a sentence should be the length of its transformational derivation, according to generative theories of the 1960s.

There were, for sure, some impressive early successes due to Miller and his students. For example, Miller & McKean (1964) showed that in a task requiring mapping between related pairs of sentences, the time required to relate active ‘kernel’ sentences to negative sentences and to passive sentences respectively was an almost perfect predictor of the time required to relate kernel sentences to passive negative sentences. (Both negation and passivization were regarded as derived via transformation in syntactic models of that time.) But subsequent work applied to a wider range of phenomena that were regarded at the time as transformationally related yielded less encouraging results. For example, Fodor and Garrett (1967) compared phrases with prenominal adjectives (3a) with postnominal relative clauses (3b). They assumed that prenominal adjectives are derived from relative clauses by a rule of *whiz deletion*, drawing upon a common analysis of the time. They found no evidence that single-word adjectives were more difficult than relative clauses. They also compared center-embedded sentences (4a) with semantically equivalent sentences that removed the center embedding via a series of passive operations (4b). The second sentence involves more transformations, but results confirmed the intuition that it is easier to comprehend. Neither of these contrasts is particularly surprising, nor troubling. Few today would make the claim that adjectives are derived from relative clauses, and given the well-

known difficulty of center-embedding, it is not too surprising that the benefits of removing center-embedding should outweigh any cost of passivization.

3. a. The tall boy
 b. The boy who is tall

4. a. The first shot the tired soldier the mosquito bit fired missed.
 b. The first shot fired by the tired soldier bitten by the mosquito missed.

As Townsend and Bever point out, “either the linking hypothesis was wrong, or the grammar was wrong, or both” (Townsend & Bever 2001, p. 30). It seems that the DTC, as tested, was never a particularly sensible linking hypothesis, if it was taken to imply that transformational operations should not merely contribute to perceptual complexity, but should be the primary contributor to perceptual complexity. Tellingly, there is little evidence of a struggle to defend this idea against the supposed counterarguments in the late 1960s, suggesting that people simply did not find DTC to be very useful, and so moved on to work on other things instead. Nowadays the basic idea behind DTC is alive and well, but rather than seeking ways of capturing the aggregate processing demands for a sentence, researchers more commonly seek to identify the specific costs of individual elements or processes, such as wh-dependency formation (Kaan et al. 2000, Phillips, Kazanina & Abada 2005), complement coercion (McElree et al. 2001; Traxler et al. 2002; Baggio et al. 2010; Kuperberg et al. 2010), or ellipsis resolution (Martin & McElree 2008).

The point of these remarks is not to ‘revive’ DTC. It is a bit of a straw-man linking hypothesis, and it is probably even less useful now than it was 40 years ago. But the key idea behind it was not discredited, and in many respects never went away. And for this reason it is surprising that it continues to be invoked as a motivation for a division of labor between perceptual systems and the grammar.

2. Is it feasible to use grammars as the core of sentence recognition devices?

In contrast to Fodor and colleagues’ argument about DTC, which has been widely cited, they presented a different, and largely overlooked, argument about the difficulty of directly using a grammar as a sentence recognition device. The argument was based on serious consideration of a 1960s generative grammar (Chomsky 1965) as a derivational model of sentence generation, and asking whether that model could be run ‘in reverse’ in comprehension. For example, if the grammar is conceived as starting with the symbol S, and then applying a sequence of rewrite rules and transformations to arrive at the output string of words, a sentence recognizer based on this grammar would start with the string of words and attempt to run the transformations and rewrite rules in reverse order to arrive back at the symbol S.

Fodor et al. describe two main challenges for such an approach. First, if the goal of the parser is to recognize the outputs of syntactic rules and use those to infer the inputs, then it is hard to account for incremental comprehension. This is because a strictly bottom-up parser cannot incrementally assemble the kind of right-branching structure that is most commonly found in natural languages. For example, in a simple sentence like *John said that Mary had left the house* the first pair of words that a recognition device of this kind would

be able to combine would be *the house*, i.e., the final words of the sentence. This is because they correspond to the output of a rule, $NP \rightarrow Det N$, and no previous word pairs in the sentence match the output of a rewrite rule. This clearly does not capture the incremental nature of language comprehension.

The second challenge was presented by sentences that had undergone transformations. Contemporary grammars specified exhaustively the list of symbols that were the output of any rewrite rule, making it feasible for a parser to search for the output of rewrite rules. In contrast, the output of transformational rules was not exhaustively specified lists of symbols that could be recovered from a surface string. Transformational rules could contain variable symbols that ranged over a variety of different environments (5a), or null elements resulting from deletion operations (5b). Recognizing either of these in the input is far from straightforward. Fodor et al. acknowledged that it is possible to invoke additional mechanisms to get around these problems, but they cautioned – probably correctly – that once those mechanisms are made explicit, they would soon come to do more work than the grammar itself, thereby undermining the goal of using the grammar as a sentence recognition device.

5. a. Wh-movement

	X	wh-NPY			
	1	2	3	→	
2	1	0	3		

b. VP-ellipsis

X	VP ₁	Y	VP ₂	Z	
1	2	3	4	5	→
1	2	3	0	5	

Condition: VP₁ = VP₂

Much more could be said on this topic, and decades of subsequent work in computational parsing models has shown both how phrase structure grammars of the kind that Fodor et al. were working with can be used incrementally (Resnik 1992; Crocker 2010), and how alternative grammar formalisms can avoid the problems that Fodor et al. raised (Pollard & Sag 1994; Steedman 2000; Kempson et al. 2001; Phillips 2003). Even the ‘minimalist’ descendants of 1960s transformational grammars are associated with explicit parsers (Stabler 2011). Moreover, it is not clear that the best way of using the grammar in a recognition device would be to ‘run it in reverse’. As discussed further in Section 5 below, it is possible that sentence structures are generated in the same (roughly) left-to-right order, using the same combination of bottom-up and top-down mechanisms, in parsing and production alike. As such, this argument for distancing grammar from language processing is weaker than it once was.

3. Does parsing rely on heuristics and strategies?

If on-line analysis cannot make direct use of the grammar, then some other system must be responsible for the general success of real-time comprehension. There is a long tradition of assuming that this success is due to a series of heuristics and strategies that are good at getting the job done quickly, at the possible cost of some degree of accuracy. Bever has argued for 40 years that comprehenders parse using a series of template-based heuristics that conform to probabilistic tendencies in the ambient language, but are only indirectly related to the grammar. Evidence for the operation of these heuristics comes from various cases where comprehenders appear to construct analyses that deviate from the grammatically sanctioned analysis.

a. Garden Paths

Heuristics such as the 'NVN' template for a clause were used to account for the existence of garden path effects in parsing, as in notorious examples like *The horse raced past the barn fell* (Bever 1970, p. 316). Such templates are described in later work as part of a *pseudo-grammar* (Townsend & Bever 2001). Similar ideas have been invoked in much other work over the years, such as strategies for the use of case marker information in the processing of verb-final clauses in German and Japanese (Bornkessel-Schlesewsky & Schlewsky 2006; Miyamoto 2002).

However, pseudo-grammatical templates are not needed to account for these effects. The garden path effects highlighted by Bever and others are readily captured by the use of the regular rules of the grammar, as was emphasized early on by Kimball, Frazier, Fodor, and others (Kimball 1973; Frazier & Fodor 1978). Importantly, the use of probabilistic cues to draw inferences about the structure of input sentences does not entail the construction of pseudo-grammatical representations, as has been made clear by much work on probabilistic parsing in computational linguistics.

It should also be noted that after 40 years there is still rather limited specification of what the heuristics are. An 'NVN' strategy will only get us so far. A quick look at the range of surface forms that comprehenders appear to be able to successfully parse rapidly and effectively suggests that a suitably elaborated account of heuristics would quickly grow so large as to encompass the entire grammar.

b. Semantic comprehension strategies.

Another idea with a long history is the suggestion that speakers often construct interpretations for sentences by simply combining the words of a sentence in a way that is semantically plausible, ignoring the constraints of the syntactic form of sentence. For example, comprehenders may mistakenly interpret the implausible sentence "The dog was bitten by the man" as describing a more plausible scenario in which a man was bitten by a dog. This proposal goes back at least to Slobin (1966), and in Bever (1970) it is described as *Perceptual Strategy C*. The proposal is echoed in Bever's more recent work (Townsend & Bever 2001, p. 151), and is also advanced in a series of studies by Fernanda Ferreira (Ferreira 2003, Ferreira & Patson 2007). A mechanism that combines words to build semantic interpretations that mismatch the syntactic form of a sentence can be described

as *Independent Semantic Composition* (Stroud & Phillips 2012a; Chow & Phillips submitted). This contrasts with the view that syntactic and semantic processes are tightly coupled, such that the structure of a sentence determines how individual words are combined to form larger meanings. This tight coupling of syntax and semantics is a standard assumption in most linguistic theories (e.g., Heim & Kratzer, 1998; Kempson et al., 2001; Pollard & Sag, 1994; Steedman, 2000), and it is also explicitly or implicitly endorsed in well-known psycholinguistic and neurocognitive theories (e.g., Frazier & Clifton, 1996; Friederici, 2002; Hagoort, 2008; MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus & Trueswell, 1995).

It is certainly true that there are situations where comprehenders arrive at interpretations that mismatch the grammatically sanctioned semantics for the sentence. We are often forced to do this when trying to understand speech in noisy environments. And when we encounter speech errors we often understand the speaker's intended message rather than what the speaker actually said, particularly if we already know what the speaker was likely to say. In the extreme case, we can take strings of words that have no syntactic form at all, and successfully recover a plausible meaning. But we probably should not conclude from this that there is a general heuristic interpretive mechanism that operates in parallel with a grammatically based interpretive mechanism.

A heuristic interpretive system that genuinely operated in parallel with the grammatical interpretive system should be able to recover plausible-but-illicit interpretations from all types of sentences. But the ability to access plausible-but-illicit interpretations is more restricted, occurring mostly in sentences with more complex syntactic form. For example, Ferreira (2003) found that although comprehenders frequently misinterpreted passive sentences, with close to 20% errors, they gave many fewer misinterpretations for active sentences, with just 5-6% errors (see also MacWhinney, Bates, & Kliegl 1984). This is similar to findings with children and agrammatic aphasic patients.

Also, the evidence for the use of plausibility-based comprehension strategies is narrower than often supposed. A classic picture-verification study by Slobin is often cited as support for a plausibility-based interpretive heuristic ("making sentences non-reversible largely washed out the difference in syntactic complexity between active and passive sentences": Slobin 1966). But Slobin's findings are readily explained without recourse to a plausibility-based heuristic. In the key conditions, participants judged sentences like *The cat is being chased by the dog* (reversible) and *The girl is being watered by the flowers* (non-reversible). They first heard a sentence, then saw a picture and judged its match to the sentence. The reaction time was measured from the presentation of the picture. But since non-reversible sentences could be classified as plausible or implausible before the picture was presented, the lack of effects of reversibility could simply reflect the fact that key processes happened before timing began.

In a more recent study, Ferreira argued that "passives are frequently and systematically misinterpreted, especially when they express implausible ideas" (Ferreira 2003, p. 164), but responses in her task that required explicit thematic role judgments ("Who was the doer?", "Who was acted on?") showed relatively small differences in error rates between reversible and non-reversible implausible passives (20% vs. 17%, respectively, in Experiment 1; 29% vs. 25% in Experiment 2).

Recent findings using event-related brain potentials (ERPs) have triggered renewed interest in independent semantic composition, due to some surprising results that have been widely interpreted as providing evidence for such a mechanism.

In contrast to many previous studies that have found that semantic anomalies elicit an N400 effect at the anomalous word in sentences like (7) (Kutas & Hillyard 1980; for reviews see Kutas & Federmeier 2000; Lau, Phillips, & Poeppel 2008), a number of recent studies have shown that sentences containing what can loosely be described as thematic role anomalies instead elicit a P600 effect (Hoeks et al., 2004; Kim & Osterhout, 2005; Kolk et al., 2003; Kuperberg et al., 2003). This is surprising, since P600 effects have standardly been associated with the detection of morphological or syntactic anomalies (Osterhout & Holcomb, 1992; Hagoort, Brown, & Groothusen 1993; Neville et al. 1991). In an important study in English Kim and Osterhout (2005) showed that sentences like (7c) elicited a P600 effect and no N400 effect, relative to control sentences like (9a-b).

- | | | |
|-------|---|---------------|
| 6. | She drinks her coffee with cream and socks. | N400 |
| 7. a. | The hearty meal was devoured by the children. | Control 1 |
| b. | The hungry boy was devouring the cookies. | Control 2 |
| c. | The hearty meal was devouring the kids. | P600, no N400 |

The P600 effect in (7c) is surprising, given that the sentence is semantically anomalous yet syntactically perfectly well-formed. Standard views on the N400 and P600 would therefore predict an N400 effect. But Kim and Osterhout suggest that this ERP response profile is exactly what is expected if comprehenders use an independent semantic composition mechanism to first interpret *the hearty meal* in (7c) as the theme argument of the verb *devour*, based on the most plausible role assignments and ignoring the fact that the sentence is syntactically in active voice. This could explain the lack of N400 effect, since independent semantic composition would initially yield a plausible interpretation. If comprehenders subsequently noticed that the surface form of the sentence does not, in fact, support this interpretation, then the anomaly could be perceived as a syntactic mismatch, thereby eliciting a P600. This account predicts that the P600 should be elicited only if the anomalous subject-verb combination in sentences like (7c) is ‘semantically attractive’, i.e., the words can plausibly be combined as verb and argument, even if not in the exact manner dictated by the sentence form. Importantly, in a follow-up experiment Kim and Osterhout confirmed this prediction. A sentence like (8), which on the surface is semantically anomalous just like (7c), but which lacks the property of semantic attraction, elicited an N400 effect at the verb, and no reliable P600 effect.

- | | | |
|----|--|---------------|
| 8. | The dusty tabletop was devouring the kids. | N400, no P600 |
|----|--|---------------|

If correct, Kim and Osterhout’s findings are potentially very important, as they would constitute clear on-line evidence for the effects of plausibility-based interpretation strategies. Unlike findings from end-of-sentence judgment tasks, the ERP results could not be dismissed as the effect of later reflective strategies. In fact, many other authors have invoked some form of independent semantic composition in response to findings similar to

(7c) (Bornkessel-Schlesewsky & Schlesewsky 2008; Kuperberg 2007; Oishi & Sakamoto 2009; Kolk et al. 2003).

However, Kim and Osterhout's argument depends on the combination of three key findings, and subsequent findings have confirmed only some of them. The key findings are: (i) the presence of a P600 in cases of thematic/selectional anomaly without syntactic violations; (ii) the lack of N400 effects in sentences with 'reversed' thematic roles; (iii) the disappearance of the 'thematic P600' in the absence of semantic attraction.

The first key finding is quite robust. Thematic anomalies in syntactically well-formed sentences consistently elicit a P600 effect. This has been replicated many times over, in different labs and in different languages (for reviews see Kuperberg 2007; Brouwer, Fitz, & Hoeks 2012). In fact, the effect has become so routine that it is news when a construction or language fails to show that effect (Bornkessel-Schlesewsky et al. 2011).

The second key finding, about the disappearance of N400 effects in sentences with reversed thematic roles, has also been replicated multiple times. (Kolk et al. 2003; Hoeks et al. 2004; van Herten et al., 2005; van Herten et al. 2006; Kuperberg et al. 2003; Kuperberg et al. 2007; Kim & Osterhout 2005; Stroud & Phillips 2012ab; Chow & Phillips submitted; Ye & Zhou 2008; Vissers et al. 2006). However, the nature of these results should be clarified. It is not the case that semantically anomalous words in role-reversed sentences fail to elicit an N400 component. As is well known, all open class words elicit an N400, whether in isolation or in a sentence context, and the amplitude of this component varies as a function of how expected or anomalous the word is in context, and due to lexical factors such as frequency and priming (Kutas & Federmeier 2000). Rather, what has been found is that the N400 at the verb in role-reversed sentences is identical in amplitude to the N400 elicited by the same verb in control sentences with the same words in canonical order. Therefore, the comparison of the two conditions yields no "N400 effect", rather than a complete lack of N400. The lack of an N400 effect could, in principle, indicate that role-reversed sentences are initially interpreted as semantically plausible, due to a plausibility-base heuristic, i.e., there is a temporary 'semantic illusion'. But the lack of N400 effect could equally reflect the fact that the verb in role-reversed sentences and their canonical counterparts is preceded by the same lexical items, and so is equally primed in both conditions. Under this alternative account, the presence of an N400 effect in (8) but not in (7c), relative to control sentences like (7a), reflects the fact that the verb in (8) is less related to the preceding words than is the verb in (7c) (for further discussion see Brouwer et al. 2012; Chow & Phillips submitted).

The third, and most important, finding in Kim and Osterhout's study is the lack of a P600 effect in thematically anomalous sentences that lack semantic attraction between the verb and the arguments. The contrast between anomalous sentences with semantically attractive and unattractive words is crucial for the argument for independent semantic composition, as semantic attraction is a property that is independent of the surface form of a sentence. This contrast has not been tested in most studies of the thematic P600, but in those studies that have tested it Kim and Osterhout's finding has not, to my knowledge, been replicated. All studies that I am aware of have found that a P600 is elicited by sentences with attractive and non-attractive lexical items alike (van Herten et al. 2006; Kuperberg et al. 2007; Oishi & Sakamoto 2009; Paczynski & Kuperberg 2011; Stroud & Phillips 2012ab; Chow & Phillips submitted). In our own studies, we have tested the contrast in 6 different ways, across 4 studies in English, Spanish, and Chinese, and we have

consistently found that the P600 effects are identical in conditions with semantically related and unrelated words. This includes one study in English that used identical target items to Kim and Osterhout’s study, albeit with a more standard RSVP presentation rate – 500ms/word in our study (Stroud & Phillips 2012b).

In one study in English (Stroud & Phillips 2012b) we tested a prediction of genuinely independent semantic composition that P600 effects would be modulated by semantic attraction not only from co-arguments of the verb, as in (9), but also by nouns that were verb-adjacent but not coarguments of the verb, as in (10) and (11).

- | | | |
|-----|--|-----------------------------|
| 9. | <i>No intervener conditions</i> | |
| a. | The large lawn was mowed ... | <i>control</i> |
| b. | The large lawn was mowing ... | <i>attractive subject</i> |
| c. | The rural house was mowing ... | <i>unattractive subject</i> |
| 10. | <i>Attractive intervener conditions</i> | |
| a. | The tall grass on the large lawn was mowed ... | <i>control</i> |
| b. | The tall grass on the large lawn was mowing ... | <i>attractive subject</i> |
| c. | The front porch beside the large lawn was mowing ... | <i>unattractive subject</i> |
| 11. | <i>Unattractive intervener conditions</i> | |
| a. | The tall grass around the rural house was mowed ... | <i>control</i> |
| b. | The tall grass around the rural house was mowing ... | <i>attractive subject</i> |
| c. | The front porch of the rural house was mowing ... | <i>unattractive subject</i> |

In none of the three sets of conditions did we find a difference in the P600 elicited by the attractive and unattractive subject conditions (b-c sentences). Nor did we find effects on the P600 of the attractiveness of the intervening noun. These findings provide no support for claims of independent semantic composition.

c. Probabilistic heuristics

The use of probabilistic knowledge in parsing is sometimes cited as a reason for putting some space between parser and grammar, and as evidence for the use of pseudo-grammatical heuristics (Townsend & Bever 2001). The effects of probabilistic information in parsing ambiguous and unambiguous strings are well documented (MacDonald et al. 1994; Trueswell 1996; Gennari & MacDonald 2009), and the probabilistic effects may in some cases be stated over strings that do not correspond to syntactic constituents in most grammatical theories (e.g., subject-verb combinations in English, bigrams). However, these effects tell us about how comprehenders select among possible representations for incoming sentences, not about the nature of the representations that they build. The use of probabilistic knowledge does not undermine the notion that comprehenders build full-fledged syntactic representations for incoming sentences.

The assignment of probabilities to various types of grammatical units, such as lexical items and verb phrases, presents no reason for putting distance between the grammar and perceptual systems. But in cases where the stored probabilities are associated with units that do not correspond to grammatically-defined objects, e.g., non-constituents, then one

could object that using the probabilistic information to build fully-fledged syntactic representations is not enough. Under this view, 'direct' use of the grammar should involve matching input sequences against the true units of grammatical representations, rather than using extra-grammatical cues to mediate between inputs and grammatical representations. The existence of such cues would mean that comprehenders store parsing-specific cues that are task-specific, and hence go beyond what the grammar provides. This is an interesting objection, but it puts minimal distance between parser and grammar, and since the cues are used to build full-fledged syntactic representations it certainly does not entail that we understand everything twice. Additionally, it could be that comprehenders do not show sensitivity to the probabilities of sequences that are not grammatically defined. Or it could be that the standard constituency claims of phrase structure grammars need to be revised in favor of accounts with more flexible or dynamic constituency (e.g., Steedman 2000, Phillips 2003).

d. Grammatical illusions

A fourth motivation for claiming that the grammar is not directly recruited in parsing and production comes from errors and 'grammatical illusions'. These are cases where speakers process sentences that they would judge to be unacceptable in off-line tasks as if they are acceptable in on-line measures. Townsend and Bever (2001, p. 184) argue that such effects are just what we might expect if on-line comprehension is guided by a series of rough-and-ready templates.

Grammatical illusions in comprehension include such phenomena as agreement attraction (Pearlmutter, Garnsey, & Bock 1999; Staub 2010; Wagers, Lau, & Phillips 2009). Comprehenders are normally highly sensitive to violations of subject-verb agreement, but accuracy rates drop sharply in sentences where an incorrect plural verb matches the number of a structurally irrelevant noun in the sentence. This is not just a case of proximity concord, as the attraction effect is similarly strong with nouns that are closer or more distant than the true subject noun (12-13).

12. a. The key to the cabinets probably are on the table.
b. The world of big NIH grants are gone.
13. a. The musicians that the driver see each morning always play their best.
b. In what ways do the hypotheses one entertain influence visual search?

Are agreement illusions the product of template matching or a pseudo-grammar? I suspect not. These illusions can be accounted for using the exact constraints of the grammar, embedded in a noisy memory retrieval architecture. Presentation of a verb triggers retrieval of an element in memory that has the properties that it is the subject of the same clause, and has person/number features that match the verb. That is exactly what the grammar requires. Partial matches to those retrieval cues can give rise to agreement illusions (Wagers, Lau, & Phillips 2009). In addition, the predictions of this mis-retrieval account can better capture the selectivity of agreement illusions, especially the *grammatical asymmetry*: ungrammatical agreement is often misperceived as grammatical, but grammatical agreement is only very rarely misperceived as ungrammatical. This

contrast is unexpected if agreement illusions are the product of a pseudo-grammar or probabilistic template matching. Also, although it is tempting to capture instances of attraction from local nouns (12) in terms of template matching, that account does not easily extend to cases like (13) in which the attractor noun is further away from the verb than is its true subject.

A second case of illusions involves judgments of negative polarity item (NPI) licensing. Negative polarity items are expressions like *any*, *ever*, *yet*, or *the slightest bit* that are generally only acceptable when they appear in the scope (c-command domain) of negation or a negative-like operator, including negative quantifiers (e.g., *few*, *rarely*), negative verbs (e.g., *doubt*, *deny*) or expressions of surprise. The examples in (14) illustrate the combination of semantic and structural properties needed for NPI licensing. (14a) has a negatively quantified subject that c-commands the NPI *ever*, and it is judged to be acceptable. (14b) lacks a negative expression, and the NPI is not licensed. (14c) is the most interesting case, as the negation is embedded inside a relative clause, from where it fails to c-command the NPI. Such sentences are reliably judged to be unacceptable in tasks where participants have ample time to make their judgment. But in speeded tasks, and in online measures using reading times or ERPs, sentences like (14c) are frequently treated as if they are acceptable, typically on 15-30% of trials. It should be emphasized that (14c) is not a case of variable grammar or of a negative quantifier that takes scope outside of its clause. Speakers agree on the unacceptability of (14c) when they are given sufficient time, and it is the mismatch between on-line and off-line measures that makes (14c) count as a grammatical illusion.

14. a. No bills [that the democratic senators supported] will ever become law.
- b. *The bills [that the democratic senators supported] will ever become law.
- c. The bills [that no democratic senators supported] will ever become law.

Do illusions like (14c) provide evidence for a pseudo-grammar? Probably not, although in this case the nature of the illusion is less certain. It is tempting to analyze NPI licensing as a kind of item-to-item dependency, similar to agreement or anaphora, but modern theories of negative polarity claim that NPIs are licensed by the semantic/pragmatic properties of entire propositions, rather than by individual words or phrases (Kadmon & Landman 1996; Chierchia 2006). In work led by Ming Xiang and Brian Dillon we argued that illusory NPI licensing reflects an over-extension of exactly the mechanisms provided by the grammar (Xiang, Dillon, & Phillips 2009), and in more recent work Xiang corroborated this by showing that speakers with greater pragmatic inferencing abilities (as defined by low scores on the verbal reasoning sub-test of an autism quotient scale) are more susceptible to illusory NPI licensing (Xiang, Grove, & Giannakidou 2011). However, Vasisht and colleagues have proposed that illusory NPI licensing is the product of a partial memory match in an item-to-item dependency, similar to our analysis of agreement illusions (Vasisht, Brüssow, Drenhaus, & Lewis 2008). If we are right that illusory NPI licensing and regular NPI licensing use the same mechanisms, then the illusions do not challenge the tight connection between grammatical constraints and on-line processes. If, on the other hand, it turns out that the semantic/pragmatic analyses of NPI licensing are correct but on-line NPI licensing treats it as an item-to-item dependency, then that would constitute a good case of parser-grammar mismatch.

A third case, and one that is highlighted by Townsend and Bever, involves notorious cases of illusory acceptability involving comparative constructions. Sentences like (15) generally sound acceptable on first encounter, but on further reflection they turn out to be semantically incoherent.

15. More people have been to Russia than I have.

Townsend and Bever regard comparative illusions like (15) as the result of blending of two grammatically possible sequences, i.e., *More people have been to Russia than I*, and *People have been to Russia more than I have*. They correctly point out that such cases should be rife if speakers rely on template matching to derive initial interpretations. But our finding, based on studies led by Alexis Wellwood, is that comparative illusions are more constrained than such an account would predict (Wellwood et al. 2009). Rather than being the product of blends, they are the result of over-application of a property of English and other languages that allows statements about numbers of individuals to be understood statements about numbers of events, as in (16) (Krifka 1990; Barker 1999). (16a) has the form of a statement about numbers of individuals, but it does not entail that four thousand distinct ships passed through the lock, only that there were that many events of passing through the lock, many of which may have been carried out by the same ship. Similarly, (16c) is compatible with a situation in which the same number of individuals simply visited Russia more frequently.

16. a. Four thousand ships passed through the lock last year.
b. 107,912,000 cars crossed the George Washington Bridge in 2007.
c. More people went to Russia in the past 10 years than in the previous decade.

The evidence that comparative illusions like (15) are dependent on an ‘event counting’ interpretation comes from the finding that the illusions are stronger in sentences with a predicate that is ‘repeatable’, i.e., it can be carried out multiple times by the same agent, than in sentences with a ‘non-repeatable’ predicate that can only be carried out once by the same agent (Wellwood et al. 2009). Sentences like (17a) involve a repeatable predicate and are rated higher in off-line judgment tasks than sentences like (17b) that involve a non-repeatable predicate. This does not mean that sentences like (15) are grammatical or consistently acceptable – speakers often reject them, and the likelihood of rejection increases with more time and exposure. Instead it suggests that the illusions are the result of a ‘semantic garden path’, in which speakers initially exploit a mechanism that is perfectly acceptable in English, and then somehow fail to notice when the sentence form continues in a way that makes that mechanism illicit.

17. a. More undergrads call their families during the week than I do.
b. More New Yorkers began law school this semester than I did.

Summarizing the discussion in this section, phenomena such as garden path sentences, thematic role reversals, and grammatical illusions all present possible empirical motivation for assuming a gap between perceptual systems for language and mental grammars. But in each case closer examination of the phenomena suggests that either the phenomena are

less pervasive than often assumed, or that they are the product of embedding well-motivated grammatical constraints in a noisy cognitive architecture.

4. Do slow grammatical analyses prove that we understand everything twice?

A fourth motivation for the view that *we understand everything twice* comes from the observation that we comprehend language rapidly, but many of the acceptability judgments that grammatical theories aim to explain are more difficult and are arrived at only after much reflection. Surely this contrast motivates the existence of different syntactic systems in the mind: one for rapid comprehension and production, the other acting as the true grammar.

Although attractive at first, this argument is less strong than it seems. We know from electrophysiological studies that very many judgments are accurately made within a few hundred milliseconds of the critical word appearing in the input (for review see Kaan 2007). In fact, findings of rapid syntactic anomaly detection are so routine in the ERP literature that it is news if an anomaly fails to be rapidly registered in ERP signals (e.g., Xiang et al. 2009). But even in the cases where normal speakers genuinely take some time to make a judgment, there is little reason to take this as evidence for independent linguistic systems that operate on different time scales. Judgments can be slow because of ambiguity and mis-parsing, which forces comprehenders to re-parse the sentence before accurately judging it. Judgments can also be slow because they depend on construction of a suitable semantic context in which to evaluate the sentence. Trained linguists are quite good at conjuring up contexts for judging decontextualized sentences, but non-linguists are generally less good at this. In either case, the source of slowness in judgments is not a slow grammatical system that comes in to take care of the errors that the quick-and-nasty pseudo-grammar made. Rather, the difference between faster and slower judgments can be accommodated by using a single grammatical system in differing contexts.

5. Distinctness of parsing & production mechanisms

If it can be shown that parsing and production engage distinct syntactic mechanisms, but that both make use of the same linguistic knowledge, then that presents a *prima facie* argument for the existence of a third system – the grammar – that they both draw upon. So the distinction between parsing and production systems constitutes an important piece of evidence for the separation between parser and grammar.

But the fact that parsing and production are different tasks does not entail that they engage distinct syntactic mechanisms. It is possible that the two tasks make use of the same structure-building mechanism, which generates structures that mediate form and meaning, in comprehension and production alike. In comprehension this system must generate a structure that matches the sound form in the perceptual input. In production the same system must generate a structure that matches the message that the speaker wishes to convey. Under this view, parsing and production are simply different ways of using the same system, subject to different external constraints.

Is this view plausible? Perhaps. As more is learned about parsing and production, more parallels emerge.

Classic views of sentence comprehension and sentence production presented the two processes as operating with a quite different temporal grain size. It has consistently been assumed that structure-building in sentence comprehension is a highly incremental process, operating on a word-by-word basis, or at least at a phrase-by-phrase level (e.g., Marslen-Wilson, 1973; Tanenhaus et al., 1995). In contrast, it has been widely assumed that structure-encoding in sentence production is rather less incremental, operating at the grain of clause-sized templates or even larger (Bock, 1982; Garrett 1980). This view was motivated by the fact that speakers can apprehend the message that they intend to convey in its entirety, prior to grammatical encoding, and also by the fact that clause-sized production templates provide a way to understand the existence of word exchanges in speech production and to explain constraints on the exchanges, e.g., the Category Constraint (Garrett, 1975; Fay & Cutler, 1977). In contrast to this less incremental view of sentence production, more recent work has called this into question. Studies by V. Ferreira and Dell, in particular, have shown that word order choices are affected by the timing of lexical retrieval (V. Ferreira, 1996; Ferreira & Dell, 2000), leading to a view in which grammatical encoding is highly incremental (for review see F. Ferreira & Engelhardt, 2006). Put simply, the proposal is that successfully retrieved lexical items are used to incrementally construct a sentence structure for production. If the grammar of the language allows this, lexical items that are retrieved first are also encoded first. Where the grammar prevents a retrieved word from being immediately entered into the structure, a bottleneck occurs, and the possibility of errors and exchanges arises. Findings on incrementality in sentence production narrow the gap between comprehension and production, but currently far less is known about the degree of incrementality and the ordering of operations in production than is known in comprehension.

There has been much recent interest in predictive mechanisms in language comprehension (e.g., DeLong, Urbach, & Kutas 2005; Lau et al. 2006; Staub & Clifton 2006), and this has sometimes been taken as a reflex of the role of sentence production mechanisms in comprehension (Federmeier 2007, Pickering & Garrod 2007).

There are also many behavioral differences between production and comprehension that present interesting challenges for the view that the two tasks are served by independent systems. For example, there are well-known dissociations between comprehension and production difficulties in language disorders (Geschwind 1965ab). Also, patterns of errors may differ between the two tasks. Errors of agreement attraction (18), in which illicit subject verb agreement is produced (Bock & Miller 1991, Eberhard, Cutting, & Bock 2005) or not detected (Pearlmutter et al. 1999, Wagers et al. 2009) show close parallels. But the corresponding phenomenon in reflexive licensing appears to be more different in the two domains, despite the fact that reflexives are licensed by a clause-mate subject, just like agreement (19). Comprehension of reflexives does not show the counterpart of agreement attraction (Nicol & Swinney 1989, Sturt 2003, Dillon et al. submitted; cf. Badecker & Straub 2002), but at least one study on the production of reflexives suggests that attraction errors are just as common with reflexives as agreement (Bock, Nicol, & Cutting 1999).

18. The key to the cabinet(s) are on the table.
19. The actor in the soap opera(s) watched themselves.

Similarly, syntactic priming effects have attracted much interest in recent work in sentence production (for review see Pickering & Ferreira 2008), and some amount of interest in sentence comprehension (for review see Tooley & Traxler 2010). There is evidence for priming from comprehension to production, suggesting some degree of coupling between the two processes (Potter & Lombardi 1998; Bock et al. 2007; Branigan et al. 2007). But there are also contrasts in the abstractness of the effects in the two domains. Verb-independent priming effects are robustly found in production, but priming effects are more commonly verb-specific in comprehension (Branigan, Pickering, & McLean 2005; Arai, Van Gompel, & Scheepers 2007; Traxler & Tooley 2008).

None of these findings settles the issue of whether it is possible to conflate the structure building component of sentence production and comprehension, and the mismatches are ripe for further investigation. But it is no longer obvious that the two systems are independent, and hence they no longer motivate an independent grammatical system as they once did.

6. Analysis-by-synthesis

A final point involves Townsend and Bever's claim that language comprehension involves a kind of 'analysis-by-synthesis' device. Although I disagree with many of Townsend and Bever's arguments for their architecture, I think that an analysis-by-synthesis account of language comprehension has much to recommend it. But I do not agree with the view that such accounts require a parser-grammar split, nor do they entail that we understand everything twice.

What Townsend and Bever (2001) mean by analysis-by-synthesis is that comprehenders converge on a correct analysis of an incoming sentence by first constructing a rough-and-ready heuristic parse of the sentence ('analysis'), and then using the information derived from that analysis to generate a subsequent analysis using the full-blown grammar ('synthesis'). The goal is for the synthesis stage to yield a form that matches the input. If that fails, then revision processes are triggered.

My own view of an analysis-by-synthesis system is that the syntactic components of the comprehension and production mechanisms are essentially identical. Comprehenders use the input to incrementally generate a representation that matches the form of the input utterance. Due to the active nature of the process, this generation stage often gets a step or two ahead of the external input, allowing it to guide the perceptual analysis of the input.

Therefore, the main difference between these two positions appears to lie in the timing and the grain size of the generation stage. For Townsend and Bever it appears to apply after sizeable amounts of input have been heuristically analyzed, at which point it becomes more feasible to generate an analysis of the sentence using a standard 'backwards' transformational grammar derivation. From my perspective, on the other hand, individual steps of structure generation are either synchronized with the input, or slightly ahead of the input.

Townsend and Bever highlight an interesting consequence of their model. They point to the long-standing finding that grammatical speech is perceived as acoustically clearer than ungrammatical or incoherent speech (Miller & Isard 1963), and argue that this is predicted by their model. Their version of analysis-by-synthesis predicts that grammatical sentences

are mentally represented twice, via the pseudo-grammatical parse and the output of the full-fledged grammar. They argue that this double representation is what is responsible for the heightened clarity of coherent speech. The alternative view on analysis-by-synthesis as a single grammatical system with a strong predictive component can probably capture the same perceptual illusion, but without appealing to multiple analyzers. If the redundancy that is inherent in coherent speech allows comprehenders to anticipate some aspects of upcoming input before it arrives, then the recognition of that input should be greatly facilitated once it actually arrives. Under this view, the perception of heightened clarity results from knowing something about the phonological input before it appears in the input. It need not be attributed to the notion that there are multiple syntactic analyzers for every sentence.

7. Conclusion

The conclusion of this discussion should be straightforward. Tom Bever deserves much credit for keeping questions about the architecture of the language faculty on the agenda, bridging the concerns of linguistics and psychology. I am as indebted to him as anybody for this. He has identified many of the empirical and theoretical questions that are most important for understanding the relation between language processing and knowledge of language. However, I disagree on the conclusions that we should draw from findings in this area. Townsend and Bever (2001, p 37) state that: "Cognitive science made progress by separating the question of what people understand and say from how they understand and say it." This is certainly true, for many practical reasons. They then go on to say that: "It is a philosophical truism that there is a difference between knowing *that X* from knowing *how to X*." This is also true, in principle. But it is an empirical question in each domain of cognition whether this distinction entails a division between separate cognitive sub-systems, one responsible for *knowing that* and others responsible for *knowing how*. In the case of human grammar, I think that the evidence suggests that no such division is needed.

Acknowledgments

For useful discussion of the material in this chapter I am grateful to Wing Yee Chow, Shevaun Lewis, Jeff Lidz, Shota Momma, Clare Stroud, and of course Tom Bever. This work was supported in part by NSF grant #BCS-0848554 to CP.

References

- Arai, Manabu, Van Gompel, Roger P. G., & Scheepers, Christoph (2007). 'Priming ditransitive structures in comprehension', *Cognitive Psychology* 54: 218-250.
- Badecker, William, & Straub Kathleen (2002). 'The processing role of structural constraints on the interpretation of pronouns and anaphors', *Journal of Experimental Psychology: Learning, Memory and Cognition* 28: 748-769.
- Baggio, Giosuè, Choma, Travis, van Lambalgen, Michiel, & Hagoort, Peter (2010). 'Coercion and compositionality', *Journal of Cognitive Neuroscience* 22: 2131-2140.
- Barker, Chris (1999). 'Individuation and quantification', *Linguistic Inquiry* 30: 683-691.

- Bever, Thomas G. (1970). 'The cognitive basis for linguistic structures', in J. R. Hayes (ed.), *Cognition and the development of language*. New York: Wiley, 279-362.
- Bever, T. G., Lackner, J. R., & Kirk, R. (1969). 'The underlying structures of sentences are the primary units of immediate speech processing', *Perception and Psychophysics* 6: 225-234.
- Blumenthal, A. L. (1967). 'Prompted recall of sentences', *Journal of Verbal Learning and Verbal Behavior* 6: 203-206.
- Bock, Kathryn (1982). 'Towards a cognitive psychology of syntax: Information processing contributions to sentence formulation', *Psychological Review* 89: 1-47.
- Bock, Kathryn, Dell, Gary S., Chang, Franklin, & Onishi, Kristine H. (2007). 'Persistent structural priming from language comprehension to language production', *Cognition* 104: 437-458.
- Bock, Kathryn, & Miller, Carol A. (1991). 'Broken agreement', *Cognitive Psychology* 23: 45-93.
- Bock, Kathryn, Nicol, Janet, and Cutting, Cooper (1999). 'The ties that bind: Creating number agreement in speech', *Journal of Memory and Language* 40: 330-346.
- Bornkessel-Schlesewsky, Ina & Schlewsky, Mathias (2006). 'The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages', *Psychological Review* 113: 787-821.
- Bornkessel-Schlesewsky, Ina, & Schlewsky, Mathias (2008). 'An alternative perspective on "semantic P600" effects in language comprehension', *Brain Research Reviews* 59: 55-73.
- Bornkessel-Schlesewsky, Ina, Kretschmar, Franziska, Tune, Sarah, Wang, Luming, Genç, Safiye, Philipp, Markus, Roehm, Dietmar, & Schlewsky, Mathias (2011). 'Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension', *Brain and Language* 117: 133-152.
- Branigan, Holly P., Pickering, Martin J., & McLean, Janet F. (2005). 'Priming prepositional phrase attachment during language comprehension', *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31: 468-481.
- Branigan, Holly P., Pickering, Martin J., McLean, Janet F., & Cleland, Alexandra A. (2007). 'Syntactic alignment and participant role in dialogue', *Cognition* 104: 163-197.
- Bresnan, Joan (1978). 'A realistic transformational grammar', in G. Miller, M. Halle, & J. Bresnan (eds.), *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press, 282-390.
- Brouwer, Harm, Fitz, Hartmut, & Hoeks, John (2012). 'Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension', *Brain Research*. doi:10.1016/j.brainres.2012.01.055
- Chierchia, Gennaro (2006). 'Broaden your views: Implicatures of domain widening and the "logicality" of language', *Linguistic Inquiry* 37: 535-590.
- Chomsky, Noam (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chow, Wing Yee, & Phillips, Colin (submitted). 'No semantic illusion in the "semantic P600" phenomenon: ERP evidence from Mandarin Chinese.'
- Crocker, Matthew (2010). 'Computational psycholinguistics', in A. Clark, C. Fox, & S. Lappin (eds.), *The handbook of computational linguistics and natural language processing*. Oxford, UK: Wiley-Blackwell.

- DeLong, Katherine A., Urbach, Thomas P., & Kutas, Marta (2005). 'Probabilistic word pre-activation during language comprehension inferred from electrical brain activity', *Nature Neuroscience* 8: 1117-1121.
- Dillon, Brian, Mishler, Alan, Sloggett, Shayne, & Phillips, Colin (submitted). 'Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence.'
- Eberhard, Kathleen M., Cutting, Cooper, & Bock, Kathryn (2005). 'Making syntax of sense: Number agreement in sentence production', *Psychological Review* 112: 531-559.
- Fay, David, & Cutler, Anne (1977). 'Malapropisms and the structure of the mental lexicon', *Linguistic Inquiry* 8: 505-520.
- Federmeier, Kara (2007). 'Thinking ahead: The role and roots of prediction in language comprehension', *Psychophysiology* 44: 491-505.
- Ferreira, Fernanda (2003). 'The misinterpretation of noncanonical sentences', *Cognitive Psychology* 47: 164-203.
- Ferreira, Fernanda, & Engelhardt, Paul (2006). 'Syntax and production', in M. A. Gernsbacher & M. J. Traxler (eds.), *Handbook of Psycholinguistics*. Oxford: Elsevier, 61-91.
- Ferreira, Fernanda, & Patson, Nikole D. (2007). 'The 'good enough' approach to language comprehension', *Language and Linguistics Compass* 1: 71-83.
- Ferreira, Victor S. (1996). 'Is it better to give than to donate?', *Journal of Memory and Language* 35: 724-755.
- Ferreira, Victor S., & Dell, Gary S. (2000). 'Effect of ambiguity and lexical availability on syntactic lexical production', *Cognitive Psychology* 40: 296-340.
- Fillenbaum, Samuel (1971). 'Psycholinguistics', *Annual Review of Psychology* 22: 251-308.
- Fodor, Jerry A., Bever, Thomas G., & Garrett, Merrill (1974). *The psychology of language*. New York: McGraw Hill.
- Fodor, Jerry A., & Garrett, Merrill (1967). 'Some syntactic determinants of sentential complexity', *Perception and Psychophysics* 2: 289-296.
- Frazier, Lyn, & Clifton, Charles (1996). *Construal*. Cambridge, MA: MIT Press.
- Frazier, Lyn, & Fodor, Janet D. (1978). 'The sausage machine: A new two-stage parsing model', *Cognition* 6: 291-325.
- Friederici, Angela D. (2002). 'Towards a neural basis of auditory sentence processing', *Trends in Cognitive Sciences* 6: 78-84.
- Garrett, Merrill (1975). 'The analysis of sentence production, in G. H. Bower (ed.), *The Psychology of Learning and Motivation*. San Diego: Academic Press, 133-175.
- Garrett, Merrill (1980). 'Levels of processing in sentence production', in B. Butterworth (ed.), *Language Production*. London: Academic Press, 177-220.
- Gennari, Silvia P., & MacDonald, Maryellen C. (2009). 'Linking production and comprehension processes: The case of relative clauses', *Cognition* 111: 1-23.
- Geschwind, Norman (1965a). 'Disconnexion syndromes in animals and man', *Brain* 88: 237-294.
- Geschwind, Norman (1965b). 'Disconnexion syndromes in animals and man', *Brain* 88: 585-644.
- Hagoort, Peter (2008). 'The fractionation of spoken language understanding by measuring electrical and magnetic brain signals', *Philosophical Transactions of the Royal Society B* 363: 1055-1069.

- Hagoort, Peter, Brown, Colin M., & Groothusen, Jolanda (1993). 'The Syntactic Positive Shift (SPS) as an ERP measure of syntactic processing', *Language and Cognitive Processes* 8: 439-484.
- Heim, Irene, & Kratzer, Angelika (1998). *Semantics in generative grammar*. Oxford: Blackwell.
- Hoeks, John C. J., Stowe, Laurie A., Doedens, Gina (2004). 'Seeing words in context: the interaction of lexical and sentence level information during reading', *Cognitive Brain Research* 19: 59-73.
- Johnson, N. F. (1965). 'The psychological reality of phrase structure rules', *Journal of Verbal Learning and Verbal Behavior* 4: 469-475.
- Kaan, Edith (2007). 'Event-related potentials and language processing: An overview', *Language and Linguistics Compass* 1: 571-591.
- Kaan, Edith, Harris, Anthony, Gibson, Edward, & Holcomb, Phillip (2000). 'The P600 as an index of syntactic integration difficulty', *Language and Cognitive Processes* 15: 159-201.
- Kadmon, Nirit, & Landman, Fred (1996). 'Any', *Linguistics and Philosophy* 16: 353-422.
- Kempson, Ruth, Meyer-Viol, Wilfried, & Gabbay, Dov (2001). *Dynamic syntax: The flow of language understanding*. Malden, MA: Blackwell.
- Kim, Albert, & Osterhout, Lee (2005). 'The independence of combinatory semantic processing: Evidence from event-related potentials', *Journal of Memory and Language* 52: 205-225.
- Kimball, John (1973). 'Seven principles of surface structure parsing in natural language', *Cognition* 2: 15-47.
- Kolk, Herman H. J., Chwilla, Dorothee J., van Herten, Marieke, & Oor, Patrick (2003). 'Structure and limited capacity in verbal working memory: a study with event-related potentials', *Brain and Language* 85: 1-36.
- Krifka, Manfred (1990). 'Four thousand ships passed through the lock: Object-induced measure functions on events', *Linguistics and Philosophy* 13: 487-520.
- Kuperberg, Gina R. (2007). 'Neural mechanisms of language comprehension: Challenges to syntax', *Brain Research* 1146: 23-49.
- Kuperberg, Gina R., Choi, Arim, Cohn, Neil, Paczynski, Martin, & Jackendoff, Ray (2010). 'Electrophysiological correlates of complement coercion', *Journal of Cognitive Neuroscience*, 22, 2685-2701.
- Kuperberg, Gina, Kreher, Donna A., Sitnikova, Tatiana, Caplan, David, & Holcomb, Phillip (2007). 'The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials', *Brain and Language* 100: 223-238.
- Kutas, Marta, & Federmeier, Kara D. (2000). 'Electrophysiology reveals semantic memory use in language comprehension', *Trends in Cognitive Sciences* 4: 463-470.
- Kutas, Marta, & Hillyard, Steven A. (1980). 'Reading senseless sentences: Brain potentials reflect semantic incongruity', *Science* 207: 203-205.
- Lau, Ellen F., Phillips, Colin, & Poeppel, David (2008). 'A cortical network for semantics: (de)constructing the N400', *Nature Reviews Neuroscience* 9: 920-933.
- Lau, Ellen F., Stroud, Clare, Plesch, Silke, & Phillips, Colin (2006). 'The role of structural prediction in rapid syntactic analysis', *Brain and Language* 98: 74-88.
- Levelt, Willem J. M. (1970). 'A scaling approach to the study of syntactic relations', in G. B. Flores d'Arcais & W. J. M. Levelt (eds.), *Advances in psycholinguistics*. New York: American Elsevier, 109-121.

- Levelt, Willem J. M. (1974). *Formal grammars in linguistics and psycholinguistics* (3 volumes). The Hague: Mouton.
- MacDonald, Maryellen C., Pearlmutter, Neal J., & Seidenberg, Mark S. (1994). 'The lexical nature of syntactic ambiguity resolution', *Psychological Review* 101: 676-703.
- MacWhinney, Brian, Bates, Elizabeth, & Kliegl, Reinhold (1984). 'Cue validity and sentence interpretation in English, German, and Italian', *Journal of Memory and Language* 23: 127-150.
- Marantz, Alec (2005). 'Generative linguistics within the cognitive neuroscience of language', *The Linguistic Review* 22: 429-445.
- Marslen-Wilson, William (1973). 'Linguistic structure and speech shadowing at very short latencies', *Nature* 244: 522-523.
- Martin, Andrea E. & McElree, Brian (2008). 'A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis', *Journal of Memory and Language* 58: 879-906.
- McElree, Brian, Traxler, Matthew J., Pickering, Martin J., Seely, Rachel E., & Jackendoff, Ray (2001). 'Reading time evidence for enriched composition', *Cognition* 78: 17-25.
- Miller, George A. & Chomsky, Noam (1963). 'Finitary models of language users', in R. D. Luce, R. R. Bush, & E. Galanter (eds.), *Handbook of mathematical psychology, volume 2*. New York: Wiley, 419-492.
- Miller, George A., & Isard, Stephen. (1963). 'Some perceptual consequences of linguistic rules', *Journal of Verbal Learning and Verbal Behavior* 2: 217-228.
- Miller, George A., & McKean, Kathryn (1964). 'A chronometric study of some relations between sentences', *Quarterly Journal of Experimental Psychology* 16: 297-308.
- Miyamoto, Edson T. (2002). 'Case markers as clause-boundary inducers in Japanese', *Journal of Psycholinguistic Research* 31: 307-347.
- Neville, Helen J., Nicol, Janet, Barss, Andrew, Forster, Kenneth I., & Garrett, Merrill F. (1991). 'Syntactically based sentence processing classes: Evidence from event-related brain potentials', *Journal of Cognitive Neuroscience* 3: 151-165.
- Nicol, Janet, & Swinney, David (1989). 'The role of structure in coreference assignment during sentence comprehension', *Journal of Psycholinguistic Research* 18: 5-19.
- Oishi, Hiroaki, & Sakamoto, Tsutomu, (2009). 'Immediate interaction between syntactic and semantic outputs: evidence from event-related potentials in Japanese sentence processing', poster presented at the 22nd annual CUNY Human Sentence Processing Conference, Davis, CA.
- Osterhout, Lee, & Holcomb, Philip J. (1992). 'Event-related potentials elicited by syntactic anomaly' *Journal of Memory and Language* 31: 785-806.
- Paczynski, Martin & Kuperberg, Gina R. (2011). 'Electrophysiological evidence for the use of the animacy hierarchy, but not thematic role assignment, during verb argument processing', *Language and Cognitive Processes* 26: 1402-1456.
- Pearlmutter, Neal J., Garnsey, Susan M., & Bock, Kathryn (1999). 'Agreement processes in sentence comprehension', *Journal of Memory and Language* 41: 427-456.
- Phillips, Colin (2003). 'Linear order and constituency', *Linguistic Inquiry* 34: 37-90.
- Phillips, Colin, Kazanina, Nina, & Abada, Shani (2005). 'ERP effects of the processing of syntactic long-distance dependencies', *Cognitive Brain Research* 22: 407-428.
- Phillips, Colin, & Lewis, Shevaun (in press). 'Derivational order in syntax: Evidence and architectural consequences', in C. Chesi (ed.), *Directions in derivations*. Elsevier.

- Pickering, Martin J., & Ferreira, Victor S. (2008). 'Structural priming: A critical review', *Psychological Bulletin* 134: 427-459.
- Pickering, Martin J., & Garrod, Simon (2007). 'Do people use language production to make predictions during comprehension?', *Trends in Cognitive Sciences* 11: 105-110.
- Pollard, Carl, & Sag, Ivan A. (1994). *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
- Potter, Mary C., & Lombardi, Linda (1998). 'Syntactic priming in immediate recall of sentences', *Journal of Memory and Language* 38: 265-282.
- Resnik, Philip (1992). 'Left-corner parsing and psychological plausibility', in *Proceedings of the fourteenth international conference on computational linguistics*, 191-197.
- Sag, Ivan A., & Wasow, Thomas (2011). 'Performance-compatible competence grammar', in R. Borsley & K. Borjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar*. Oxford: Wiley-Blackwell.
- Slobin, Dan I. (1966). 'Grammatical transformations and sentence comprehension in childhood and adulthood', *Journal of Verbal Learning and Verbal Behavior* 5: 219-227.
- Stabler, Edward (2011). 'Computational perspectives on minimalism', in C. Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*. Oxford University Press, 617-642.
- Staub, Adrian (2010). 'Response time distributional evidence for distinct varieties of number attraction', *Cognition* 114: 447-454.
- Staub, Adrian, & Clifton, Charles (2006). 'Syntactic prediction in language comprehension: Evidence from *either ... or*', *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32: 425-436.
- Steedman, M. (2000). *The syntactic process*. Cambridge, MA: MIT Press.
- Stroud, Clare, & Phillips, Colin (2012). 'Examining the evidence for an independent semantic analyzer: An ERP study in Spanish', *Brain and Language* 120: 108-126.
- Stroud, Clare, & Phillips, Colin (2012b). 'An investigation of the effect of thematic fit on ERP components', ms. University of Maryland.
- Sturt, Patrick (2003). 'The time-course of the application of binding constraints in reference resolution', *Journal of Memory and Language* 48: 542-562.
- Suci, G., Ammon, P., & Gamlin, P. (1967). 'The validity of the probe-latency technique for assessing structure in language', *Language and Speech* 10: 69-80.
- Tanenhaus, Michael K., Spivey-Knowlton, Michael, Eberhard, Kathleen, & Sedivy, Julie (1995). 'Integration of visual and linguistic information during spoken language comprehension', *Science* 268: 1632-1634.
- Tanenhaus, Michael K., & Trueswell, John C. (1995). 'Sentence comprehension', in P. Eimas & J. D. Miller (eds.), *Handbook in perception and cognition, volume 11: Speech, language, and communication*. San Diego: Academic Press, 217-262.
- Tooley, Kristen M. & Traxler, Matthew J. (2010). 'Syntactic priming effects in comprehension: A critical review', *Language and Linguistics Compass* 4: 925-937.
- Townsend, David J., & Bever, Thomas G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.
- Traxler, Matthew J., Pickering, Martin J., & McElree, Brian (2002). 'Coercion in sentence processing: Evidence from eye-movements and self-paced reading', *Journal of Memory and Language* 47: 530-547.
- Traxler, Matthew J. & Tooley, Kristen M. (2008). 'Syntactic priming in comprehension: Strategic or syntactic?', *Language and Cognitive Processes* 23: 609-645.

- Trueswell, John C. (1996). 'The role of lexical frequency in syntactic ambiguity resolution', *Journal of Memory and Language* 35: 566-585.
- Van Herten, Marieke, Chwilla, Dorothee J., & Kolk, Herman H. (2006). 'When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception', *Journal of Cognitive Neuroscience* 18: 1181-1197.
- Van Herten, Marieke, Kolk, Herman H. J., & Chwilla, Dorothee (2005). 'An ERP study of P600 effects elicited by semantic anomalies', *Cognitive Brain Research* 22: 241-255.
- Vasishth, Shrvan, Brüssow, Sven, Lewis, Richard L., & Drenhaus, Heiner (2008). 'Processing polarity: How the grammatical intrudes on the ungrammatical', *Cognitive Science* 32: 685-712.
- Visser, Constance T. W. M., Chwilla, Dorothee J., & Kolk, Herman H. J. (2006). 'Monitoring in language perception: The effect of misspellings of words in highly constrained sentences', *Brain Research* 1106: 150-163.
- Wagers, Matthew W., Lau, Ellen F., & Phillips, Colin (2009). 'Agreement attraction in comprehension: Representations and processes', *Journal of Memory and Language* 61: 206-237.
- Wanner, Eric (1968). *On remembering, forgetting, and understanding sentences: A study of the deep structure hypothesis*. PhD dissertation, Harvard University.
- Wellwood, Alexis, Pancheva, Roumyana, Hacquard, Valentine, & Phillips, Colin (2011). 'Deconstructing a comparative illusion', ms., University of Maryland and University of Southern California.
- Xiang, Ming, Dillon, Brian, & Phillips, Colin (2009). 'Illusory licensing effects across dependency types: ERP evidence', *Brain and Language* 108: 40-55.
- Xiang, Ming, Grove, Julian, & Giannakidou, Anastasia (2011). 'Interference licensing of NPIs: Pragmatic reasoning and individual differences', poster presented at the 24th annual CUNY Conference on Human Sentence Processing. Stanford, CA.
- Ye, Zheng & Zhou, Xiaolin (2008). 'Involvement of cognitive control in sentence comprehension: Evidence from ERPs', *Brain Research* 1203: 103-115.