Syntax and Production

Fernanda Ferreira

and

Paul E. Engelhardt

Department of Psychology and Cognitive Science Program

Michigan State University

To appear in *Handbook of Psycholinguistics*, edited by Matthew Traxler and Morton Ann Gernsbacher

Address correspondence to:

Fernanda Ferreira
Department of Psychology and Cognitive Science Program
Michigan State University
East Lansing, MI 48824-1116
fernanda@eyelab.msu.edu

23 January 2006

Syntax and Production

I. Introduction

Our capacity for language enables us to communicate virtually any thought or idea, and this is in large part because the language faculty evolved a syntactic component. Syntax allows words to be combined to create unique combinations of meaning. And although all human languages share some universal syntactic properties – for example, all allow the creation of some type of plural – the constraints on how constituents may be generated vary substantially. Thus, an English speaker knows that verbs usually come before their objects, whereas Japanese speakers learn the opposite setting of this parameter for arranging words and arguments.

But even though the grammar of any particular language constrains the way words may be put together, all languages also give the speaker some freedom of choice. Even English, which is often characterized as a language that offers little in the way of word ordering options (MacWhinney, Bates, & Reinhold, 1984), allows the speaker to choose from among at least a few different forms to express the same essential idea. A proposition involving a cat, a dog, and a state of fear may be grammatically encoded as *my cat terrifies the dog next door*, or *the dog next door is terrified of my cat*, or *it's my cat that terrifies the dog next door*, and so on. An important insight that has emerged from work on language production is that these syntactic options are used not just to convey different information structures (e.g., that the cat and not the dog is given information, and therefore the cat should take the subject position of the sentence), but also to take advantage of the states of activation within the language and cognitive systems at the moments when speakers make their syntactic decisions. Thus, any model of language production must explain how

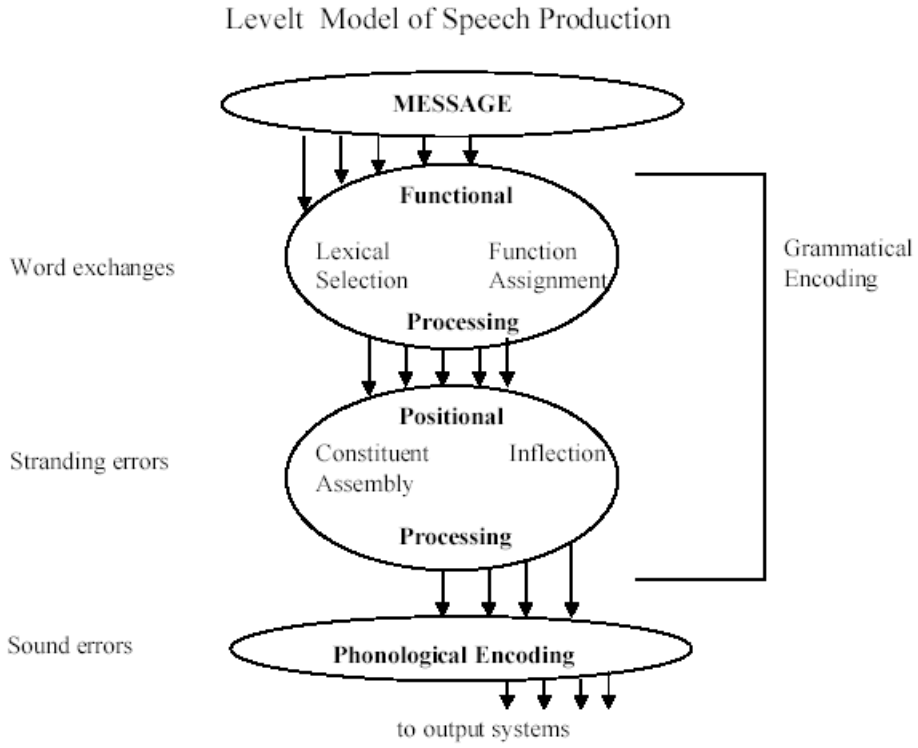
speakers create utterances optimally given constraints on processing capacity, while at the same time respecting grammatical constraints, at least most of the time.

This chapter is a review of syntax in language production, and focuses on three critical issues. The first addresses what a syntactic representation for production looks like. We consider to what extent it is hierarchical and ordered. Then we discuss whether it contains empty categories such as traces of constituent movement, as would be expected according to theories of syntax which assume that questions, relative clauses, passives, and other complex structures are formed via movement. We also evaluate whether grammatical information is represented and accessed in a way that is purely form-based, or whether structures are actually small trees associated with specific lexical items which get combined to create utterances (F. Ferreira, 2000). In the second major section, we ask whether grammatical encoding is an automatic or a resource-demanding process, and we consider whether certain structures are inherently more difficult to produce than others. We also try to determine whether the system generates structure incrementally, or whether production involves planning units larger than a single word (e.g., clauses). In the third section, we review the now quite extensive literature on syntactic choice in production, examining how people choose between syntactic forms such as active and passive, and how they decide whether to include optional words like complementizers (e.g., *Mary knows (that) her cat terrifies the dog*). In the final section, we briefly consider the question of syntax affects prosody during production. We conclude that syntactic structure is created in two stages, using lexically stored syntactic templates that include representations of empty categories. This process of grammatical encoding is computationally demanding, which is why the system makes optimal use of scarce resources by being moderately incremental. And the prosody of a sentence reflects its syntactic organization, both because syntax affects prosodic constituency, and because difficulties associated with syntactic planning can affect phonological phrasing.

II. Syntactic representations in production

A. Two-stage models of grammatical encoding

As has been argued since the earliest days of generative grammar, syntax is an interface between meaning and sound (articulation / phonetic form). A word such as *cat* has a particular meaning, but the expressive power of language is enhanced immeasurably by our ability to create meanings compositionally, by putting words together – for example, our ability to say *not a cat* or *that's my cat*. Models of production instantiate this basic architecture fairly transparently. Consider the Bock and Levelt (1994) model, which was described in the previous *Handbook of Psycholinguistics*. The model, henceforth referred to as BL, is shown below (reprinted from the 1994 chapter).



The process of speaking begins with a message-level representation, which captures the idea the speaker wishes to convey. This message becomes sound at the other end of the model, at a stage called phonological encoding. Linking the message and phonological levels are two stages of syntactic processing (or grammatical encoding, as it is called in the model), one called functional processing, and the other called positional processing. Notice that the basic linguistic architecture in which syntax mediates between meaning and form is replicated in the BL model of production. Yet an important difference is that syntactic operations are factored into two components. This two-stage architecture originated with Garrett (1975), who argued from speech error data that the production system first creates the global, syntactically functional structure for a sentence, and in a separate stage determines phrasal details such as serial order (Footnote 1).

In BL, grammatical encoding begins with functional-level processing. Abstract lexical entries termed *lemmas*, which contain information about a word's meaning and its syntactic requirements but do not represent its phonology, are retrieved and assigned grammatical functions such as subject and object (Footnote 2). For example, for the utterance *my cat terrifies the dog next door*, the lemmas for CAT, DOG, NEXT, DOOR, and TERRIFY would be retrieved, and CAT would be assigned the role of subject and DOG (modified by NEXT DOOR) the role of object. At this point, then, the speaker has committed to some type of active structure in which CAT will be the subject; a passive structure is ruled out, because in any type of passive DOG would be the subject. But notice that a structure such as *the dog next door, my cat terrifies him* is still possible, because in this form CAT is the grammatical subject and DOG is the object (*dog* is the object in both the preposed position and in the pronominal form *him*). The difference between the regular active and this left-dislocated construction is a matter of constituent ordering, which is left undecided at this stage of grammatical encoding.

The second component of syntactic processing takes place at the positional level, which operates on the functional-level representation. At this point, serial order is imposed on the utterance. Beginning with the initial constituent, each grammatical function created earlier (e.g., subject, object, modifier) is translated into a linearized constituent. The grammatical encoder retrieves a prestored phrasal frame, which contains slots for all the elements of that phrase – the determiner *my* and the noun *cat*, in the current example. Inflectional affixes are represented as an intrinsic part of the frame, so that if the subject were plural, the plural morpheme would already be in place and would therefore not have to be separately retrieved and inserted.

Because the language production system is assumed to be incremental (see section III.C for a more thorough discussion of incremental production), the order in which lemmas are 'worked on' determines the overall order of the phrases in the utterance (F. Ferreira, 2000). So if the lemma for DOG were processed before the one for CAT, then the resulting structure might be the left-dislocation form given above or perhaps a topicalized form such as *Mary my cat loves. (*Although this construction is disfavored in most dialects of English, it can be acceptable given the appropriate context). Thus, positional-level processing determines both the serial order of phrases and the order of elements within any given phrase, and all inflectional processing takes place at this stage of processing as well. For many ordering decisions, the processor simply obeys grammatical constraints such as the requirement that determiners initiate an NP, that adjectives precede nouns, and that verbs precede objects but follow subjects (for English). But because languages give speakers some ordering options, there may be decisions about order that still need to be made, particularly at the within-phrase level. One good example is the sequencing of conjuncts, illustrated in *the dog and cat slept soundly*. The other order of *dog* and *cat* is equally grammatical, and so the choice about how to sequence the conjuncts within the overall NP

subject must be based on extra-grammatical considerations (Pinker & Birdsong, 1979; Bock, 1987).

B. Evidence for two-stage models

The evidence for this two-stage architecture separating functional and positional level processing comes from two sources: speech error analyses and data from experiments designed to shed light on how structure is created. The argument from speech errors is as follows. First, speakers sometimes make semantic substitutions, as in *my cat terrifies the boy next door* when *the girl next door* was intended. These errors almost invariably respect a form-class constraint: nouns substitute for nouns, verbs for verbs, and so on. Speakers also sometimes make word exchange errors, illustrated by *my boy terrifies the cat next door*. The interacting elements in these errors usually come from different phrases, and the words tend to be of the same form class. Semantic substitutions and word exchange errors indicate that there is a level of processing at which grammatical category is relevant and at which the roles for lemmas are decided. In contrast, errors such as phonological substitutions and stranding occur as well, but they have quite different characteristics. In a phonological substitution, a word with a similar sound is incorrectly assembled and made part of the utterance. In stranding errors, content morphemes end up misordered but inflectional material does not, as in *I went to get my park trucked* (Garrett, 1980). Notice that the morphemes *truck* and *park* swapped places, but the suffix *–ed* is in its correct location. Phonological substitutions and stranding errors indicate that there is a level of processing at which sound and serial order are decided, and stranding suggests that the inflectional morpheme is an intrinsic part of the phrasal frame.

The experimental evidence for distinguishing functional and positional level processing comes from priming studies, both lexical and syntactic. Let’s begin with lexical

priming. Notice that the two-stage architecture divides up lexical processing so that word *meanings* become available at the functional level, but word *phonology* only gets generated after (and probably after most positional-level processing takes place as well; F. Ferreira, 1993). This is because the sounds of words are (arguably) not relevant for deciding grammatical functions such as subject and object, but (again, arguably) phonology may help the system decide how to sequence words, as suggested by the finding that, in conjuncts, short words tend to precede longer ones (Cooper & Ross 1975; Bock, 1987), for example. Experiments in which words are either semantically or phonologically primed have demonstrated that making a lemma available (i.e. semantically priming a concept) causes the constituent containing that lemma to be the subject of the sentence. In contrast, phonological priming has either weak effects or leads to late positioning of the constituent containing the word (Bock, 1987; cf. Cleland & Pickering, 2003). This pattern is typically taken to support a division of labor between functional and positional level processing, because the idea is that only a manipulation which affects the lemma can influence processes hypothesized to be taking place at the functional level. Of course, this interpretation is somewhat compromised by the finding that phonological priming sometimes leads to late constituent placement, but the effect is much smaller and has been argued to reflect a late stage in production where an utterance is evaluated and then changed if it is judged to be deficient before it is articulated (Levelt, 1989).

Another source of evidence for the two-stage architecture comes from syntactic priming. If a speaker produces or even simply hears an utterance with a particular structural form, he or she is likely to mimic that structure in a subsequent utterance. The classic demonstration (Bock, 1986b) involves both the active/passive and the prepositional/double-object dative alternations. Speakers will tend to produce a passive sentence after hearing or producing one themselves (Levelt & Kelter, 1982; Schenkein, 1980); the same goes for the prepositional dative (*the driver showed the overalls to the mechanic*) and the double-object

dative construction (*the driver showed the mechanic the overalls*). (It is not clear that the active can be primed, possibly because of ceiling effects due to its high frequency.) Hartsuiker and Westenberg (2000) discovered using Dutch that a very low-level ordering decision (the sequencing of an auxiliary and a main verb at the end of a sentence) can be primed, leading them to argue for a two-stage model of syntactic processing where a 'dominance-only' representation (i.e. one that is not linearized) can prime a representation that is ordered.

Further evidence for the two-stage architecture comes from the way speakers compute agreement relations during sentence production. Consider the fragment *the spokesman who defended the actions*. If this fragment is the subject of a sentence, then it must agree with the main verb. In English, this agreement process is visible mainly on forms of *to be* and *to have* (particularly in the past tense), but in other languages agreement is overt on a wide range of verbs and other words. Carefully designed experiments have revealed that agreement errors occasionally happen, particularly in examples like *the spokesman who defended the actions*, in which the head noun *spokesman* is singular but there is another noun in the subject (*actions*) that is plural (Bock & Eberhard, 1993; Bock & Miller, 1991). Agreement errors turn out to be just as likely in yes/no questions as in declaratives, suggesting that agreement is computed on a representation that specifies dominance but not linear relations (Vigliocco & Nicol, 1998). This argument can be seen in the contrast between *the helicopter for the flights are safe* and *Are the helicopter for the flights safe*, where the linear positions of the head noun are different but the likelihood of an agreement error is the same. This result suggests that agreement relations are computed from a syntactic representation created before linearization takes place. Notice that this particular finding is consistent not only with a two-stage view of syntactic processing, but also some version of a transformational / derivational account of grammar, because the linearization process at issue here is the one that moves the verbal material to the front of

the sentence to create an interrogative construction. This general idea will be discussed in section II.D when we consider the question whether syntactic representations created during production show evidence of processing attributable to constituent movement. Additional evidence for the idea that hierarchical position but not linear order is critical for computing agreement can be found in a study of complex NPs such as *the computer with the program(s) of the experiment(s)* (Franck, Vigliocco, & Nicol, 2002). Agreement errors were found to be more likely when the medial noun *program* was plural compared to the more proximate noun *experiment*, indicating that position in a hierarchical structure has more effect on agreement than does linear position.

Thus, evidence from speech errors, from syntactic and lexical priming, and from the process of computing subject-verb agreement seem to converge on the idea that syntactic structure is generated in two distinct stages during production. Nevertheless, this architecture has been challenged, and we turn now to evidence that is argued to support a single-stage model of grammatical encoding.

C. Evidence challenging two-stage models

First, recall the effects of semantic and phonological priming on grammatical form. If it had turned out that only semantic primes could affect the establishment of grammatical relations, then an architecture separating syntactic generation into a stage that uses only lemma information to assign roles such as subject and object, and a separate stage that uses sound to determine linear order, would have been supported. But recall that phonological primes do have a small but significant effect (Bock, 1987). For example, if participants encountered the word *trump* and then a picture of a truck towing a car, they were likely to say *the car is being towed by a truck*, because the phonological relationship between *trump* and *truck* leads to some type of inhibition. Thus, the effect of a phonological

prime appears to be opposite from one that is semantic, but the important point is that according to the classic two-stage architecture, it should have no effect at all. Therefore, it may be argued that this finding undermines two-stage models.

However, there are two problems with this argument. The first was briefly mentioned earlier: It is possible that this effect of the phonological prime occurs not during grammatical encoding but during a stage at which the utterance is checked for overall acceptability (the so-called monitor; see Hartsuiker, Corley, & Martensen, 2005, for a recent discussion of its properties). The second problem with this argument is that the inhibitory effect of phonological primes only challenges the assumptions regarding lexical processing during grammatical encoding – specifically, the idea that lexical retrieval occurs in two stages, with only the second including access of phonology. It is possible that semantic, syntactic, and phonological information about words is all retrieved simultaneously, but that dominance and linear relations are nonetheless computed separately. An important question too is why a phonological prime should be inhibitory rather than facilitatory. Bock (1987) suggested that the effect could be due to lateral inhibition among phonological competitors, but some studies of lexical processing have shown that phonological primes facilitate processing (Grainger & Ferrand,1996; Tanenhaus, Flanigan, & Seidenberg, 1980). The Bock (1987) finding clearly should be pursued further, and indeed, it has not yet even been replicated.

The second finding that has been argued to undermine the two-stage model of syntactic processing concerns priming in the dative structure (Pickering, Branigan, & McLean, 2002). Consider once again the prepositional-dative, illustrated in *the driver showed the overalls with the stains to the mechanic*. Another grammatical alternative is the shifted form in which the prepositional phrase (PP) precedes the object (*the driver showed to the mechanic the overalls with the stains*), a structure that is more likely to be generated the longer and heavier the object (Wasow, 1997). Shifted and non-shifted prepositional datives share the same hierarchical or dominance relations but differ in how the NP and the

PP are ordered. Thus, on a two-stage view in which dominance relations are computed separately, the shifted version should prime the non-shifted version. However, such priming does not occur. Based on these results, Pickering et al. (2002) argued for a single stage model in which dominance and linear relations are computed simultaneously. But the results could be attributed to the peculiarities of the shifted dative form, which is not only fairly rare (even in the Pickering et al. experiments in which measures were taken to elicit them) but also seems to require fairly strict discourse conditions to be felicitous (Hawkins, 1994). These properties of the shifted prepositional dative might compromise its ability to prime any other construction. It would be very useful to see whether this result can be found using a less marked structure. Exploring this possibility might require consideration of languages that allow more flexibility in constituent ordering than English does.

D. Do syntactic structures contain evidence of constituent movement?

Perhaps the most distinctive characteristic of generative grammar compared to other approaches to syntax is its assumption that syntactic structures are generated by movement. Anyone who has taken even an undergraduate course in cognitive psychology knows that in the earliest versions of this theory, noncanonical structures such as passives were created by rearranging the basic active structure (Chomsky, 1965). Somewhat less well known is the transition to the Government and Binding (GB) theory, which assumed that syntactic representations contain evidence of movement. For instance, a passive such as *the dog next door$_i$ was bitten t$_i$ by my cat* requires movement of the NP *the dog next door* from the post-verbal position to the subject position, but the starting position of the phrase is marked in the representation with a trace (indicated with the *t*). The trace allows the phrase to be interpreted as the object of the verb *bite* even though it is no longer in object position in the surface structure. The same holds for structures such as wh-questions and relative

clauses: A sentence such as *which dog$_i$ did my cat bite t$_i$?* is created by moving the wh-NP (*which dog*) to the top of the tree, again leaving behind a trace so that *dog* can be interpreted as the object of *bite*. It has been common in the psycholinguistic literature to refer to traces as gaps and to moved constituents as fillers (J. D. Fodor, 1978, 1989, 1991), and so we will follow this convention for the rest of our chapter.

The question we now turn to is, do the syntactic structures that people create when they talk contain any evidence of constituent movement? It is widely believed that they do not. For example, it has been argued that one way to conceptualize the two-stage architecture for grammatical encoding is to assume that the first stage creates a 'deep-structure' representation and that the second creates a 'surface structure' representation. It is important to note, however, that the concept of a 'deep structure' has really not been part of generative grammar for the last 25 years, and so it would be surprising to find any evidence for it in language production. And, not unexpectedly, we do not.

Bock, Loebell, & Morey (1992) used the syntactic priming technique to distinguish between the direct and mediated (first a deep structure is computed, then a surface structure) approaches to syntactic generation. Participants heard sentences and then repeated them, and then they had to describe an unrelated picture of a simple transitive event. The critical feature of the study was that if the pictures were described in the active voice, the subject would be inanimate and the object animate (e.g., *the clock woke up the boy*). The prime for the picture description (i.e., the heard and repeated sentence) was either active or passive, and it had either an animate or an inanimate subject. The assumptions behind the design were that both structural form and animate placement would be mimicked in the picture descriptions. Bock et al. reasoned that if the 'deep structure then surface structure' hypothesis is right, then both an active with an inanimate object AND a passive with an inanimate surface subject would prime the active picture description, because the passive actually has an inanimate object at deep structure. However, this

pattern was not observed; instead it was the surface placement of the animate entity that determined the degree of animacy priming. Hence, Bock et al. concluded that the direct mapping account is correct.

But these results might simply indicate that the old-fashioned transformational model in which passives are generated by modifying an active kernel sentence is indeed no longer viable. More problematic for the view that constituent movement is psychologically real in production is the finding that a passive such as *the plane was landed by the pilot* can be primed with a non-passive such as *the plane was landing by the tower* (Bock & Loebell, 1990). This result suggests that the newer GB version of generative grammar is also wrong for production, because the presence of the gap in the passive sentence does not seem to matter for the priming effect. And this lack of evidence for gaps is consistent with the BL model of production, which is based not on a generative grammar but rather on Lexical-Functional Grammar (LFG; Bresnan & Kaplan, 1985), since LFG does not assume the existence of movement operations or gaps in syntactic representations.

At the same time, there is also some compelling evidence for the presence of gaps in the syntactic representations that speakers generate. One type of evidence is admittedly intuitive, but we believe the effect is so strong, an experiment is hardly necessary (but see F. Ferreira, 1988, for experimental evidence). To begin, consider the way we normally pronounce a sequence such as *to the party* – notice that the preposition *to* undergoes vowel reduction so that it ends up sounding more like ‘t-schwa’ than the citation form /too/. This phonological process is essentially mandatory; It would be odd to pronounce the sentence as ‘too-the-party’, especially at normal rates of speech. But now consider a sentence involving wh-movement, such as *John knows who$_i$ Mary talked to t_i at the party*. The wh-word *who* is semantically and lexically the object of the preposition *to*, which according to GB theory means that a gap must occur in that position if *who* moves (as it must in English). But now notice how the word *to* is pronounced: it is not reduced but rather lengthened, so it

ends up sounding more or less like its citation form /too/ (F. Ferreira, 1988). This means that the trace was represented in some way in the syntactic structure guiding the creation of prosodic structure, which allowed it to block the normal process of vowel reduction. The same seems to hold for NP-gaps as well, as in *The boy$_i$ was spoken to t_i by his teacher*.

Recent evidence from errors of subject-verb agreement also support the idea that traces are mentally represented during grammatical encoding (Franck, Lassi, Frauenfelder, & Rizzi, in press). An experiment designed to elicit such errors from French speakers showed that displaced direct objects in a cleft construction (*It's the deputy that the senators welcome t*) determine whether errors of agreement occur, even though the object does not intervene between the head and the verb in the surface word sequence. Franck et al. argue that their results can only be accounted for if we assume not just a single transformational process that turns a set of lexical items into a surface structure, but rather a grammar consistent with the Minimalist Program (Chomsky, 1995) in which syntactic structures are generated through a series of operations termed MERGE, MOVE, and AGREE (only the latter two are relevant for our purposes; MERGE simply refers to the combining of lexical items). According to Minimalism, elements in a syntactic structure cyclically move up the tree until they are in the appropriate position to allow an agreement relation to be checked. Each movement, even intermediate ones, leaves behind a trace. Franck et al. argue that the pattern of agreement errors they observe in French and Italian can only be explained if it is assumed that gaps of intermediate movement interfere with the process of computing agreement. These results are some of the strongest evidence to date that gaps are generated as part of the normal process of creating a syntactic structure for a sentence.

What should be said, then, about the work suggesting that gaps do not exist? At this point, the most unbiased assessment of the state of our knowledge is that this entire issue needs to be examined in much more detail. In fact, it is worth noting that no experiment has ever been conducted to test directly whether gaps are psychologically real in language

production (although the Franck et al. paper comes close). One potentially useful observation about our current state of knowledge is that the evidence for and against traces comes from different sources – the evidence against the reality of gaps is based largely on results from syntactic priming, and the evidence for them has come from studies of subject-verb agreement as well as the process of translating syntactic structures to prosodic constituents.

Of course, the traces of wh- and NP-movement are not the only types of empty categories that have been proposed in generative grammar; another important phonetically null category results from ellipsis, as in *Mary can tie her shoes and Natalie can too.* Comprehension studies have shown that people reconstruct the missing material, eventually obtaining the interpretation that Natalie can tie her own shoes (perhaps by first entertaining but then rejecting the so-called 'strict' reading on which Natalie ties Mary's shoes; Shapiro, Hestvik, Lesan, & Garcia, 2003). The production of ellipsis has not been studied at all, so the extent to which the omitted or deleted material is mentally represented is not known. Consider our shoe-tying example. It is clear that lexical forms corresponding to the second verb phrase *tie her shoes* are not retrieved. But at the message level, the speaker almost certainly generates the idea that Natalie is capable of tying her own shoes. The question those interested in grammatical encoding might ask is, what about the levels in-between? Is a VP generated in the second clause and then left unpronounced because the lemmas do not point to any word-forms, as Levelt (1989) suggests? Another intriguing question is what leads speakers to choose one type of ellipsis over another; for example, an alternative to the form above is *Mary can tie her shoes and so can Natalie*. Clearly, many important issues concerning the grammatical encoding of empty categories remain to be even formulated in the field of language production.

E. Are syntactic structures lexically anchored?

Another question that has been of interest both in formal linguistics and in psycholinguistics is whether syntactic structures are linked to words – specifically, lemmas. The classic work by Levelt (1989) argued for lexical generation of syntax. In contrast, the BL model of production assumes a non-lexical view of syntactic structure. In BL, trees are conceptualized as 'control hierarchies' which contain no lexical content but instead coordinate the insertion of lexical material which is retrieved and assembled separately (BL, pp. 947-948). This conceptualization is in part based on Bock's earlier findings suggesting that lexical overlap does not enhance syntactic priming (Bock, 1989; Bock & Loebell, 1990). For example, the amount of priming for a sentence such as *the girl handed the paintbrush to the man* is the same given a prime like *the secretary baked a cake for her boss* and *the secretary gave a cake to her boss*, even though in the latter case the prepositional phrases share the same prepositional head (*to*).

However, more recent work using the syntactic priming paradigm suggests that verb identity does increase the magnitude of priming (Pickering & Branigan, 1998, 1999; Cleland & Pickering, 2003). One motivation for examining this issue in careful detail is that many formal theories of syntactic structure assume that words and the syntactic environments in which they may occur are lexically linked. In the earliest versions of transformational grammar, for example, verbs specified the syntactic environments in which they could occur via their subcategorization frames (Chomsky, 1965). A verb such as *put* would be represented as requiring an object and a locative prepositional phrase. The theory of Government and Binding essentially dissolved phrase structure rules altogether in favor of lexical storage of constituent structure, so that all words were represented in the lexicon with their associated arguments (Chomsky, 1981; Stowell, 1981). Retrieval of a word would then automatically bring along its associated structures. This elimination of phrase structure rules was a logical extension of X-bar theory (Jackendoff, 1977), which described a universal

format for all phrases regardless of their type. Other models of syntax such as Lexical-Functional Grammar (Bresnan & Kaplan, 1985), Categorical Grammar (Moorgat, 1988; Steedman, 2000), and Generalized Phrase Structure Grammar (Gazdar, Klein, Pullum, & Sag, 1985) also connect words and their syntactic environments. The same trend is evident in Tree-Adjoining Grammar (Joshi, Levy, & Takahashi, 1975; Joshi, 1985). In TAG, the primitive objects of the grammar are treelets which consist of a word (a lexical head such as a verb) and the arguments the head licenses.

Given these theoretical perspectives on the representation of words and syntactic structures, and in particular verbs and arguments, it makes sense to expect that syntactic priming would be greater when the main verb in the prime and target sentences overlap. This issue was investigated in a study designed to assess whether priming would be observed in simple dialogue situations (Branigan, Pickering, & Cleland, 2000), and which manipulated verb identity. A confederate and a genuine experimental subject described pictures to one another, and it was found that the naïve participant tended to use the same construction as the confederate. In addition, priming was greater when the verb in the prime and target sentences was the same – the effect was about twice as large. The model of syntactic generation offered by Pickering and Branigan (1998) assumes that words such as verbs are linked to the phrases with which they may combine (termed *combinatory nodes*). Cleland and Pickering (2003) demonstrated remarkably similar priming effects for noun phrase structure, including enhancement when the head of the noun phrase (the noun) was shared, indicating that this form of representation and the process for creating structures are similar both for clauses and phrases.

These results, then, tell us that the syntactic representations used for language production are ones in which structures may be generated directly from lemmas rather than through the accessing of contentless phrasal templates (Levelt, 1989). F. Ferreira (2000) presented a model for human language production which uses TAG as the database for

creating structures through lemmas (see also F. Ferreira, Lau, & Bailey, 2004). All heads, including verbs, nouns, prepositions, and adjectives, are represented with the arguments that they license. These elementary trees consisting of a head and its licensed arguments are combined to form utterances (see F. Ferreira, 2000 for a description of the operations that combine elementary trees). The model proposed by Pickering and his colleagues is somewhat different, but it shares the basic insight that words and syntactic structures are representationally linked. Thus, although it may be possible for grammatical encoding to take place using control structures that have no lexical content, perhaps via extraction of some type of general schema for forming particular construction types, in general it appears that the syntactic structures used for production are lexically anchored.

III. Processing resources for grammatical encoding

Thus far, we have considered mainly representational issues, focusing particularly on the properties of syntactic structures and the format in which syntactic information is stored. The question we turn to in this section is how these structures are formed, and in particular, how the process of grammatical encoding draws on processing resources and how computational load is managed. This will lead us to consider the degree to which grammatical encoding is incremental.

A. Is Grammatical Encoding Automatic?

Talking generally feels effortless, but even the most fluent speakers occasionally experience some difficulty formulating their utterances. The classic work of Goldman-Eisler (1968) demonstrated that almost half of most people's speaking time is devoted to pauses and disfluencies such as *um* and *er*. Ford (1982) measured spontaneous speech and

observed that about 20% of all clauses are preceded by a pause of about one second in duration. This finding suggests that the process of creating syntactic structure is resource-demanding, but the result is not definitive; clauses are both major syntactic and semantic junctures, and it might be that grammatical encoding is automatic but semantic processing requires planning and can therefore be resource demanding. This view was articulated by Levelt (1989). He adopts Kempen and Hoenkamp's (1982, 1987) model of grammatical encoding which assumes that syntactic procedures are modular, and thus have the characteristics that J. A. Fodor (1983) views as typical of cognitive modules: grammatical encoding processes consult a proprietary vocabulary, they operate whenever they recognize their standard input, and they operate automatically. Thus, it is to message level planning that resources are devoted during production; syntactic decisions are made automatically, and measures of processing load such as reaction time and pausing will not be affected by the complexity of syntactic operations.

Work conducted since the publication of Levelt's book, however, does appear to suggest that syntactic planning demands computational resources, especially as structural complexity increases. F. Ferreira (1991) had people memorize declarative sentences, which they then had to produce upon receipt of a visual cue; latency to begin speaking was measured. The variable that was manipulated was the syntactic complexity of the sentential subject: It was either short (*the river*) or long, and in the long conditions, it was either of low (*the large and raging river*), medium (*the river near their city*), or high (*the river that stopped flooding*) syntactic complexity. Complexity was defined in terms of a node count, so that the more syntactic nodes the subject needed in its representation, the greater its complexity. Ferreira found that as complexity increased, so did production latencies. Interestingly, memorization times were not affected by this variable, suggesting that the effect was particular to the task of speaking. In a second experiment, she orthogonally varied the syntactic complexity of the subject and object in subject-verb-object sentences. Once again,

latencies to begin speaking increased with the complexity of the subject, and the object's characteristics had no effect. However, when both the subject and the object were syntactically complex, speakers tended to pause within the sentence, and the preferred pause location was the subject-verb phrase boundary.

Because speakers were not required to generate any of the sentences' content, these effects of syntactic complexity cannot be attributed to any semantic complexity that might be correlated with the syntactic manipulation. Moreover, as speakers had not chosen the syntactic forms themselves either, the effect cannot be attributed to the need to make syntactic decisions. Instead, it appears that simply saying a sentence with a complex structure takes up processing resources. The results also demonstrate that if both the subject and object are complex, speakers divide the utterance into two processing units, one consisting of the subject, and the other consisting of the verb phrase. Notice that this division into processing units respects the syntactic structure of the sentence, and indeed Ferreira observed almost no cases in which participants paused after the main verb rather than before it (see also F. Ferreira 1993 for further discussion of these issues). This finding that the processing units are syntactic constituents is consistent with the assumption that the difficulty in processing is localized to the syntactic level.

In another study, Smith and Wheeldon (2001) tested whether speakers plan their utterances before saying them. Participants were asked to describe pictures, and they were primed with sentences such as *The spoon and the car move up*. When participants uttered sentences that were syntactically similar to the prime sentence, a reliable 50 ms advantage to begin speaking was observed. Smith and Wheeldon also tested the scope of this effect and found that it held only for the first phrase of an utterance, consistent with F. Ferreira's (1991) second experiment demonstrating that only the complexity of the subject affects utterance initiation times. The advantage that Smith and Wheeldon found for sentences that had been syntactically primed suggests that we should revisit the phenomenon of syntactic

priming in light of this question concerning processing resources. Recall that speakers are more likely to produce a particular construction when they have just heard it or produced it themselves. Based on the finding that syntactic priming is particularly robust in dialogue, Pickering and colleagues have suggested that syntactic priming is used by the production system as a tool for "reducing the load associated with syntactic processing" (Pickering & Branigan, 1999, p. 136). What this idea assumes, of course, is that syntactic generation is a resource-demanding process, so much so that speakers try to find ways of managing and reducing the computational burden.

It should be noted that studies predating the recent psycholinguistic era also demonstrate that syntactic processing is computationally demanding. Johnson (1966) compared the generation of sentences such as *The person who jumped over there is good* and *The person over there who jumped is good*. Because the structure of the second sentence is right-branching, it is less complex according to the Yngve (1960) complexity metric, and so Johnson predicted it would take less time to initiate. This prediction was confirmed. Second, although Rochester and Gill (1973) did not find any effects of what they termed "syntactic complexity" on speech hesitations and disruptions, they did find that such disruptions in speech varied along with the type of nominal modifier people produced. Specifically, speakers were more likely to show speech disruptions before a noun phrase complement (e.g., "The fact that the woman was aggressive threatened the professors") than before a relative clause (e.g., "The book that was written by Millet was lauded by all"). Goldman-Eisler (1968), who like Rochester and Gill (1973) failed to find effects of syntactic *complexity* on hesitation, also found hesitation differences before different types of syntactic forms. If disruptions in speech are a measure of mental load, and more disruptions occur before one particular ordering of words than another, then one structure must have required the use of more mental resources than the other. We turn next to a more detailed consideration of this question concerning the inherent difficulty of certain forms.

B. Are Some Constructions Difficult to Generate?

In this section, we ask a question that has received surprisingly little attention from experimental psycholinguists: Are some syntactic constructions inherently difficult to produce, or does difficulty arise only when a structure must be generated in an infelicitous discourse context? To see what is at issue here, consider the passive construction, which is often viewed as more complex than the active, and is certainly more difficult to understand (F. Ferreira, 2003). The passive may be harder to produce than the active because it has a noncanonical structure, because it is less frequent, or because it is more complex, in the sense of requiring more syntactic nodes in its phrase-structure representation. Alternatively, it has been argued that the passive may be the *right* construction for particular discourse situations. For example, Tomlin (1983) observed that passives are very common in hockey broadcasts, because what the commentator tries to do is make the player in possession of the puck the subject of the sentence. If that player is affected in some way (e.g., gets checked), then the sentence form the commentator will use is a passive (*Gretzky was checked by his opponent*), because that is the form that allows the topic to be maintained as subject, even when it is not an agent. But Tomlin's study did not examine whether passives are harder to produce even when they are licensed by the discourse. To answer this question, it is necessary to measure processing load rather than just frequency of occurrence.

Tannenbaum and Williams (1968) conducted one relevant study. Speakers first read a story that was either about trains, cars, or a topic that was relatively neutral. They then saw a picture of a train hitting a car, and their task was to produce either an active or a passive sentence (cued by a letter superimposed on the picture). They found that latencies to produce the active were fastest in the subject-focus condition and about equally long in

the object-focus and neutral conditions. Passives were produced fastest in the object-focus condition, next fastest in the neutral condition, and most slowly in the subject-focus condition. This finding would appear to suggest that as long as a construction occurs in the appropriate context, it is easy to produce. However, a closer examination of their data suggests a different conclusion. Although this pattern was observed, it was also found that actives in the “wrong” discourse were produced as quickly as passives in the “right” discourse; indeed, in no condition were passives initiated faster than actives. The picture that emerges, then, is that noncanonical structures can be inherently hard to say, even in proper contexts. A construction that is rare or that is syntactically complex (or both, as these two characteristics tend to co-occur) requires a specific sort of context but is still difficult to generate, perhaps because more syntactic nodes take more processing resources to create, or because the production system has less experience generating forms such as the passive.

This finding is compatible with a study that investigated whether certain verbs license passives more easily than others, as might be expected on a lexicalist view of grammatical encoding (F. Ferreira, 1994). Participants were asked to generate sentences out of three visually presented words – e.g., LAYOFFS MANAGER WORRIED. The verb either had a theme-experiencer argument structure (as in *worried*) or a more conventional agent-patient structure (LAYOFFS MANAGER ORDERED). Speakers produced passives more often when the verb was theme-experiencer, which was predicted based on the idea that speakers attempt to place the more prominent thematic role in the subject position of the sentence, and experiencers are more prominent than themes (Grimshaw, 1990). Nevertheless, passives took longer to formulate than actives, suggesting that even though certain lexical conditions might license them, they still seem to take more time to grammatically encode.

Clearly, however, this issue needs to be examined in more detail, particularly now that there is such intense interest in the idea that the frequency of exposure to a syntactic construction affects how easily it can be comprehended (MacDonald, Pearlmutter, & Seidenberg, 1994; Mitchell & Cuetos, 1991). As Race and MacDonald (2003) have pointed out, these distributional patterns come from speakers – they reflect the choices speakers make in different circumstances. The Race and MacDonald approach to processing assumes that comprehension and production must be examined together, and they predict that the forms that are hard for people to produce are also the ones that are hard to understand. This parity is based on speakers' tendency to avoid difficult structures, thus creating distributional patterns. But this interesting research program is predicated on the idea that some forms are inherently difficult to produce – for example, an object relative such as *The story the quiet boy read was long*, which does not contain the relative pronoun *that* – and that is why it is less likely to be said. The empirical question that arises is whether these sentences are harder to say when they are generated, or whether the discourse conditions which obtained at the time the sentence was said in fact made the structure easy to encode and articulate.

C. Incremental Production

Incremental production may be viewed as a way to reduce the processing resources required for production. The idea is that at particular points in time certain concepts may be more available to a speaker than others, and the grammatical encoder tends to begin with those accessible lemmas. Incrementality is viewed as optimizing the use of processing resources, because it allows the system to begin with the 'easy bits', so to speak, and to deal with the more difficult portions of the utterance during articulation (F. Ferreira & Henderson, 1998). Of course, as a reviewer of an earlier version of this chapter pointed out,

incrementality might create a situation in which an accessible constituent forces a syntactic structure that is computationally demanding (e.g., the passive). But the reason incrementality will generally still lead to efficient processing is that the difficulty of making a passive can be 'spread out' over the entire utterance rather than being localized entirely to the point of its initiation. As a result, there need not be any hesitation or disfluency before utterance production, and the demands of managing the rest of the structure can be distributed over the remaining constituents, with planning going on in parallel with articulation.

Recent work suggests that the degree to which the system is incremental is under strategic control, as would be expected if incrementality is a way for the production system to manage its resources (F. Ferreira & Swets, 2002). Participants were asked to calculate the answers to arithmetic problems and to provide the answer in the form of a sentence (*The answer is 58*). The problems always included at least one two-digit addend (e.g., 53+5), so participants were unlikely to be able to retrieve the sum. In one experiment, participants were allowed to begin to speak whenever they felt ready, and the data provided no evidence for incremental production: Initiation times were longer the more difficult the entire problem, but durations were unaffected. This pattern indicates that the entire utterance was planned before articulation. In a follow-up experiment, speakers were required to begin to speak before a deadline (indicated by a punishment 'beep'). This manipulation dramatically reduced initiation times overall, from over 2 seconds in the first experiment to about 700 ms in the experiment with the deadline (interestingly, accuracy was not compromised). Nevertheless, initiation times still reflected the difficulty of computing the sum. At the same time, the duration of the earlier part of the utterance was also affected by problem difficulty, suggesting that speakers postponed some planning of the sum until they were actually speaking. This study suggests that the degree to which the system is incremental depends on the speaker's strategy. If a premium is placed on beginning to

speak quickly, then the production system does indeed become more incremental; but if speakers have the opportunity to plan, they seem to prefer to do so. Moreover, the system engages in some planning even under conditions most conducive to incremental production – that is, when there is a premium on initiating speech quickly.

On some views of incrementality, constructions are chosen indirectly; they emerge from the speaker's attempt to place a highly accessible concept in the most prominent syntactic position. If that concept happens to be a theme or patient, in a relatively fixed word order language such as English a passive structure will need to be produced to accommodate a thematic patient in subject position. Under this view of human sentence production, the lemmas associated with the most accessible concepts automatically grab the earliest positions in utterances. However, Bock (1986b) questioned this radical version of incrementality: "Typically, speakers do not simply produce words in the order in which they come to mind….Rather, the syntactic forms of sentences seem to be changed so as to accommodate word order variations without altering the intended meaning" (p. 359). But a radically incremental model is assumed, for example, by van Nice and Dietrich (2003), who interpret their German data as supporting the view that "the first-conceptualized referent will continue onward as the first-lexicalized and, ultimately, as the first in word order"(pp. 829). This view, they point out, is also held by Kempen and Hoenkamp (1987) as well as de Smedt (1996).

Christianson and F. Ferreira (in press) attempted to resolve this controversy by examining production in Odawa, an Algonquin language which allows constituents to be ordered freely (i.e., any arrangement of subject, verb, and object is grammatically licensed). Speakers were asked questions about a pictured event. The questions topicalized either the agent, the patient, or neither (this latter question was simply, *What happened?*). Even though Odawa speakers have access to any word order arrangement of subject, verb, and object, their descriptions were similar to those observed for English speakers. Given the no-

topic and agent-topicalizing questions, actives were the forms most commonly produced; in the patient-topicalizing question condition, passives were preferred. Thus, even though speakers of Odawa could have produced active sentences with the patient in the first position (i.e., OSV or OVS sentences) when the patient was topicalized, they in fact chose to produce passives, which not only are about as rare in Odawa as they are in English, but also require the omission of the agent argument altogether (because passives in Odawa do not permit any type of by-phrase). Thus, a highly available constituent primes a particular syntactic form, and if that constituent is a patient, the form that will be generated is a passive.

These findings are inconsistent with extreme versions of incremental production, and instead support V. Ferreira and Dell (2000), who argued that the lexically driven picture of production—in which the most accessible lexical item wins a figurative “race” out of the mouth—might not be sufficient to accurately describe their results. Instead, they proposed that speakers choose a syntactic structure without necessarily first deciding between alternative lexical items. The structure, then, is what is really primed by pictures, sentences, and questions. On this view, incrementality applies to the filling of available NP nodes in the primed structure (F. Ferreira, 2000). Incrementality encourage the selection of a syntactic structure that allows accessible material to be mentioned sooner (V. Ferreira & Dell refer to this as ‘lexical-syntactic interactionism’).

Another way to think about incrementality in production is to ask what sorts of planning units the system uses. The most extreme versions of incrementality assume that there is little or no look-ahead, and so predict that planning units will be essentially non-existent (i.e., utterances are planned more or less word-by-word). On the other hand, non-incremental views assume that the system does engage in look-ahead over some multi-word domain. Clauses have been classically assumed to serve as planning units for grammatical encoding (Boomer, 1965; Ford & Holmes, 1978; Garrett, 1975; Lashley, 1951).

The idea is that the system organizes an entire clause (i.e., a verb and its arguments) before engaging in any phonological encoding.

One way that this issue has been addressed is by examining how speakers compute grammatical agreement between complex subjects and inflected verbs. A variety of studies have demonstrated that a distractor noun in the subject can be an attractive lure for agreement, especially if it is plural (Bock & Eberhard, 1993; Bock & Miller, 1991). Bock and Cutting (1992) used this phenomenon to determine whether the unit of grammatical encoding is the clause. They varied whether the constituent that intervened between the head noun of the subject and the main verb was a prepositional phrase modifier or a relative clause (e.g., *the editor of the history books* vs. *the editor who rejected the books*). They reasoned that if clauses are planning units for grammatical encoding, then agreement errors (plural inflections on a form of *to be*) should be less common in the relative clause condition. This prediction follows because the relative clause would be planned separately, and thus the head noun and the verb would be more closely linked during processing. This prediction was confirmed: Agreement errors were more common when a relative clause came between the head noun of the subject and the main verb, consistent with the classic idea that the unit of syntactic planning is the clause. This finding is inconsistent with radical incrementality or any type of production system which generates utterances on a word-by-word basis, but it can be reconciled with more limited incrementality (Christianson & F. Ferreira, in press; F. Ferreira, 2000).

Finally, there is evidence that the production system operates more efficiently when it has syntactic options that allow potentially different states of activation to be taken into account during grammatical encoding. V. Ferreira (1996) compared the production of sentences headed by a verb such as *give*, which alternates between a double-object and a prepositional dative form, and verbs such as *donate*, which only allow the prepositional dative (e.g., **The widow donated the library her entire collection*). He found that sentences

with syntactically flexible verbs such as *give* were generated more quickly and more fluently than sentences with more restrictive verbs. He argued that flexibility allows the system to accommodate lemmas' potentially different states of activation over time. For example, if a speaker has said *The widow gave* and then finds that direct object hard to retrieve, he or she can continue processing by working on the indirect object instead, because the verb *give* permits this flexibility. Thus, one benefit of syntactic freedom of choice is that it enhances the efficiency of language production. In the next section, we focus specifically on the issue of how speakers make syntactic choices.

IV. Syntactic choice

A. Choice of Syntactic Construction

As already mentioned, work by Carroll (1958), Bock (1986a, b), Bock and Warren (1985), and others has shown that, in English, when a noun phrase is made accessible by showing someone a picture of a semantically related item, asking a focusing question, or establishing a context, speakers tend to begin their sentences with that primed NP. Bock and Warren's (1985) work on the production of passives and dative structures in English indicates that the most accessible entity claims not only an early position in the string, but also the most prominent syntactic function (i.e., subject, or non-oblique dative in ditransitive structures). A similar finding is that passives tend to occur with theme-experiencer verbs, because the passive allows the experiencer to be placed in subject position (F. Ferreira, 1994). This effect is larger when the experiencer is human and the theme is not, indicating that an animacy contrast perhaps helped to distinguish the conceptual prominence of the two entities even more than just their thematic role status. Spanish speakers also are sensitive to accessibility when they choose syntactic constructions (Prat-Sala & Branigan,

2000). Spanish syntax includes a dislocated active structure (OSV, along the lines of ‘Cheese I love to eat’), which allows the effects of inherent accessibility (animacy) to be distinguished from those of derived accessibility (discourse prominence). Spanish speakers tended to place the more salient entities in higher syntactic positions, making use of both passives and the dislocated active structure. In general, then, syntactic forms are chosen to allow speakers to line up conceptual and syntactic prominence.

B. Inclusion of Optional Functional Elements

Thus far, we have considered how speakers decide on a syntactic form – active versus passive, double-object versus prepositional dative, and so on. Now we ask a slightly different question: How do speakers decide whether to include an optional function word such as the complementizer *that* in a sentence like *The weary traveler claimed (that) his luggage had been stolen*? If the complementizer is omitted, an ambiguity about the status of the noun phrase *the luggage* is created for the comprehender: *the luggage* could be either the direct object of *claimed* or the subject of a complement clause. The presence of the complementizer essentially disambiguates the structure, making it clear that the noun phrase is a subject. (It is possible for *that* to be a determiner, as in *that luggage*, not *your luggage*, but Roland, Elman, and V. Ferreira (in press) have demonstrated that post-verbally, the word *that* is almost always a complementizer, and the parser is likely tuned to this distributional information.) If speakers attempt to produce utterances that are easy for their listeners to understand, one might predict that, the greater the chance of a misinterpretation, the greater the likelihood that speakers will include the complementizer. For example, if the verb preceding the ambiguous noun phrase subcategorized for only clausal complements, the *that* is unnecessary, and so it might be omitted; but if the verb

takes both direct objects and clausal complements, the *that* would help the speaker avoid making a parsing error.

The evidence suggests that speakers' needs motivate complementizer inclusion. This has been shown in a variety of experiments by V. Ferreira (2003; V. Ferreira & Dell, 2000), which demonstrate that *that* is more likely to be included in complement and relative clause structures when the speaker is having difficulty retrieving the word that would follow *that*. Two different mechanisms can be proposed to account for this relationship: Alleviation and Signaling (Jaeger, 2005). According to the Alleviation hypothesis, speakers include *that* to give themselves time to plan (Race & MacDonald, 2003), making the complementizer essentially like a filler term such as *uh*. The alternative hypothesis, Signaling, assumes that the complementizer is a signal or at least a symptom of upcoming difficulty. The two hypotheses make opposite predictions about the distribution of complementizers and filler disfluencies. If Alleviation is right, then the presence of a complementizer should reduce the likelihood of a filler. If Signaling is correct, then *that*s and fillers should be positively correlated. Jaeger (2005) and V. Ferreira and Firato (2002) found results consistent with the second pattern, which supports the Signaling hypothesis. It is important to note, however, that the data are compatible with the idea that the inclusion of a complementizer is merely a symptom of difficulty – that is, that the same factors that lead to disfluencies may lead to complementizer inclusion as well.

This pattern has emerged in other studies as well - speakers in dialogue tasks fail to make use of either optional words or disambiguating prosody to avoid ambiguity (Allbritton, McKoon, & Ratcliff, 1996; Arnold, Wasow, Asudeh, & Alrenga, 2004; Kraljic & Brennan, 2005). One exception is a recent study reported by Haywood, Pickering, and Branigan (2005), who found that speakers did provide more disambiguating *that's* when they were describing objects to a conversational partner. However, as V. Ferreira, Slevc, and Rogers (2005) argue, the effect may be due to the visual properties of the situation the interlocutors

were presented with. The situation which led to ambiguity in the Haywood et al. study was one in which there was more than one object of the same type, thus inviting the use of a disambiguating modifier (e.g., *the penguin THAT'S in the cup on the star*), and it is in these situations that the word *that* tended to be included. Ferreira et al. argue that perhaps speakers were simply sensitive to the existence of more than one token of the same type and in those cases produced more explicit utterances. At the same time, it must be acknowledged that even if this interpretation is correct, it still appears that Haywood et al. have indeed observed the altruistic rather than egocentric use of optional functional elements. Moreover, as highly skilled speakers are probably better able to avoid ambiguity than those who are less practiced, it is clear that some mechanisms must exist to allow speakers to monitor their speech and include optional elements in just those situations when they might be helpful to listeners. What we do not know is how this process, which is potentially quite resource-demanding, is coordinated with the other tasks performed by the production system.

V. Syntax and Prosody in Production

Utterances have regular rhythmic and tonal properties, and these prosodic features of a sentence are influenced directly and indirectly by syntactic structure. Rhythm results from the way syllable stress and duration change over the course of an utterance, and intonation is the result of changes in fundamental frequency (F0) or tone (see F. Ferreira, 1993, for a description of what constitutes prosody). Consider the sentence *If she goes, I go too*. It would likely be spoken as two intonational phrases separated at the boundary indicated with a comma. The word *goes* would tend to have a long duration and would receive greater stress than it would in a non-clause-final position, and the same effects

would be observed on *too*, perhaps even more markedly, because *too* ends not only a clause but an entire sentence.

It is clear that utterance stress and timing have something to do with a sentence's syntactic structure. Phonologists have debated whether the correct characterization of these effects appeals directly to syntactic constituents, or instead makes reference to prosodic entities such as phonological words, phrases, and intonational phrases. According to the syntactic view (Cooper & Paccia-Cooper, 1980; Odden, 1990; Selkirk, 1984; Wagner, 2005), the amount of lengthening and stress assigned to a given word can be directly related to syntax. For example, the more syntactic right brackets that terminate on a word, the longer and more stressed it will tend to be. On the prosodic constituency view, syntax is used to create prosodic constituents, but then it is features of prosodic constituency that determine timing and stress (Gee & Grosjean, 1983; Inkelas & Zec, 1990; Levelt, 1989; Selkirk, 1986). Disentangling these two approaches to rhythm can be challenging, because prosodic and syntactic constituency are highly correlated (F. Ferreira, 1993), but one important theoretical difference between them is that prosodic structure is generally viewed as flatter and less articulated than syntactic structure, because prosodic constituency is generally thought not to permit recursion (Selkirk, 1986; cf. Gee & Grosjean, 1983; Ladd, 1986; Wagner, 2005). The idea is that, in syntax, a clause may have another clause inside of it (for example), but in prosody, such self-embedding is forbidden. As a result, prosodic structures are flatter than syntactic ones, allowing prosody to serve as an interface between hierarchical and recursive syntactic/semantic representations and the sequential speech channel through which articulation must take place. Another important difference between the two types of structures is that prosodic representations pay attention to the distinction between function and content words. Therefore, a phrase consisting of just a pronoun, for instance, would typically not behave the same way as a full lexical NP.

Intonation is related to syntax too, but again, it has long been known that the intonational phrasing of a sentence may not be isomorphic to its syntactic constituency. One famous example is *This is the cat that chased the rat that swallowed the cheese*…, which tends to be phrased as *(this is the cat) (that chased the rat) (that swallowed the cheese*), even though the major syntactic boundary is between *is* and *the cat*. Other more realistic examples include (*Mary left)(after the party)* and *(Mary gave the book) (to her brother who lives in Ohio*). In both these cases, the major intonational boundary comes not between the subject and verb phrase, but after the verb. To account for these cases, Selkirk (1984) proposed the Sense Unit Condition, which states that the constituents making up a single intonational phrase must be in either a head-argument or head-modifier relation. The Sense Unit Condition rules out apparently malformed examples such as *(Ten mathematicians) (in ten derive a lemma*) (Steedman, 2000), because *in ten* is a modifier of *mathematicians* and is not a head, argument, or modifier of *derive a lemma*. Steedman (2000) argues that a grammar such as Combinatory Categorial Grammar which allows a wider range of syntactic constituents than other approaches captures these sorts of facts and eliminates the need for a separate and stipulative Sense Unit Condition. The important point for our purposes, however, is that even though the intonational phrasing of a sentence might ultimately deviate from its syntactic structure, the well-formedness of the intonational phrasing appeals to syntactic concepts such as head, modifier, and argument.

So far we have considered only aspects of prosody that can be directly related to linguistic structures, either prosodic or syntactic. But syntax affects the sound properties of a sentence in another way, which we will roughly characterize as having to do with performance effects. For example, hesitations and pauses due to planning difficulty tend to cluster at clause boundaries (Ford, 1982; Goldman-Eisler, 1968). In addition, it has been argued that the most probable location for pauses and prosodic breaks can be predicted from algorithms which assume that break points are jointly determined by the complexity of

material to the left and to the right of the boundary (Gee & Grosjean, 1983; Watson & Gibson, 2004). F. Ferreira (1991) demonstrated that the syntactic complexity of upcoming material affected pause duration, and she argued that the effects were due to the difficulty of planning the upcoming material. Thus, the sound pattern of a sentence has at least two possibly distinct sources: One is the syntactic and prosodic representation which might mandate breaks in particular locations, and the other is the speaker's need for more time to plan upcoming material.

An important question for future research on the syntax-prosody interface in language production is whether these two sources are indeed distinct, or whether prosody and performance phenomena can be reduced to the same underlying causes. Another critical issue is whether it is necessary to postulate a distinct level of prosodic constituency to account for phenomena related to rhythm and intonational phrasing, or whether syntactic structure is sufficient to explain prosodic patterns in spoken sentences. One limitation of work that has been conducted up to this point is that virtually all studies investigating prosody in production have used simple reading or repetition tasks to elicit utterances. The reason is that in order to test prosodic and syntactic hypotheses adequately, it is necessary to precisely control what the speaker says. But unless more naturalistic tasks are used that allow speakers to talk relatively normally, it will be impossible to assess to what extent the need to plan affects the sound features of a sentence, and to evaluate how incrementality in production affects the distribution of hesitations, pauses, and even intonational boundaries across an utterance.

VI. Conclusions

What is currently known about the process of grammatical encoding indicates that the syntactic structures used in language production have the following characteristics. First,

they are generated in two separate stages, the first one creating a representation that represents hierarchical relations but not necessarily linear order, and a second stage in which linearization within phrases takes place. Second, the structures for creating both the global form of the entire utterance and the form of the individual phrases are generated from trees anchored to specific lexical heads. Third, there is some evidence that syntactic representations contain gaps or traces. Admittedly this is a point on which there little consensus and almost no data, but recent evidence about the computation of subject-verb agreement (Franck et al., in press) as well as data concerning the blocking of function word reduction following a gap suggests that gaps are in fact mentally represented at some stage in production. Moreover, all of these features of grammatical encoding can be captured using TAG as the representational format for syntactic information (F. Ferreira, 2000), which again assumes lexical generation of structure. The main verb of an utterance provides an overall clausal template constrained by the verb's phrase-taking properties, and then each specific phrase is fleshed out and attached as its head (e.g., a noun for a noun phrase) is accessed. TAG represents gaps not via movement but as part of the treelet anchored to the lexical item, thus explaining phenomena such as the blocking of function word reduction in the hypothesized vicinity of a gap.

In addition, although the concept of processing resources is somewhat vague (as MacDonald & Christiansen, 2002 argue), we can resort to an operational definition and say that processing resources are what measures such as initiation time and pause probability / duration reflect. With this assumption, we conclude that grammatical encoding requires resources, and that some constructions appear to be difficult to generate even in felicitous contexts. In addition, the bulk of the evidence indicates that production is incremental in the sense that the most accessible concept will tend to capture the syntactically most prominent position in a functional level structure. This tendency towards moderate incrementality reduces the computational burden on the grammatical encoder because the system can

begin with what is already accessible and wait for other elements to become available as processing unfolds. In addition, if it indeed turns out that syntactic priming is particularly robust in dialogue because it makes the task of generating a syntactic structure easier (Pickering & Garrod, 2004; Pickering & Branigan, 1999), then we have further evidence that grammatical encoding requires significant processing resources.

Recall that the original argument against this idea was that syntactic processing was assumed to be modular (Levelt, 1989), and one of the characteristics of a module is that it operates automatically (J. A. Fodor, 1983). Do our conclusions undermine this assumption? Not necessarily. Fodor's conception of automaticity appears to have more to do with whether a person's conscious goals and intentions can influence processing than with whether the process is computationally costly. Moreover, it is clear that many specialized systems call upon working memory resources, and one point of debate has been whether the working memory that is involved is domain-general or entirely devoted to just that module. Thus, a system might be modular but still draw on working memory, and the resource pool that is used could itself be modular, in the sense that it is dedicated to processing in that one domain. These are topics for further investigation.

We also conclude that syntactic choices are made largely for the benefit of the speaker. The decision about what syntactic construction to use is at least in part based on the accessibility of the lemmas that will comprise the utterances. Optional function words such as complementizers are left out when the speaker has time to retrieve the immediately following word but included when retrieval is slow and difficult. There is also evidence that speakers obey the Gricean Maxim of Quantity only to a limited extent, in part because they have a tendency to describe objects in whatever way is salient to them, neglecting to take into account the effect the description might have on the comprehender (Engelhardt, Bailey, & Ferreira, in press). For example, even though the relevant discourse might include just a single hat, studies show that more than 25% of the time, the hat will be described as *the red*

hat or *the hat with the feather*. These over-descriptions likely occur because, from the speaker's point of view, the object IS a red hat or a hat with a feather. Because those features of the object are salient, they have a good chance of making it into the message-level representation. In these situations, it would require extra effort for the speaker to produce concise descriptions, because he or she would have to remove content to make sure that information did not get grammatically encoded.

Finally, the syntactic structure of a sentence affects the way it is spoken. For example, the presence of a gap in surface structure affects whether a preceding word is reduced or lengthened. More generally, syntax has profound effects on all aspects of prosody, including the duration and stress level of words, the location and duration of pauses, and the intonational tune and phrasing of the sentence. An unresolved question is whether syntax is directly responsible for these effects, or whether they are mediated through prosodic constituency. Another is how linguistic structure and performance limitations play off of each other to help establish a sentence's overall prosodic form. In addition, it is still not clear how lemma retrieval, word-form activation, and functional and positional level processing are coordinated with the tasks of creating prosodic constituents, generating intonational contours, and implementing a phonetic plan (F. Ferreira, 1993). Moreover, essentially the same questions can be asked about prosody that we considered with respect to syntax in the present review: What sorts of computational resources does the process of creating prosodic representation draw upon, and how do speakers manage and even take advantage of optionality in prosody (for some discussion, see Steedman, 2000; Watson & Gibson, 2004). Unfortunately, although we can ask these questions, there is still not enough evidence to allow us to provide even tentative answers. Ultimately, a complete understanding of syntax in production will require consideration of issues relating to prosody. We therefore hope that the next decade will see an integration of research on syntax, prosody, and language production.

References

Allbritton, D. W., McKoon, G., & Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, & Cognition , 22,* 714-735.

Arnold, J. E., Wasow, T., Asudeh, A., & Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory and Language, 51,* 55-70.

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

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

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

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

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language, 31,* 99–127.

Bock, J. K., & Eberhard, K. M. (1993). Meaning, sound, and syntax in English number agreement. *Language and Cognitive Processes, 8,* 57–99.

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

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

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

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23,* 45-93.

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

Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech, 8,* 148-158.

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

Bresnan, J. & Kaplan, R. (1985). Introduction: Grammars as mental representations of language. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 17-52). Cambridge, MA: MIT Press.

Carroll, J. B. (1958). Communication theory, linguistics, and psycholinguistics. *Review of Educational Research, 28,* 79–88.

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

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1995). *The minimalist program.* Cambridge, MA: MIT Press.

Christianson, K., & Ferreira, F. (in press). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition*, 1-31.

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

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech.* Cambridge, MA: Harvard University Press.

Cooper, W. E., & Ross, J. R. (1975). World Order. In R. E. Grossman, L. J. San, and T. J.

Vance (Eds.), *Papers from the parasession on functionalism* (pp. 63-111). Chicago Linguistic Society.

de Smedt, K. (1996). Computational models of incremental grammatical encoding. In A. Dijkstra & K. de Smedt (Eds.) (1996). *Computational psycholinguistics: AI and connectionist models of human language processing* (pp. 279-307). London: Taylor & Francis.

Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (in press). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language,*

Ferreira, F. (1988). Planning and timing in sentence production: The syntax-to-phonology conversion. Unpublished Doctoral Dissertation, University of Massachusetts, Amherst, MA.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language, 30,* 210-233.

Ferreira, F. (1993). The creation of prosody during sentence processing. *Psychological Review, 100,* 233-253.

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

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. Wheeldon (Ed.), *Aspects of language production* (pp. 291–330). London: Psychology Press.

Ferreira, F. (2003). The misinterpretation of noncanonical sentence. *Cognitive Psychology, 47,* 164-203.

Ferreira, F., & Henderson, J. M. (1998). Linearization strategies during language production. *Memory & Cognition, 26,* 88-96.

Ferreira, F., Lau, E. F., & Bailey, K. G. D. (2004). Disfluencies, language comprehension, and tree adjoining grammars. *Cognitive Science, 28,* 721-749.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arthimetic sums. *Journal of Memory and Language, 46,* 57-84.

Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35,* 724–755.

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

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology, 40,* 296–340.

Ferreira, V.S., & Firato, C.E. (2002). Proactive interference effects on sentence production. *Psychonomic Bulletin & Review*, *9*, 795-800.

Ferreira, V. S., Slevc, L. R., & Rogers, E. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition, 96,* 263-284.

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

Fodor, J.D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry*, *11*, 759-770.

Fodor, J.D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes*, *4*, 155-209.

Fodor, J. D. (1991). Sentence Processing and the Mental Grammar. In P. Sells, S. Sheiber, and T. Wasow, (Eds)., *Foundational Issues in Natural Language Processing* (pp. 83-114). Cambridge, MA: MIT Press.

Ford, M. (1982). Sentence planning units: Implications for the speaker’s representation of meaningful relations underlying sentences. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 798-827). Cambridge, MA: MIT Press.

Ford, M., & Holmes, V. M. (1978). Planning units and syntax in sentence production. *Cognition, 6,* 35-53.

Franck, J., Lassi, G., Frauenfelder, U. H., & Rizzi, L. (in press). Agreement and movement: A syntactic analysis of attraction. *Cognition,* 1-79.

Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject–verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes, 17,* 371–404.

Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 9, pp. 133–177). New York: Academic Press.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 177–220). London: Academic Press.

Gazdar, G., Klein, E., Pullum, G. K., & Sag, I. A. (1985). *Generalized phrase structure grammar.* Cambridge, MA: Harvard University Press.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology, 15, 411-458.*

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech.* New York: Academic Press.

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language*, *35*, 623-647.

Grice, P. (1975). Logic and Conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics: Speech acts* (Vol. III, pp. 41-58). New York: Academic Press.

Grimshaw, J. (1990). *Argument structure.* Cambridge, MA: MIT Press.

Hartsuiker, R.J., Corley, M., & Martensen, H. (2005). The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Related beply to Baars et al. (1975). *Journal of Memory and Language*, *52*, 58-70.

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

Hawkins, J. (1994). *A performance theory of order and constituency.* Cambridge: Cambridge University Press.

Haywood, S. L., Pickering, M. J., & Brennan, S. E. (2005). Do speakers avoid ambiguities during dialogue? *Psychological Science, 16,* 362-366.

Inkelas, S., & Zec, D. (1990). *The phonology-syntax connection.* Chicago: University of Chicago Press.

Jackendoff, R. (1977). *X-bar Syntax: A Study of Phrase Structure.* Cambridge MA: MIT Press.

Jaeger, T.F. (2005). Optional *that* indicates production difficulty: Evidence from disfluencies. *Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop. 10–12 September 2005, Aix-en-Provence, France*, 103-109.

Johnson, N. F. (1966). On the relationship between sentence structure and the latency in generating the sentence. *Journal of Verbal Learning and Verbal Behavior, 5,* 375-380.

Joshi, A. K. (1985). How much context-sensitivity is required to provide reasonable structural descriptions: Tree adjoining grammars. In D. Dowty, L. Kartunnen, and A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives*. (pp. 206–250). Cambridge: Cambridge University Press.

Joshi, A. K., Levy, L., & Takahashi, M. (1975). Tree adjunct grammars. *Journal of the Computer and System Sciences, 10,* 136–163.

Kempen, G., & Hoenkamp, E. (1982). Incremental sentence generation: Implications for the structure of a syntactic processor. In J. Horecky (Ed.), *Proceedings of the Ninth International Conference on Computational Linguistics.* Amsterdam: North-Holland.

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

Kraljic, T., & Brennan S. E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee. *Cognitive Psychology, 50,* 194-231.

Ladd, D. R. (1986). Intonational phrasing: The case for recursive prosodic structure. *Phonology Yearbook, 3,* 311-340.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-136). New York: John Wiley & Sons.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

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

MacDonald, M.C., & Christiansen, M.H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*, 35-54.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical Nature of Syntactic Ambiguity Resolution. *Psychological Review, 101,* 676-703.

MacWhinney, B., Bates, E., & Reinhold, K. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior, 23,* 127-150.

Mitchell, D. C., & Cuetos, F. (1991). The origins of parsing strategies. In C. Smith (Ed.), *Current issues in natural language processing* (pp. 1-12). Austin TX: University of Texas, Center for Cognitive Science.

Moortgat, M. (1989). *Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus.* Foris, Dordrecht.

Odden, D. (1990). Syntax, lexical rules and postlexical rules in Kimatummbi. In S. Inkelas & D. Zec (Eds.), *The phonology-syntax connection* (pp. 259-278). Chicago: University of Chicago Press.

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

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

Pickering, M. J., Branigan, H. P., & McLean, J. F. (2002). Constituent structure is formulated in one stage. *Journal of Memory and Language, 46,* 586–605.

Pickering, M.J., & Garrod, S. (2004). Toward a mechanistic theory of dialogue. *Behavioral & Brain Sciences*, *27*, 169-226.

Pinker, S., & Birdsong, D. (1979). Speaker's sensitivity to rules of frozen word order. *Journal of Verbal Learning and Verbal Behavior, 18,* 497-508.

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

Race, D. S., & MacDonald, M. C. (2003). The use of "that" in the production and comprehension of object relative clauses. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*.

Rochester, S. R., & Gill, J. (1973). Production of complex sentences in monologues and dialogues. *Journal of Verbal Learning and Verbal Behavior, 12,* 203-210.

Roland, D., Elman, J. L., & Ferreira, V. S. (in press). Why is *that*? Structural prediction and ambiguity resolution in a very large corpus of English sentences. *Cognition,* 1-28.

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

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure.*

Cambridge, MA: MIT Press.

Shapiro, L.P., Hestvik, A., Lesan, L., & Garcia, A. R. (2003). Charting the time-course of VP-ellipsis sentence comprehension: Evidence for an initial and independent structural analysis. *Journal of Memory and Language*, *49*, 1-19.

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

Steedman, M. (2000). *The Syntactic Process*. Cambridge, MA: MIT Press.

Stowell, T. (1981). *Origins of phrase structure.* Unpublished doctoral dissertation, MIT.

Tanenhaus, M. K., Flanigan, H. P., & Seidenberg, M. S. (1980). Orthographic and phonological activation in auditory and visual word recognition. *Memory & Cognition*, *8*, 513-520.

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

Tomlin, R. S. (1983). On the interaction of syntactic subject, thematic information, and agent in English. *Journal of Pragmatics, 7,* 411-432.

van Nice, K. Y., & Dietrich, R. (2003). Task sensitivity of animacy effects: Evidence from German picture description. *Linguistics, 41,* 825-849.

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

Wagner, M. (2005). *Long distance effects on prosody.* Poster presented at the 18th Annual Meeting of the CUNY conference on Human Sentence Processing. Tucson, Arizona.

Wasow, T. (1997). Remarks on grammatical weight. *Language Variation and Change*, *9,* 81–105.

Watson, D., & Gibson, E. (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes*, *19*, 713-755.

Yngve, V. H. (1960). A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society, 104,* 444-466.

Notes

[1]The Fromkin model of speech errors, which is a forerunner of BL and of Garrett's work, included more than two stages of syntactic processing.

[2]Here the models deviate from generative approaches to syntax in assuming that the notions 'subject' and 'object' are syntactic primitives, whereas since the earliest days of generative grammar, grammatical relations have been computed from the geometry of a phrase structure representation. For example, the subject is the Noun Phrase (or Determiner Phrase) immediately dominated by the sentence node. Other approaches to formal syntax such as Lexical-Functional Grammar and Relational Grammar do treat notions such as subject and object as primitives.