



The Cross-linguistic Study of Sentence Production

T. Florian Jaeger^{1*} and Elisabeth J. Norcliffe²

¹*Brain and Cognitive Sciences, Computer Science, University of Rochester; and*

²*Linguistics, Stanford University*

Abstract

The mechanisms underlying language production are often assumed to be universal, and hence not contingent on a speaker's language. This assumption is problematic for at least two reasons. Given the typological diversity of the world's languages, only a small subset of languages has actually been studied psycholinguistically. And, in some cases, these investigations have returned results that at least superficially raise doubt about the assumption of universal production mechanisms. The goal of this paper is to illustrate the need for more psycholinguistic work on a typologically more diverse set of languages. We summarize cross-linguistic work on sentence production (specifically: grammatical encoding), focusing on examples where such work has improved our theoretical understanding beyond what studies on English alone could have achieved. But cross-linguistic research has much to offer beyond the testing of *existing* hypotheses: it can guide the development of theories by revealing the full extent of the human ability to produce language structures. We discuss the potential for interdisciplinary collaborations, and close with a remark on the impact of language endangerment on psycholinguistic research on understudied languages.

1 Introduction

In order to communicate, speakers need to encode the messages they intend to convey into an acoustic signal. Speakers need to select the words necessary to convey the intended message, determine the functional dependencies between them, arrange them in an acceptable order, and retrieve the phonological information necessary to initiate articulation. Despite the obvious complexity of this task, speakers usually master it in real time, while (more or less) obeying language-specific grammatical constraints, and while maintaining a high degree of fluency. Psycholinguistic research on sentence production seeks to understand the linguistic processes, representations, and interfaces involved in the encoding of pre-linguistic messages into linguistic form and, ultimately, into articulation. In this paper, we describe how evidence from the cross-linguistic study of language production (work on different languages) – albeit still rare – has advanced our theoretical understanding

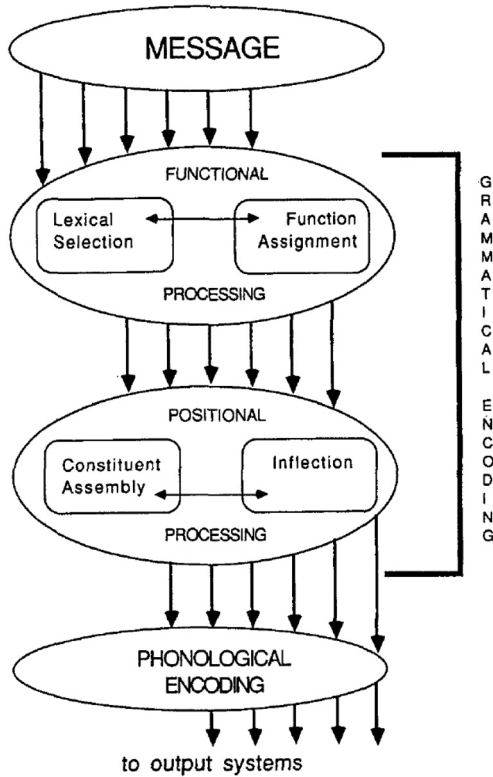


Fig. 1. Model of sentence production (reprinted from Bock and Levelt 1994).

of this area of psycholinguistics. We discuss some of the challenges of cross-linguistic work and suggest future directions for this research program.

We limit ourselves to what is commonly referred to as *grammatical encoding*, the first ‘steps’ in the process from thought to articulation. Grammatical encoding is the process of selecting the lexical entries and syntactic frames necessary to encode the intended pre-linguistic message. This process is commonly assumed to involve two stages (Garrett 1980; Kempen and Hoenkamp 1989; Levelt 1989; Bock and Levelt 1994; but see Kempen and Harbusch 2004; Branigan, Pickering, and Tanaka 2008), *functional processing* and *positional processing*, as illustrated in Figure 1 (reprinted from Bock and Levelt 1994).

During functional processing, speakers are assumed to select lexical entries, retrieve their syntacto-semantic information, and assign grammatical functions (i.e. determine which referent bears the subject function, etc.). During positional processing, the retrieved lexical items are then inflected (e.g. subject-verb agreement), assembled into constituents, and linearized.

Most psycholinguistic research assumes that the architecture of the language production system is universal and hence does not depend on

speakers' native languages (e.g. Bock et al. 2001; Bock et al. 2006, but see discussions in Bates and MacWhinney 1982; Bates and Devescovi 1989; MacWhinney and Bates 1989; and references below). Indeed, at least at a very general level, accounts originally based on evidence mostly from English, Dutch, and German (Garrett 1980; Kempen and Hoenkamp 1989; Levelt 1989; Bock and Levelt 1994) seem to be compatible with evidence from other languages. Several production phenomena, including accessibility-sensitive and complexity-sensitive sentence production, syntactic persistence (discussed below), have been attested in all languages investigated so far.

However, the set of languages that have been studied psycholinguistically is small. Out of more than 5000 languages spoken across the world, less than 30 have been subjected to controlled psycholinguistic investigation on sentence production, most of which resemble each other typologically. Even for this typologically rather homogeneous sample, there is evidence that language-specific properties affect language processing and that not all aspects of language production are fully universal (cf. evidence from agreement processing, discussed below). But cross-linguistic work can do more than just check whether our fundamental assumptions about language production are justified. We give examples from work on grammatical encoding where cross-linguistic evidence has distinguished between competing psycholinguistic accounts.

We believe that the value of cross-linguistic work goes even further, beyond the direct comparison of *existing* psycholinguistic theories. Cross-linguistic work on sentence production and, more generally, psycholinguistics drives progress by broadening the empirical base in need of explanation. In addition, we argue, cross-linguistic work provides a great opportunity (and need) for interdisciplinary collaborations between linguists and cognitive scientists.

2 Current Findings in Cross-linguistic Research on Sentence Production

We summarize three widely studied research themes on grammatical encoding, focusing in particular on the impact that cross-linguistic work has had on theoretical development. First, we summarize research on how the relative accessibility of referents involved in a message affects sentence production and what this reveals about the extent to which sentence production is incremental. Second, we summarize research on agreement (e.g. subject–verb agreement), which aims to answer to whether conceptual information, which is generally considered pre-linguistic, is available during grammatical encoding. Finally, we summarize research on how the complexity of constituents affects constituent order. We conclude this section by briefly mentioning other areas of cross-linguistic research on sentence production.

2.1 ACCESSIBILITY AND GRAMMATICAL ENCODING

Much cross-linguistic work on sentence production has focused on so-called *conceptual accessibility* effects (Bock and Warren 1985) and implications

about the time course of incremental sentence production. In its most general sense, conceptual accessibility refers to the ease of retrieval of a referent and the corresponding expression from memory. Conceptual accessibility has been linked to both inherent properties of referents (accessibility increases with *imageability*, Bock and Warren 1985; *prototypicality*, Kelly, Bock, and Keil 1986; Onishi, Murphy, and Bock 2008; *animacy/humanness*, Byrne and Davidson 1985; Bock, Loebell, and Morey 1992; McDonald, Bock, and Kelly 1993; F. Ferreira 1994; Prat-Sala and Branigan 2000; van Nice and Dietrich 2003; Kempen and Harbusch 2004; Rosenbach 2005; Bresnan et al. 2007; Dennison 2008) and contextually conditioned properties (accessibility increases with *previous mention* of the same referent, MacWhinney and Bates 1978; Bock and Irwin 1980; V. S. Ferreira and Yoshita 2003; Prat-Sala and Branigan 2000; Bresnan et al. 2007; *semantic similarity* to recently mentioned words, Bock 1986b; Igoa 1996; *visual salience*, Gleitman et al. 2007; Myachykov 2007; Myachykov, Garrod, and Scheepers, forthcoming; Myachykov, Posner, and Tomlin 2007). Production researchers have been interested in conceptual accessibility, because it affects speakers' word order choices, for example, in active vs. passive voice (*Lightning struck the church* vs. *The church was struck by lightning*) or the ditransitive alternation (*John handed a book to Mary* vs. *John handed Mary a book*). Early findings suggested that more accessible material is generally produced earlier in such environments (Bock and Warren 1985), which has been taken to suggest what Ferreira and Dell (2000: 289) call the Principle of Immediate Mention: 'Production proceeds more efficiently if syntactic structures are used that permit quickly selected lemmas to be mentioned as soon as possible'. Production is assumed to be radically incremental, greedily proceeding with whatever material is available first, wherever speakers have the choice to do so (i.e. where grammar permits it). In this view, conceptual accessibility affects word order *directly* (i.e. during positional processing; see also Branigan and Feleki 1999; Prat-Sala and Branigan 2000; V. S. Ferreira and Yoshita 2003; Kempen and Harbusch 2004). According to an alternative proposal, conceptual accessibility affects word order *indirectly* by affecting grammatical function assignment to referents (during functional processing, Bock and Warren 1985; F. Ferreira 1994). According to indirect accounts of conceptual accessibility effects, a passive becomes more likely when the patient (the entity being acted on) is more accessible than the agent (e.g. *The boy was struck by lightning* vs. *The church was struck by lightning*) because speakers prefer to assign the subject function to the most accessible referent (Bock and Warren 1985). Unfortunately, direct and indirect accounts are hard to distinguish based on English data alone, because 'structures separating grammatical function from string prominence are rare in English' (Christianson and Ferreira 2005: 109; see also Branigan, Pickering, and Tanaka 2008; Prat-Sala and Branigan 2000). For example, English passivization not only assigns the subject function to the patient, the patient is also mentioned earlier than in the active.

Prior to cross-linguistic work on conceptual accessibility, the distinction between the two accounts rested on one piece of data: the fact that animacy affects word order variations involving different mappings between referents and grammatical function (such as the voice and ditransitive alternations), while apparently not affecting word order in noun phrase (NP) coordination, where the two word orders do not correspond to difference in grammatical function assignment (*The lost hiker fought time and winter* vs. *The lost hiker fought winter and time*; Bock and Warren 1985; McDonald, Bock, and Kelly 1993; Osgood and Bock 1977). Other studies have found conflicting evidence (Kelly, Bock, and Keil 1986; Gleitman et al. 2007; Onishi, Murphy, and Bock 2008). Yet others have pointed out that NP coordination is a syntactically unusual structure for independent reasons (Branigan, Pickering, and Tanaka 2008), making it all the more desirable to have additional evidence distinguishing direct and indirect accessibility accounts.

This evidence has come from the cross-linguistic study of accessibility effects (already called for in Osgood and Bock 1977: 133), including work on Fijian (Byrne and Davidson 1985), German (van Nice and Dietrich 2003; Kempen and Harbusch 2004), Greek (Branigan and Feleki 1999), Hungarian and Italian (MacWhinney and Bates 1978), Japanese (V. S. Ferreira and Yoshita 2003; Tanaka, Branigan, and Pickering, in prep), Korean (Dennison 2008), Odawa (Christianson and Ferreira 2005), Russian (Myachykov 2007), Spanish (Prat-Sala and Branigan 2000), and other languages (Sridhar 1988 presents evidence from Cantonese, Hebrew, Finnish, Slovenian, English, Spanish, Hungarian, Japanese, Kannada, and Turkish; see also Chang, Lieven, and Tomasello 2008). For example, while grammatical function and word order are fairly highly correlated in English, languages with case-marking tend to have more flexible word order (presumably because they do not need to rely on word order as a cue to grammatical function):

- (1a) Sta dimokratika politevmata, ton politis seve te to sindagma (Greek)
 in democratic regimes the citizen_{ACC} respects the law_{NOM}
 'In democratic regimes, the law respects the citizen'.
- (1b) Sta dimokratika politevmata, to sindagma seve te ton politis
 in democratic regimes the law_{NOM} respects the citizen_{ACC}
 'In democratic regimes, the law respects the citizen'. (Branigan and Feleki 1999: 4–5)

Studies on such languages in which word order can vary more freely without changes in grammatical function assignment have found that animacy affects word order even after grammatical function assignment is controlled for (see Kempen and Harbusch 2004 for German; Branigan and Feleki 1999 for Greek; Tanaka, Branigan, and Pickering in prep, for Japanese). Similar effect have been observed for givenness (MacWhinney and Bates 1978 for Italian and Hungarian; V. S. Ferreira and Yoshita 2003 for Japanese). This provides evidence for a direct effect of accessibility on

word order. However, independent effects of accessibility on grammatical function assignment have also been observed.

Consider Japanese, a case-marking verb-final language, that allows both subject-object-verb (SOV) and object-subject-verb (OSV) sentences and does so in both the active and the passive voice (distinguished on the verb).

(2a) Active SOV sentence (Japanese)

booto-ga gonin no hito-o hakonda.
 boat-NOM five people-ACC carry-PAST
 'A boat carried five people'.

(2b) Active OSV sentence

gonin no hito-o booto-ga hakonda.
 five people-ACC boat-NOM carry-PAST
 'Five people, a boat carried'.

(2c) Passive SOV sentence

gonin no hito-ga booto-niyotte hakobareta.
 five people-NOM boat-OBL carry-PAS-PAST
 'Five people were carried by the boat'.

(2d) Passive OSV sentence

booto-niyotte gonin no hito-ga hakobareta.
 boat-OBL five people-NOM carry-PAS-PAST
 'By the boat, five people were carried'. (Tanaka, Branigan, and Pickering in prep)

These properties make it possible to study both indirect and direct accessibility effects on the same language. As expected by direct accounts, Japanese speakers prefer to realize animate referents earlier in the sentence, independent of whether they carry the subject or object function (Tanaka, Branigan, and Pickering in prep, Experiment 1). However, as expected by indirect accounts, Japanese speakers also prefer to assign the subject function to animate referents (independent of word order, Tanaka, Branigan, and Pickering in prep, Experiment 2; see also Christianson and Ferreira 2005 on Odawa).

In sum, cross-linguistic work provides strong evidence that conceptual accessibility affects word order both indirectly during functional processing and directly during positional processing, which has led some researchers to reject the assumption that these two 'stages' of grammatical encoding are independent (Kempen and Harbusch 2004; Branigan, Pickering, and Tanaka 2008; Tanaka, Branigan, and Pickering in prep) – a conclusion that would be hard or impossible to arrive at without cross-linguistic data on sentence production.

2.2 INCREMENTALITY

Cross-linguistic investigation has also contributed significantly to another area of research that is tightly linked to the question of whether accessibility

affects production directly or indirectly (e.g. Dutch and French: Brysbaert, Fias, and Noel 1998; German: van Nice and Dietrich 2003; Odawa: Christianson and Ferreira 2005; Spanish: Brown-Schmidt and Konopka 2008). The presence of direct accessibility effects has been taken to indicate that grammatical encoding is a strongly incremental process: presumably to maintain fluency, speakers continue with whatever material becomes available first whenever grammar permits it (cf. Ferreira and Dell 2000: 289; see also V. S. Ferreira 1996; Kempen and Harbusch 2003; Kempen and Hoenkamp 1989; Levelt 1989; Roelofs 1998; Wheeldon and Lahiri 1997; but see Bock 1986a: 359). But just how incremental is grammatical encoding (Christianson and Ferreira 2005; F. Ferreira and Swets 2002)?

Brown-Schmidt and Konopka (2008) take advantage of an asymmetry between Spanish and English to demonstrate that speakers initiate pronunciation of a noun phrase before all its parts are retrieved. While adjectives precede the noun they modify in English (*the small butterfly*), they follow the noun in Spanish (*la mariposa pequeña*). Brown-Schmidt and Konopka show that Spanish speakers retrieved the adjective later in Spanish than in English, consistent with the hypothesis that sentence production is highly incremental (see also Brysbaert, Fias, and Noel 1998 on Dutch and French; van Nice and Dietrich 2003 on German; but see Christianson and Ferreira 2005; F. Ferreira and Swets 2002). Results like these demonstrate the power of cross-linguistic psycholinguistic investigations, but further work is necessary to see whether the finding extends beyond the planning and production of noun phrases. Some studies on English suggest that speakers need to plan the subject and verb of a sentence before they initiate articulation, arguing for a special status of the sentence subject and predicate and against radical incrementality (Lindsay 1975; also Smith and Wheeldon 1999). It is an open question how this finding would extend to verb-final languages or languages where the subject usually follows the verb. For example, do Japanese speakers initiate articulation only after the sentence-final verb is planned? Some evidence that previous studies (Lindsay 1975; also Smith and Wheeldon 1999) have underestimated incrementality comes from studies showing that utterances can be initiated when as little as one word has been fully planned (Griffin 2003; Wheeldon and Lahiri 1997; also Brysbaert, Fias, and Noel 1998). One possible reason for the conflicting data may be that incrementality is to some extent under speakers' strategic control: under time pressure, speakers initiate articulation earlier (F. Ferreira and Swets 2002 on English; van Nice and Dietrich 2003 on German). This area will most likely benefit from further cross-linguistic work on languages with different base word orders than most German languages.

2.3 CONCEPTUAL VS. GRAMMATICAL INFLUENCES DURING AGREEMENT PROCESSING

Psycholinguists have studied what information is available at different processing stages. While, for example, modularist accounts assume restrictive

interfaces between the different stages of production such as conceptual processing, functional and positional processing, and phonological encoding (Garrett 1976, 2000; Levelt 1989), others take the view that information can spread more freely between levels (Vigliocco and Franck 2001; also cascading activation and connectionist accounts of lexical access and grammatical encoding, Caramazza 1997; Chang, Dell, and Bock 2006; Dell 1986; Dell, Chang, and Griffin 1999; Stemberger 1985). One phenomenon that has been studied extensively with regard to this question is *agreement* (e.g. subject–verb agreement, as in *The road is wet* vs. *The roads are wet*). Agreement provides a window into what information is available during positional processing (the hypothesized second stage of grammatical encoding, during which agreement/inflection takes place, see Figure 1). In particular, psycholinguists have investigated what role conceptual (a.k.a. notional) information plays in agreement processing, and whether the availability of this information might vary across languages (Bock, Carreiras, and Meseguer, forthcoming; Vigliocco and Franck 1999; Bock et al. 2001, 2001; Vigliocco and Hartsuiker 2002; Eberhard, Cutting, and Bock 2005; Lorimor et al. 2008).

Consider the word *cats*. At the conceptual level, speakers begin with a notion of the numerosity of the referent, in this case, a multitude of cats. At the grammatical level, this preverbal notion undergoes linguistic encoding as a word bearing grammatical number features. *Grammatical* number refers to the linguistic agreement properties of a word. For example, we know that *cats* is grammatically plural because words that agree with it, e.g. finite verbs, must themselves be plural. Typically, grammatical number and notional number converge, though not always. *Scissors* for example, is grammatically plural, but notionally singular, that is, it is conceived of as a single object (Bock et al. 2001). Conversely, *clothing* can denote a multitude, yet is grammatically singular. Mismatches between grammatical and notional number extend to noun phrases as well. For example, although the phrase *the picture on the postcards* is grammatically singular, it has a distributive interpretation (multiple postcards bearing the identical picture), which can favor notional plurality in a speaker's mental model (Eberhard 1999; R. J. Hartsuiker, Kolk, and Huinck 1999; Vigliocco, Butterworth, and Garrett 1996; Vigliocco, Butterworth, and Semenza 1995).

Experimental investigations of conflicts between notional and grammatical number provide evidence for the existence of two separate components of number information in agreement resolution. This has been observed most often for distributive construals of noun phrases. For example, Humphreys and Bock (2005) found that English speakers were more likely to use plural verbs after subject noun phrases like *the gang on the motorcycles*, which have a distributive construal, than after minimally different phrases like *the gang near the motorcycles*, which don't. Distributivity has been shown to affect subject–verb agreement in a number of languages, including English, Dutch, French, Spanish, and Russian (Eberhard 1999; Lorimor et al. 2008; Vigliocco, Butterworth, and Garrett 1996; Vigliocco et al. 1996).

This research has raised questions about the universality of the agreement mechanism: Is agreement processing in all languages affected by the same types of information (Bock et al. 2001, 2006; Lorimor et al. 2008) or do language-specific properties mediate processing (Vigliocco, Butterworth, and Garrett 1996; Vigliocco et al. 1996)? In particular, some studies have suggested that the morphological richness of a language affects to what extent conceptual information affects agreement processing. While some early studies have suggested that languages with rich verb–number morphology (e.g. Italian or French compared to English) exhibit stronger effects of notional number (Vigliocco, Butterworth, and Garrett 1996; Vigliocco et al. 1996), more recent evidence suggests, if anything, the opposite (Berg 1998 on German and English; Lorimor et al. 2008 on Russian; for a possible account, see also Bock, Carreiras, and Meseguer forthcoming). In either case, this research suggests that language-specific properties can interact with production mechanisms to lead to cross-linguistic differences with regard to information flow during sentence production (Vigliocco, Butterworth, and Garrett 1996; Vigliocco et al. 1996), a conclusion which receives independent support from cross-linguistic research on other aspects of noun phrase production (e.g. determiner selection, Schiller and Caramazza 2003; Schriefers 1993 on Dutch; Schiller and Caramazza 2003; Schriefers and Teruel 2000 on German; Miozzo and Caramazza 1999 on Italian; Costa, Sebastian-Galles, Miozzo, and Caramazza 1999 on Catalan and Spanish; for an overview, see Caramazza et al. 2001).

Another tempting possibility is that what appear to be language-specific differences actually point to deeper cross-linguistic generalizations. For example, there is evidence that within a language, too, reduction of morphological complexity (e.g. the elision of pronouns) correlates with increased availability of conceptual information during agreement processing (Foote 2006; Lorimor et al. 2008), suggesting ‘an unexpected and scientifically interesting consistency in language processes across different languages’ (Lorimor et al. 2008: 791). While it remains to be seen, whether this generalization holds for a typologically more varied set of languages than studied so far, cross-linguistic research has already contributed significantly to our understanding of agreement processing.

2.4 COMPLEXITY

It has been widely documented, and experimentally demonstrated, that English speakers exhibit a strong preference to order lighter constituents before heavier ones, grammar permitting. In the dative alternation, for example, in which speakers have a choice between the double-object construction (*I gave [Tom] [the book]*) and the prepositional dative construction (*I gave [the book] [to Tom]*), heavy constituents tend to be produced later (*I gave the book to the man riding the bicycle*; Hawkins 1994; Arnold et al. 2000; Bresnan et al. 2007; Wasow 2002).¹ Such tendencies have also been

observed in English for prepositional phrase (PP) ordering (*I saw him* [_{TEMP} *yesterday*] [_{LOC} *in the park*] vs. *I saw him* [_{LOC} *there*] [_{TEMP} *just a couple of days ago*], Hawkins 1999), verb particle shift (*Give* [_{OBJ} *me*] *up* vs. *Give up* [_{OBJ} *any hope*], Lohse, Hawkins, and Wasow 2004) and heavy NP shift (*Put* [_{OBJ} *the apple*] [_{LOC} *in the basket*] vs. *Put* [_{LOC} *in the basket*] [_{OBJ} *the apple that's lying on the napkin*], Arnold et al. 2000; Wasow 1997). Short-before-long preferences have also been widely documented for other (Germanic) languages, including Dutch (Haeseryn 1997) and German (Uszkoreit 1987), which has contributed to the assumption that it may be a universal feature of the production mechanism. One prominent account for the asymmetry is framed in terms of the Levelt (1989; see Bock and Levelt 1994) model of sentence production: the ordering of constituents is determined by the processing time required to actually produce them. Short constituents can be formulated faster, and hence are selected earlier for production (de Smedt 1994; Wasow 1997, 2002; cf. the Principle of Immediate Mention, Ferreira, and Dell 2000: 289, discussed above).

However, investigations of other languages have undermined the postulated universality of these preferences. For example, in Japanese (Dryer 1980; Hawkins 1994; Yamashita and Chang 2001) and Korean (Choi 2007) speakers seem to prefer to produce long phrases *before* short phrases (see also Matthews and Yeung 2001 for comprehension evidence for a similar preference in Cantonese Chinese).

(3a) Long-before-short

Kunye-ka kkoli-lul huntunun kay-eykey ppye-lul cwu-ess-ta. (Korean)
 she tail wagging dog-to bone-ACC gave
 'She gave a dog wagging his tail a bone'.

(3b) Short-before-long

Kunye-ka ppye-lul kkoli-lul huntunun kay-eykey cwu-ess-ta.
 she bone-ACC tail wagging dog-to gave
 'She gave a bone to a dog wagging his tail'. (Choi 2007: 209)

These language-specific ordering preferences have been linked to an interaction between general principles of efficient processing and differences in the headedness of the languages (Hawkins 1994, 2004, 2007). While the languages investigated thus far that exhibit a tendency to order short before long are head-initial (i.e. they order heads of syntactic constituents canonically before their complements), those that exhibit the opposite tendency (Japanese, Korean, Cantonese prenominal relative clauses) are head-final. According to Hawkins (2004), this correlation is due to the preference for choosing constituent orders that allow comprehenders to recognize immediate constituents as quickly as possible. In a head-initial language like English, shifting a heavy theme to follow the recipient in a sentence like *I introduced* [_{RECIPIENT} *to Mary*] [_{THEME} *some friends that John had brought to the party*] (Hawkins 1994) allows the two constituents of the VP to be recognized more quickly than in the reverse ordering, *I introduced*

some friends that John had brought to the party [to Mary], in which all words in the long theme NP must be processed before the verb's recipient can be identified. In the case of head-final structures, the mirror image of this is predicted to be preferred, because this similarly will keep the distance between the heads of the two constituents short. This type of theory allows for a mediating role of language-specific structure in determining the outcome of universal processing preferences.

Production-oriented accounts that acknowledge the need for language-specific differences have also been proposed: it is possible that Japanese speakers are more sensitive to conveying meaning (putting enriched material earlier), while English speakers prefer to sequence forms (putting easier to produce, e.g. shorter, words earlier, Yamashita and Chang 2001, 2006, based on Bock 1982). This of course leaves the question open as to what might account for these cross-linguistic differences in early accessing of form vs. meaning in incremental processing.

Regardless of which of these accounts turns out to be correct, data from non-Germanic languages challenge production accounts that attribute constituent order preferences to availability-based production, because these accounts predict a universal preference for short-before-long ordering (de Smedt 1994; Wasow 1997; see also references in Section 2.1). Thus, the relation between constituent complexity and constituent ordering is yet another example demonstrating the crucial role of cross-linguistic evidence both in ruling out postulated processing universals and in advancing our understanding of the mechanisms underlying language production.

2.5 OTHER CROSS-LINGUISTIC RESEARCH ON GRAMMATICAL ENCODING

In addition to the research themes discussed above, several other topics have been studied cross-linguistically, although they have received considerably less attention. For example, the tendency for speakers to reuse recently processed syntactic structures (syntactic priming, Bock 1986c; Pickering and Branigan 1998) has been replicated for a variety of syntactic structures in several languages (e.g. R. J. Hartsuiker and Westenberg 2000 on auxiliary-past participle ordering in Dutch; R. J. Hartsuiker and Kolk 1998 on Dutch passives and ditransitives; Scheepers 2003 on German RC attachment). This supports the view that syntactic processing of different languages shares at least some general, language-independent mechanisms, although it is important to keep in mind that the languages investigated so far are typologically very similar (we return to this point in the next section).

Syntactic priming has also been used to investigate the organization of linguistic knowledge in speakers that are proficient in two languages (bilinguals). Several studies have shown syntactic priming for bilinguals from one language to another (e.g. Dutch-English: Salamoura and Williams 2006; Dutch-German: Bernolet, Hartsuiker, and Pickering 2007; German-English: Loebell and Bock 2003; Greek-English: Salamoura and Williams

2007; Spanish-English: R. J. Hartsuiker, Pickering, and Veltkamp 2004; Hernandez, Bates, and Avila 1996; Meijer and Fox Tree 2003; for an overview, see R. J. Hartsuiker and Pickering 2008). This research has accumulated evidence that bilinguals' representations of syntactic structures that are present in both languages are shared between the two languages (for recent overviews, see Desmet and Duyck 2007; Salamoura and Williams 2007).

Beyond experimental psycholinguistic research, rich traditions in sociolinguistic, variationist, and comparative work have yielded evidence relevant to psycholinguistic theories of language production. An adequate discussion of the literature in these areas is beyond the scope of this paper (for examples, see Bresnan and Hay 2007; Strunk 2005; Tagliamonte and Smith 2005; Tagliamonte, Smith, and Lawrence 2005).

3 *Expanding the Empirical Base*

To get an estimate of how much work has been conducted on languages other than English, we elicited references from two international language news lists, read all suggested papers, and followed all references to work on other languages in them. While this survey is admittedly itself biased (being posted in English on lists *we* were aware of), it may suffice to make our point: We found psycholinguistically controlled production research on fewer than 30 of the world's 5000 to 10,000 languages. A sizable literature on sentence production (more than five papers) seems to exist for only seven languages: English, Dutch, German, French, Spanish, Italian, and Japanese, of which six fall into two language families (Germanic and Romance), both of which have developed from Indo-European and have been spoken in close geo-cultural proximity for many centuries.² This focus on a few genetically and areally related languages has resulted in a striking lack of typological diversity in the empirical base against which theories of sentence production are evaluated. By far most of the world's 200 language families have not been investigated psycholinguistically, leaving entire continents close to unexplored (including almost all native American, African, and Austronesian language families; see below for exceptions).

In the first part of this paper, we have given examples of production research where cross-linguistic differences provided crucial theoretical insights. But there are further reasons why psycholinguistic research will benefit from a typologically more diverse empirical base. We briefly elaborate on a few of them.

Different languages come with different structures and structural choices. For example, languages may differ in terms of word order and word order flexibility (e.g. English with rather fixed word order vs. Warlpiri, which allows many different constituent orders and even discontinuous constituents):

- (4a) Ngarrka-ngku ka wawirri pantirni (Warlpiri)
 man-_{ERG} _{PRES.IMP} kangaroo spear-Npst
 'The man is spearing the kangaroo'.
- (4b) Wawirri ka pantirni ngarrkangku
 Pantirni ka ngarrkangku wawirri
 Ngarrkangku ka pantirni wawirri
 Pantirni ka wawirri ngarrkangku
 Wawirri ka ngarrkangku pantirni (Hale 1983: 3)

Languages may have little inflectional morphology (e.g. English) or have complex morphological processes (e.g. Yup'ik Inuit, a polysynthetic language):

- (5) tuntussuqatarniksaitengqiggtuq (Yup'ik Inuit)
 tuntu -ssur -qatar -ni -ksaite -ngqiggte -uq
 reindeer -hunt -FUT -say -NEG -again -3SG:IND
 'He had not yet said again that he was going to hunt reindeer'. (Eliza Orr, cited by Payne 1997: 27–28)

They may signal argument functions via dependent-marking (e.g. case-marking on arguments as in, e.g. German) or via head-marking (e.g. clitics or pronominal inflection on the verb as in, e.g. Bulgarian or Yucatec Mayan). They may differ in terms of morphosyntactic alignment (e.g. ergative languages mark subjects of intransitive verbs like objects of transitive verbs, and distinctly from subjects of transitive verbs, while accusative languages treat subjects of both transitive and intransitive verbs alike); they may differ in the size of the lexicon; and so on. In other words, speakers of different languages are faced with different choices when encoding their message into an utterance, and the work of encoding is distributed differently across different levels of linguistic processing. This raises the question as to what extent speakers of different languages may employ different strategies in formulating their utterances.

A typologically narrow empirical base comes with the risk that mechanisms observed in all languages studied so far – and hence assumed to be universal – are in actuality due to typological features shared by all those languages. Crucially, there is evidence that even similar constructions may be processed differently in different languages, as seen, for example, in the discussion of agreement processing. Another example comes from animacy effects on grammatical encoding. Prat-Sala and Branigan (2000) found that Spanish speakers are more likely to left-dislocate animate patients compared to inanimate patients as in *The boy, the swing hit (him)*. Snider and Zaenen (2006) find the opposite, an anti-animacy effect, for English. Similar observations have been made by others. Myachykov (2007) summarizes: '[accessibility]-driven choices of word order are realized differently in different syntactic structures and in languages with different grammatical systems' (see also Bates and Devescovi 1989; Branigan,

Pickering, and Tanaka 2008; Chang, Lieven, and Tomasello 2008; V. S. Ferreira and Yoshita 2003; MacWhinney and Bates 1978; Rosenbach 2008; Sridhar 1988).

Just as linguistic theories have been transformed by cross-linguistic evidence, psycholinguistic theories are shaped by the empirical basis against which they are evaluated. By studying languages that differ typologically from those psycholinguistic researchers are familiar with, not only are we able to compare already well-explored phenomena across a broader range of language *types*, but we are bound to discover new *phenomena* themselves in need of explanation, data points that existing theories do not make predictions about. Consider, for example, a type of variation peculiar to dependent marking languages, so-called differential object marking (e.g. in Hindi, Aissen 2003b; de Hoop and Narasimhan 2005), where speakers have a choice between explicitly marking a direct object with case morphology or omitting that morphology.

- (7a) Ravii-ne kaccaa kelaa kaat̪aa (Hindi)
 Ravi-_{ERG} unripe banana cut._{PERF}
 Ravi cut the/an unripe banana.
- (7b) Ravii-ne kacce kele-ko kaat̪aa
 Ravi-_{ERG} unripe banana-_{ACC} cut._{PERF}
 Ravi cut the unripe banana. (Mohanan 1994: 87–88)

In the linguistics literature, whether an object is marked or not has been argued to be affected by a range of factors including its animacy, definiteness, topicality, and person (e.g. Aissen 2003b; de Hoop and Narasimhan 2005; Morimoto 2002). We have already seen how these factors figure prominently in psycholinguistic theories of accessibility, but it remains unclear what predictions (if any) these theories would make for differential object marking, given that the variation involves neither the subject function (Bock and Warren 1985) nor linear ordering (V. S. Ferreira and Dell 2000). Psycholinguistic research must draw from a wider typological base. We argue in the next section that this can be best achieved through collaboration between psycholinguistics and linguists.

4 *Cross-Linguistic Psycholinguistics as an Interdisciplinary Program*

Adequate cross-linguistic research on sentence production is impossible without detailed typological knowledge of the range of cross-linguistic variation that exists. It is also impossible without a detailed knowledge of the structural choices available in the particular language(s) under study. The ‘same’ word or structure translated into another language may come with different connotations, subtle different meanings, different subcategorization biases, and so on. For example, there is reason to believe that some apparent cross-linguistic differences found in early work on agreement processing (Vigliocco, Butterworth, and Garrett 1996) vanish or change direction

once translation equivalence of the stimuli is achieved (Bock et al., forthcoming; see also Section 2.3 above). While this does not invalidate cross-linguistic psycholinguistics, it supports our argument that cross-linguistic psycholinguistics requires collaboration between researchers familiar with psycholinguistic theory and methodology and researchers familiar with the target languages.

Also, while much linguistic work has not provided quantitative data (at least in the past), it has provided introspective evidence as to what factors influence speakers' choices, as shown above with the example of differential object marking. Indeed, the harmonic alignment accounts proposed for differential object marking in the linguistic literature (Aissen 2003a,b; see also linguistic alignment accounts for voice or word order variations, e.g. Bresnan, Dingare, and Manning 2001; Bresnan and Hay 2007) correspond closely to the indirect accessibility account of how conceptual accessibility affects word order (Bock and Warren 1985; Christianson and Ferreira 2005; F. Ferreira 1994) if one assumes not necessarily alignment to the subject function, but rather to particular unmarked/prototypical/frequent structural types.

An equally important benefit psycholinguistic researchers could receive from collaborations with linguists is methodological in nature. Since the vast majority of the world's languages (especially those typologically different from previously studied languages) are not spoken in the vicinity of psycholinguistic labs, it will be necessary to collect quantitative data in the field. Field-based psycholinguistics comes with a set of social and methodological challenges that, while not insurmountable, require expertise that is not commonly part of psycholinguistic training (Jaeger, Norcliffe, Bohnemeyer, and Nikitina 2008).

Psycholinguistics, too, brings much into the collaborative effort of cross-linguistic psycholinguistics. In addition to the obvious methodological advances that have been made in collecting quantitative data, psycholinguistic theory links linguistic phenomena to cognitive mechanisms, thereby (at least potentially) providing cognitive plausibility to linguistic theory.

Hawkins (2007) provides probably the most compelling argument for collaboration between psycholinguistics and typologists (see also Hawkins 1994, 2004; see Bender 2009 for a similar argument for collaborations between researchers in natural language processing and typologists). Hawkins summarizes evidence that many of the typological patterns observed across the world's languages reflect the processing preferences observed within speakers of one language. For example, Hawkins derives the typologically observed preference for SVO languages from the same principles of efficient processing that predict the short-before-long preference in head-initial languages like English (Arnold et al. 2000; Bresnan and Hay 2007; Hawkins 1999) and the long-before-short preference in head-final languages like Japanese (Yamashita and Chang 2001) and Korean (Choi 2007). This means that typological work can inform psycholinguistic theory and conversely,

psycholinguistic findings about processing mechanisms can provide explanations for cross-linguistically observed typological patterns (Hawkins 2004, 2007, 2009; Newmeyer 2005).

5 *The Dwindling Sample*

Cross-linguistic psycholinguistic research assumes an increased importance as much of the data on which theorizing can be based is disappearing. It is estimated that 90% of the world's languages will be extinct or moribund by the end of the twenty-first century (Hale, Krauss, and Watahomijie 1992; for a recent overview, see Romaine 2007). While many of these language communities realistically are already too small to conduct standard psycholinguistic research on them (Chung 2008), the majority still have enough speakers (70% of all languages are still spoken by more than 1000 speakers, about 40% by more than 10,000 speakers, Romaine 2007: 121). The alarming rate of language death lends additional urgency to the psycholinguistic study of these language communities while we still can.

Acknowledgments

We would like to thank I. Arnon, T. Wasow, A. Fine, M. Brown, N. Snider, K. Bock, and R. Hartsuiker, as well as our reviewers for feedback on this paper. We are also grateful to all those who replied to our postings and sent us references.

Short Biographies

T. Florian Jaeger is an Assistant Professor in the Brain and Cognitive Sciences at the University of Rochester. His research mostly falls within the domain of psycholinguistics, focusing on to what extent speakers and comprehenders process language efficiently. Together with his collaborators, Florian has investigated language production at a various levels of linguistic representation, from speech rate decisions via morphosyntactic variation to inter-clausal planning. His lab approaches these questions using both laboratory experiments and corpus-based computational methods, while aiming to extend the cross-linguistic empirical base against which psycholinguistic theories are evaluated. Florian holds an MA in Linguistic and Computer Science from Humboldt University, Berlin, and a PhD in Linguistics with a designation in Cognitive Science from Stanford University. Other education includes Psychology (post-doc, UC San Diego), Linguistics (MIT, UC Berkeley), and computer science (TU Berlin).

Elisabeth Norcliffe is a doctoral student in Linguistics at Stanford University. Her research is centered on syntactic variation from a processing and cross-linguistic perspective. Her dissertation draws on experimental and quantitative methods in order to investigate morphosyntactic variation

in Yucatec (Mayan) long-distance dependencies. Before starting her doctoral work at Stanford, she received a BA in German and an MA in Linguistics from the University of Canterbury in New Zealand. She is currently an ACLS/Mellon doctoral fellow.

Notes

* Correspondence address: Brain and Cognitive Sciences, Computer Science, University of Rochester, Meliora Hall, Box 270268, Rochester, NY 14627-0268, USA. E-mail: fjaeger@bcs.rochester.edu.

1 What exactly is meant by ‘heavy’ can be variously interpreted, e.g. as syntactic complexity (e.g. phrasal nodes), or as simple length. Wasow (2002) and Szmrecsányi (2004) report that structural complexity has no greater effect on ordering phenomena than simple phrasal length. Wasow and Arnold (2005) summarize further efforts to tease apart length and complexity as influences on ordering. In most research on the subject it is simple length that tends to be used as a measure of grammatical weight (Choi 2007).

2 The situation looks even more dire once dialects are considered. Languages usually consist of many dialects, which can show considerable structural variation. Yet hardly any psycholinguistic production studies have compared dialects (but see Bock et al. 2006).

Works Cited

- Aissen, J. (2003a). Differential Coding, Partial Blocking, and Bidirectional OT. Paper presented at the BLS 29, Berkeley, CA.
- . (2003b). Differential object marking: iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3).435–83.
- Arnold, J. E., T. Wasow, T. Losongco, and R. Ginstrom (2000). Heaviness vs. newness: the effects of structural complexity and discourse status on constituent ordering. *Language* 76(1).28–55.
- Bates, E. A., and A. Devescovi (1989). Crosslinguistic studies of sentence production. *The Crosslinguistics Study of Sentence Processing*, ed. by B. MacWhinney and E. A. Bates, pp. 225–53. Cambridge, UK: Cambridge University Press.
- Bates, E. A., and B. MacWhinney (1982). Functionalist approaches to grammar. *Language Acquisition: The State of the Art*, ed. by W. E. and L. R. Gleitman, pp. 173–218. Cambridge, UK: Cambridge University Press.
- Bender, E. M. (2009). Linguistically Naïve != Language independent: why NLP needs linguistic typology. *Proceedings of the EACL 2009 Workshop on the Interaction Between Linguistics and Computational Linguistics: Virtuous, Vicious or Vacuous?* pp. 26–32. Athens, Greece.
- Berg, T. (1998). The resolution of number conflicts in English and German agreement patterns. *Linguistics* 36.41–70.
- Bernolet, S., R. J. Hartsuiker, and M. J. Pickering (2007). Shared syntactic representations in bilinguals: evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(5).931–49.
- Bock, J. K. (1982). Toward a cognitive psychology of syntax: information processing contributions to sentence formulation. *Psychological Review* 89.1–47.
- . (1986a). Exploring levels of processing in sentence production. *Natural Language Generation: New Results in Artificial Intelligence, Psychology, and Linguistics*, ed. by G. Kempen, pp. 351–64. Dordrecht, the Netherlands: Martinus Nijhoff Publishers.
- . (1986b). Meaning, sound, and syntax: lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12(4).575–86.
- . (1986c). Syntactic persistence in language production. *Cognitive Psychology* 18.355–87.
- Bock, J. K., and D. E. Irwin (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior* 19.467–84.

- Branigan, H. P., and E. Feleki (1999). Conceptual accessibility and serial order in Greek language production. *Proceedings of the 21st Conference of the Cognitive Science Society, Vancouver*.
- Bock, J. K., and R. Warren (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21.
- Bock, J. K., and W. J. M. Levelt (1994). Language production: grammatical encoding. *Handbook of Psycholinguistics*, ed. by M. Gernsbacher, pp. 945–84. San Diego, CA: Academic Press.
- Bock, J. K., H. Loebell, and R. Morey (1992). From conceptual roles to structural relations: bridging the syntactic cleft. *Psychological Review* 99.150–71.
- Bock, J. K., M. Carreiras, and E. Meseguer (forthcoming). Number meaning and number agreement in English and Spanish.
- Bock, J. K., K. M. Eberhard, J. C. Cutting, A. S. Meyer, and H. Schriefers (2001). Some attractions of verb agreement. *Cognitive Psychology* 43.83–128.
- Bock, J. K., S. Butterfield, A. Cutler, J. C. Cutting, K. M. Eberhard, and K. R. Humphreys (2006). Number agreement in British and American English: disagreeing to agree collectively. *Language*, 82.64–113.
- Branigan, H. P., M. J. Pickering, and M. Tanaka (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118(172–89).
- Bresnan, J., and J. Hay (2007). Gradient grammar: an effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118(2) (Special Issue on Animacy).245–59.
- Bresnan, J., S. Dingare, and C. D. Manning (2001). Soft constraints mirror hard constraints: voice and person in English and Lummi. *Proceedings of the Lexical Functional Grammar Conference 2001 (LFG'01)*, ed. by M. Butt and T. Holloway King. Stanford, CA: CSLI Publications.
- Bresnan, J., A. Cueni, T. Nikitina, and H. Baayen (2007). Predicting the dative alternation. *Cognitive Foundations of Interpretation*, ed. by G. Boume, I. Kraemer and J. Zwartz, pp. 69–94. Amsterdam, The Netherlands: Royal Netherlands Academy of Science.
- Brown-Schmidt, S., and A. E. Konopka (2008). Little houses and casas pequeñas: message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition* 109.274–80.
- Brysaert, M., W. Fias, and M.-P. Noel (1998). The Whorfian hypothesis and numerical cognition: Is 'twentyfour' processed in the same way as 'four-and-twenty'? *Cognition* 66.51–77.
- Byrne, B., and E. Davidson (1985). On putting the horse before the cart: exploring conceptual bases of word order via acquisition of a miniature artificial language. *Journal of Memory and Language* 24(4).377–89.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology* 14.177–208.
- Caramazza, A., M. Miozzo, A. Costa, N. O. Schiller, and F.-X. Alario (2001). A cross-linguistic investigation of determiner production. *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler*, ed. by E. Dupoux, pp. 209–26. Cambridge, MA: MIT Press.
- Chang, F., G. S. Dell, and K. Bock (2006). Becoming syntactic. *Psychological Review* 113(2).234.
- Chang, F., E. Lieven, and M. Tomasello (2008). Automatic evaluation of syntactic learners in typologically different languages. *Cognitive Systems Research*, 9(3).198–213.
- Choi, H.-W. (2007). Length and order: a corpus study of Korean dative-accusative construction. *Discourse and Cognition* 14.207–27.
- Christianson, K., and F. Ferreira (2005). Planning in sentence production: evidence from a free word-order language (Ojibwa). *Cognition* 98.105–35.
- Chung, S. (2008). How much can understudied languages really tell us about how language works? Paper presented at the Annual Meeting of the Linguistic Society of America, Chicago, January.
- Costa, A., N. Sebastian-Galles, M. Miozzo, and A. Caramazza (1999). The gender congruity effect: evidence from Spanish and Catalan. *Language and Cognitive Processes* 14.381–91.
- de Hoop, H., and B. Narasimhan (2005). Differential case-marking in Hindi. *Competition and Variation in Natural Languages: The Case for Case*, 321–45.
- de Smedt, K. J. M. J. (1994). Parallelism in incremental sentence generation. *Parallelism in Natural Language Processing*, ed. by G. Adriaes and U. Hahn, pp. 421–47. Norwood, NJ: Ablex.

- Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review* 93.283–321.
- Dell, G. S., F. Chang, and Z. M. Griffin (1999). Connectionist models of language production: lexical access and grammatical encoding. *Cognitive Science*, 23(4).517–42.
- Dennison, H. Y. (2008). Universal versus language-specific conceptual effects on shifted word-order production in Korean: evidence from bilinguals. *Working Papers in Linguistics: University of Hawaii at Manoa* 39(2).1–16.
- Desmet, T., and W. Duyck (2007). Bilingual language processing. *Language and Linguistics Compass* 1(3).168–94.
- Dryer, M. S. (1980). The positional tendencies of sentential noun phrases in universal grammar. *Canadian Journal of Linguistics* 25.123–95.
- Eberhard, K. M. (1999). The accessibility of conceptual number to the processes of subject-verb agreement in English. *Journal of Memory and Language* 41.560–78.
- Eberhard, K. M., J. C. Cutting, and K. Bock (2005). Making syntax of sense: number agreement in sentence production. *Psychological Review* 112(3).531–59.
- Ferreira, F. (1994). Choice of passive voice is affected by verb type animacy. *Journal of Memory and Language* 33.715–36.
- Ferreira, F., and B. Swets (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic 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* 35.724–55.
- Ferreira, V. S., and G. S. Dell (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology* 40(4).296–340.
- Ferreira, V. S., and H. Yoshita (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research* 32(6).669–92.
- Foote, R. (2006). A Psycholinguistic Investigation of Agreement in Spanish and English Monolinguals and Bilinguals. University of Illinois, Urbana-Champaign.
- Garrett, M. F. (1976). Syntactic processes in sentence production. *New Approaches to Language Mechanisms (North-Holland Linguistic Series)*, ed. by R. Wales and E. Walker, pp. 133–77. Amsterdam, the Netherlands: Elsevier.
- . (1980). Levels of processing in sentence production. *Language Production*, ed. by B. Butterworth, Vol. 1 (Speech and Talk), pp. 177–220. London, UK: Academic Press.
- . (2000). Remarks on the architecture of language processing systems. *Language and the Brain*, ed. by Y. Grodzinsky, L. Shapiro and D. Swinney, pp. 31–69. San Diego, CA: Academic Press.
- Gleitman, L. R., D. January, R. Nappa, and J. C. Trueswell (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language* 57(4).544–69.
- Griffin, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin & Review* 10(3).603–9.
- Haeseryn, W. (Ed.). (1997). *Algemene Nederlandse Spraakkunst*. Groningen: Nijhoff.
- Hale, K. L. (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1(1).5–47.
- Hale, K. L., M. Krauss, and L. J. Watahomijie (1992). Endangered languages. *Language* 68.1–42.
- Hartsuiker, R. J., and H. H. J. Kolk (1998). Syntactic persistence in Dutch. *Language and Speech* 41(2).143–84.
- Hartsuiker, R. J., and M. J. Pickering (2008). Language integration in bilingual sentence production. *Acta Psychologica* 128.479–89.
- Hartsuiker, R. J., and C. Westenberg (2000). Word order priming in written and spoken sentence production. *Cognition* 75(2).B27–39.
- Hartsuiker, R. J., H. H. Kolk and W. J. Huinck (1999). Agrammatic production of subject-verb agreement: the effect of conceptual number. *Brain and Language* 69.119–60.
- Hartsuiker, R. J., M. J. Pickering, and E. Veltkamp (2004). Is syntax separate or shared between languages? *Psychological Science* 15(6).409–14.
- Hawkins, J. A. (1994). *A Performance Theory of Order and Constituency*. Cambridge, UK: Cambridge University Press.

- . (1999). The relative ordering of prepositional phrases in English: going beyond manner-place-time. *Language Variation and Change* 11.231–66.
- . (2004). *Efficiency and Complexity in Grammar*. Oxford, UK: Oxford University Press.
- . (2007). Processing typology and why psychologists need to know about it. *New Ideas in Psychology* 25.87–107.
- . (2009). Language universals and the performance-grammar correspondence hypothesis. *Language Universals*, ed. by C. Collins, M. H. Christiansen and S. Edelman. Oxford: Oxford University Press.
- Hernandez, A. E., E. A. Bates, and L. X. Avila (1996). Processing across the language boundary: a cross-modal priming study of Spanish-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(4).846–64.
- Humphreys, K. R., and J. K. Bock (2005). Notional number agreement in English. *Psychonomic Bulletin and Review* 12.689–95.
- Igoa, J. M. (1996). The relationship between conceptualization and formulation processes in sentence production: some evidence from Spanish. *Language Processing in Spanish*, ed. by M. Carreiras, J. E. Garcia-Albea and N. Sebastián-Gallés, pp. 305–51. Mahwah, NJ: Lawrence Erlbaum Associates.
- Jaeger, T. F., E. Norcliffe, J. Bohnemeyer, and T. Nikitina (2008). *Studying Language Production in the Field: Accessibility Effects on Variation*. Rochester: University of Rochester.
- Kelly, M. H., J. K. Bock, and F. C. Keil (1986). Prototypicality in a linguistic context: effects on sentence structure. *Journal of Memory and Language* 25.59–74.
- Kempen, G., and K. Harbusch (2003). Word order scrambling as a consequence of incremental sentence production. *Mediating Between Concepts and Grammar*, ed. by H. Härtl and H. Tappe, pp. 141–64. Berlin, Germany: Mouton De Gruyter.
- . (2004). A corpus study into word order variation in German subordinate clauses: animacy effects linearization independently of grammatical function assignment. *Multidisciplinary Approaches to Language Production*, ed. by T. Pechmann and C. Habel, pp. 173–81. Berlin, Germany: Mouton De Gruyter.
- Kempen, G., and E. Hoenkamp (1989). An incremental procedural grammar for sentence formulation. *Cognitive Science* 11(2).201–58.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Lindsay, J. R. (1975). Producing simple utterances: how far ahead do we plan? *Cognitive Psychology* 7.1–19.
- Loebell, H., and J. K. Bock (2003). Structural priming across languages. *Linguistics* 41.791–824.
- Lohse, B., J. Hawkins, and T. Wasow (2004). Processing domains in English verb-particle construction. *Language* 80(2).238–61.
- Lorimor, H., J. K. Bock, E. Zalkind, A. Sheyman, and R. Beard (2008). Agreement and attraction in Russian. *Language and Cognitive Processes* 23(6).769–99.
- MacWhinney, B., and E. A. Bates (1978). Sentential devices for conveying givenness and newness: a cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behavior* 17.539–58.
- . (1989). *The Cross-linguistic Study of Sentence Processing*. Cambridge, UK: Cambridge University Press.
- Matthews, S., and L. Y. Y. Yeung (2001). Processing motivations for topicalization in Cantonese. *Cognitive-Functional Linguistics in an East Asian Context*, ed. by K. Horie and S. Sato, pp. 81–102. Tokyo, Japan: Kuroosio Publishers.
- McDonald, J. L., K. Bock, and M. H. Kelly (1993). Word and word order: semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology* 25(2).188–230.
- Meijer, P. J., and J. E. Fox Tree (2003). Building syntactic structures in speaking: a bilingual exploration. *Experimental Psychology* 50(3).184–95.
- Miozzo, M., and A. Caramazza (1999). The selection of lexical-syntactic features in noun phrase production: evidence from the picture-word interference paradigm. *Journal of Experimental Psychology: Learning Memory and Cognition* 25.907–22.
- Mohanan, T. (1994). *Argument Structure in Hindi*. Stanford, CA: Center for the Study of Language and Information.
- Morimoto, Y. (2002). Prominence mismatches and differential object marking in Bantu.

- Proceedings of the LFG'02 Conference, ed. by M. Butt and T. Holloway King, Stanford, CA: Center for the Study of Language and Information Publications.
- Myachykov, A. (2007). Integrating Perceptual, Semantic and Syntactic Information in Sentence Production. Glasgow, UK: University of Glasgow.
- Myachykov, A., S. Garrod, and C. Scheepers (forthcoming). Attention and syntax in sentence production: a critical review. *Discours*.
- Myachykov, A., M. I. Posner, and R. S. Tomlin (2007). A parallel interface for language and cognition: theory, method, and experimental evidence. *The Linguistic Review* 24.457–75.
- Newmeyer, F. J. (2005). *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford, UK: Oxford University Press.
- Onishi, K. H., G. L. Murphy, and J. K. Bock (2008). Prototypicality in sentence production. *Cognitive Psychology* 56.103–41.
- Osgood, C. E., and J. K. Bock (1977). Salience and sentencng: some production principles. *Sentence Production – Developments in Research and Theory*, ed. by S. Rosenberg, pp. 89–140. Hillsdale, NJ: Erlbaum.
- Payne, T. E. (1997). *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge, UK: New York, NY: Cambridge University Press.
- Pickering, M. J., and H. P. Branigan (1998). The representation of verbs: evidence from syntactic priming in language production. *Journal of Memory and Language* 39.633–51.
- Prat-Sala, M., and H. P. Branigan (2000). Discourse constraints on syntactic processing in language production: a cross-linguistic study in English and Spanish. *Journal of Memory and Language* 42.168–82.
- Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: verb–participle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 94.904–21.
- Romaine, S. (2007). Preserving endangered languages. *Language and Linguistics Compass* 1(1–2).115–32.
- Rosenbach, A. (2005). Animacy versus weight as determinants of grammatical variation in English. *Language* 81.613–44.
- . (2008). Animacy and grammatical variation – findings from English genitive variation. *Lingua* 118.151–71.
- Salamoura, A., and J. N. Williams (2006). Lexical activation of cross-language syntactic priming. *Bilingualism: Language and Cognition* 9.299–307.
- . (2007). Processing verb argument structure across languages: evidence for shared representations in the bilingual mental lexicon. *Applied Psycholinguistics* 28.627–60.
- Scheepers, C. (2003). Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production. *Cognition* 89.179–205.
- Schiller, N. O., and A. Caramazza (2003). Gender or determiner selection interference? Evidence from noun phrase production in German and Dutch. *Journal of Memory and Language* 48(1).169–94.
- Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19.841–50.
- Schriefers, H., and E. Teruel (2000). Grammatical gender in noun phrase production: the gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26.1368–77.
- Smith, J., and L. Wheeldon (1999). High level processing scope in spoken sentence production. *Cognition* 73.205–46.
- Snider, N., and A. Zaenen (2006). Animacy and syntactic structure: fronted NPs in English. *Intelligent Linguistic Architectures: Variations on Themes* by Ronald M. Kaplan, ed. by M. Butt, M. Dalrymple and T. H. King, Stanford, CA: Center for the Study of Language and Information Publications.
- Sridhar, S. N. (1988). *Cognition and Sentence Production: A Cross-linguistic Study*. New York, NY: Springer Verlag.
- Stemberger, J. P. (1985). Bound morpheme errors in normal and agrammatic speech: One mechanism or two? *Brain and Language* 25.246–56.
- Strunk, J. (2005). The role of animacy in the nominal possessive constructions of modern Low

- Saxon. Paper presented at the Pionier Workshop on 'Animacy', Radboud University, Nijmegen, 19–20 May.
- Szmrecsányi, B. M. (2004). On Operationalizing Syntactic Complexity. Paper presented at the Le Poids des mots. 7th International Conference on Textual Data Statistical Analysis, Louvain-la-Neuve.
- Tagliamonte, S., and J. Smith (2005). No momentary fancy! The zero 'complementizer' in English dialects. *English Language and Linguistics* 9(02).289–309.
- Tagliamonte, S., J. Smith, and H. Lawrence (2005). No taming the vernacular! Insights from the relatives in northern Britain. *Language Variation and Change* 17.75–112.
- Tanaka, M. N., H. P. Branigan, and M. J. Pickering (forthcoming). Conceptual influences on word order and voice in Japanese sentence production.
- Uszkoreit, H. (1987). *Word Order and Constituent Structure in German*. Stanford, CA: Center for the Study of Language and Information.
- van Nice, K. Y., and R. Dietrich (2003). Task-sensitivity of animacy effects: evidence from German picture descriptions. *Linguistics* 5.825–49.
- Vigliocco, G., and J. Franck (1999). When sex and syntax go hand in hand: gender agreement in language production. *Journal of Memory and Language* 40.455–78.
- . (2001). When sex hits syntax: gender agreement in sentence production. *Journal of Memory and Language* 45.368–90.
- Vigliocco, G., and R. J. Hartsuiker (2002). The interplay of meaning, sound, and syntax in language production. *Psychological Bulletin* 128.442–72.
- Vigliocco, G., B. Butterworth, and C. Semenza (1995). Constructing subject–verb agreement in speech: the role of semantic and morphological factors. *Journal of Memory and Language* 34.186–215.
- Vigliocco, G., B. Butterworth, and M. E. Garrett (1996). Subject–verb agreement in Spanish and English: differences in the role of conceptual constraints. *Cognition* 61.261–98.
- Vigliocco, G., R. J. Hartsuiker, G. Jarema, and H. H. Kolk (1996). One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes* 11(407–42).
- Wasow, T. (1997). End-weight from the speaker's perspective. *Journal of Psycholinguistics Research* 26.347–62.
- . (2002). *Postverbal Behavior*. Stanford, CA: Center for the Study of Language and Information Publications.
- Wasow, T., and J. Arnold (2005). Intuitions in linguistic argumentation. *Lingua* 115(11).1481–96.
- Wheeldon, L., and A. Lahiri (1997). Prosodic units in speech production. *Journal of Memory and Language* 37.356–81.
- Yamashita, H., and F. Chang (2001). Long before short preference in the production of a head-final language. *Cognition* 81(2).B45–55.
- . (2006). Sentence production in Japanese. *Handbook of East Asian Psycholinguistics*, ed. by M. Nakayama, R. Mazuka and Y. Shirai, Volume 2, Japanese. Cambridge, UK: Cambridge University Press.