Chapter 5
Vision and Language in Cross-Linguistic Research on Sentence Production

Elisabeth Norcliffe and Agnieszka E. Konopka

5.1 Introduction

A classic assumption in the cognitive sciences has been the existence of an invariant cognitive architecture underlying language processing. Perhaps for this reason, the early days of modern psycholinguistics saw little engagement with cross-language data. Indeed, despite the vitality of the discipline in the 1960s and 1970s, it is striking that the language in which experiments were conducted were seldom explicitly considered as an important potential variable of interest. As Cutler (1985) observes, during this period, experiments undertaken in one language could be supported or refuted by experiments in another language, without consideration of whether the particular language could have affected the processes under study. Similarly, the major psycholinguistic textbooks of the day did not make reference to cross-linguistic data or comparative argumentation (see, e.g., Fodor et al. 1974; Glucksberg and Danks 1975).[1]

Beginning in the 1980s, this picture began to change, as researchers in the field of speech perception began to consider cross-linguistic differences more closely. In the process, the blanket universalist assumption was challenged as it was revealed that certain critical cross-linguistic differences can, in fact, affect language processing (Cutler et al. 1983, 1986, 1989; Mehler et al. 1993, 1996). Mehler et al. (1981) found, for example, that French speakers responded faster in

[1]It is interesting to observe that this contrasts starkly with the flurry of cross-linguistic research that was undertaken in the sister field of language acquisition during the same period, in large part due to the pioneering work of Dan Slobin.

E. Norcliffe (✉) · A.E. Konopka
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
e-mail: elisabeth.norcliffe@mpi.ni

R.K. Mishra et al. (eds.), *Attention and Vision in Language Processing*, DOI 10.1007/978-81-322-2443-3_5

a syllable detection task when the target corresponded to the first syllable of the stimulus word, compared to targets that consisted of a segment longer or shorter than the first syllable. This was evidence that speakers segment the speech into syllabic units prior to lexical access. Strikingly, however, the results did not generalise to English: Cutler et al. (1983) failed to find a syllable advantage effect for English speakers when they performed an equivalent task. On the basis of these different response patterns, Cutler et al. (1986) proposed that speakers of languages with different rhythmic properties parsed speech differently: for languages such as French, with clear syllabic boundaries, listeners use a syllabic representation to segment speech. For stress languages, such as English, such a representation is not used. Thus, listeners from different language backgrounds appear to rely on different processing routines. This and other seminal work launched a new focus on reconciling the language specific and the universal in speech processing.

In contrast with speech processing, most current models of language production have tended to continue to emphasise the universal aspects of the process (Levelt 1989). Here too, however, there has been scattered evidence of language-mediated processes. To date, most of these findings come from studies of noun phrase production. For example, Costa et al. (1999), Miozzo and Caramazza (1999), and Caramazza et al. (2001) demonstrate cross-linguistic differences in the time-course of determiner production. In German and Dutch, determiners agree in gender with the noun they combine with. For example, in Dutch, *de* is used for common gender nouns (e.g. *de tafel*, 'the table'), and *het* is used for neuter gender nouns (e.g. *het boek*, 'the book'). In Romance languages, such as Catalan, Spanish, Italian and French, determiners are selected not only on the basis of the gender of the noun, but also on the phonological form of the onset of the word that immediately follows it. For example, in Catalan, the masculine determiner *el* is used when the following word begins with a consonant (*el got*, 'the glass', *el meu ull*, literally "the my eye"), and with *l'* when the following word begins with a vowel (*l'ull* 'the eye'). In picture-word interference tasks, Dutch and German speakers are slower to name pictures of objects when a printed distractor phrase mismatches in gender with that of the depicted object (Schriefers 1993; Schiller and Caramazza 2003). While this so-called 'gender congruency' effect is robustly attested for German and Dutch, it has not been found for speakers of Romance languages. Miozzo and Caramazza (1999; see also Caramazza et al. 2001) suggest that this may reflect differences in the timing of determiner selection during noun phrase production in the two sets of languages. The authors propose that the gender congruency effect reflects competition between determiners; the selection of the target determiner is slowed when a different determiner is simultaneously activated by the distractor word. In Romance languages, the lack of gender congruency effect reflects the fact that determiner selection takes place later in the production process because it requires access not just to lemma-level information of the noun (gender information), but also the phonological form of the following word. As a result of this time lag, potentially conflicting information from the distractor determiner has already dissipated. In sum, this body of research suggests that languages differ in the extent to which information at different levels of

processing (conceptual, grammatical and phonological) interacts during the selection of closed class words.

Similarly, cross-linguistic differences in the *order* of words within noun phrases also influence the timing of word retrieval. Janssen et al. (2008) reported that linear word order can be responsible for differences in the timing of phonological activation in noun-adjective combinations; in French, where nouns typically precede adjectives, noun phonology was activated earlier than in English, where adjectives appear before nouns. This indicates that noun phrase production can proceed in a highly incremental fashion (word for word), allowing for language-specific variation in the time-course of formulation.

Interestingly, cross-linguistic differences in word order appear to influence not just the timing of grammatical-level processes, but even conceptual formulation. Brown-Schmidt and Konopka (2008) tested whether speakers of English and Spanish incorporate size information into modified expressions like *the small butterfly* and *la mariposa pequeña* at different points in time, consistent with the surface linear order of nouns and modifiers in the two languages. Eye-tracked speakers described highlighted object pictures (e.g., *butterfly*) presented in large displays. Speakers were expected to mention object size if they noticed a second referent in the display that differed from the target referent only in size (i.e., a size contrast: a large butterfly). The results showed that, when speakers produced modified noun phrases, fixations to the size contrast occurred on average later in Spanish (where modifiers follow nouns) than in English (where modifiers precede nouns). This suggests that linear word order afforded more flexibility in the timing of generation of the preverbal message in Spanish than in English.

Studies such as these suggest that, at least at the phrasal level, cross-linguistic differences in the nature of the syntactic dependencies between elements in phrases, or differences in the relative ordering of those elements, can influence processing routines. Such cross-linguistic differences can systematically affect the time-course of the computations involved in conceptual and linguistic planning.

In this chapter we consider the question of whether and how the planning processes involved in producing whole sentences—i.e., utterances with multiple referents and a more complex hierarchical structure—might be fine-tuned to language-specific properties. Specifically, we ask to what extent language structure can affect the breadth and order of the conceptual and linguistic encoding operations that take place during sentence formulation. We survey the small body of cross-linguistic research that bears on this question, focusing in particular on recent evidence from eye-tracking studies. The relatively recent application of visual world eye-tracking techniques to language production research (Griffin and Bock 2000; Griffin 2004; Meyer et al. 1998) has yielded important insights into the time-course of sentence formulation. Because eye-tracking methods provide a very fine-grained temporal measure of how conceptual processing and utterance planning unfold in real time, they serve as an important complement to standard approaches based on coarser temporal measures such as speech onset latencies, or offline measures such as structure choice. Significantly, the development of portable eye-trackers in recent years has, for the first time, allowed eye-tracking

techniques to be used with language populations that are located far away from university laboratories. This has created the exciting opportunity to extend the typological base of vision-based psycholinguistic research. Illustrating these advances, we describe results from studies carried out in the field with two verb-initial languages: Tzeltal (Mexico) and Tagalog (Philippines).

5.2 Incremental Sentence Formulation

Producing spoken language requires transforming an abstract idea or a communicative intention into a linear string of words. According to most models of language production (e.g., Levelt 1989), the first stage involves formulating a *message*, i.e. an abstract, preverbal representation of the information that the speaker wants to express. The message must then undergo *linguistic encoding* in preparation for articulation. Linguistic encoding itself involves several component processes, including selecting and retrieving the words appropriate for conveying the message, and integrating them into a sentence structure. It is generally assumed that this entire process unfolds *incrementally* (Kempen and Hoenkamp 1987; Levelt 1989): we do not plan everything we say in advance of opening our mouths. Rather, in an incremental system, speech can begin once some minimal chunk of the utterance is prepared, with the planning of subsequent material taking place as speaking unfolds over time.

Given the assumption of incrementality, the central debates in sentence production research revolve largely around the question of how much information speakers *can* and *do* plan at the conceptual level (i.e., the preverbal message) and at the sentence level before initiating overt production. This question has a venerable tradition in psycholinguistic research, dating back to the very genesis of the discipline. One view, first articulated by Paul (1886/1970), and reflected in modern 'word-driven' or 'linearly incremental' approaches to formulation, holds that speaking is a highly opportunistic process, in which the relative availability of individual concepts in a message determines the order in which words are retrieved. On this view, increments at the message level and at the sentence level can be very small (perhaps as small as a single concept or word): speakers may encode as little as one content word before speech onset, and the structure of the rest of the sentence is automatically constrained by whichever content word happens to be retrieved first. An alternative view, first espoused by Wundt (1900), and recapitulated in modern 'structure-driven' theories of formulation, holds that formulation begins with the generation of a larger conceptual representation of the message and, from there, a structural representation of the sentence. This structural plan in turn guides the order of subsequent word retrieval operations. Empirically, differences between these views have been addressed by considering implications of different planning strategies for the selection of *starting points* (MacWhinney 1977): when preparing an utterance, what (and how much) do speakers plan first?

Linearly incremental ('word-driven') formulation is supported by the robust cross-linguistic finding that speakers make structural choices that allow them to place *accessible* (roughly, more easily retrievable) information earlier in sentences (Arnold et al. 2000; Bock and Warren 1985; Branigan and Feleki 1999; Ferreira and Yoshita 2003; MacWhinney and Bates 1978; see Jaeger and Norcliffe 2009, for a review). Accessibility may depend, for example, on a referent's *perceptual* salience and can be enhanced by exogenous attention-capturing cues (Gleitman et al. 2007; Ibbotson et al. 2013; Myachykov and Tomlin 2008; Tomlin 1995, 1997). Referents may also differ in their *conceptual accessibility*, which includes such features as animacy, imageability or givenness. For example, speakers of English are more likely to produce passive structures (where the patient is expressed as the sentence-initial subject), when the patient is animate (Bock et al. 1992) or imageable (Bock and Warren 1985). For example, they are more likely to say "the man was hit by the ball", than "the ball hit the man", as this allows the human argument to be expressed as the sentence-initial subject. Effects like these are compatible with the view that speakers begin formulation by retrieving the first readily available word, and that this initial choice constrains the structure of the rest of the sentence.

This interpretation is supported by evidence from visual-world eye-tracking studies. In these paradigms, speakers describe pictures of simple events while their eyes are tracked. As noted by Bock et al. (2004), eye-tracking is particularly well suited to examining theories about incrementality in sentence formulation because speakers typically look at the things they want to talk about. Thus, the distribution of attention and the timing of gaze shifts to the various characters in an event provide fine-grained temporal information about when various elements of the message and sentence are planned (see discussions in Henderson and Ferreira 2004). Gleitman et al. (2007) used this method in combination with an implicit visual cueing procedure. Participants' attention was directed to one or another character in the event by means of a fleeting, subliminal visual cue (a black square). Gleitman et al. found that English speakers preferentially fixated the visually cued character within 200 ms of picture onset and that they tended to select that character to be the first-mentioned referent in their sentence. This result suggests that sentence formulation can indeed begin with the conceptual and linguistic encoding of as little as a single referent. Such results are also generally consistent with theories assuming that the order in which various encoding operations are performed depends on relative states of *activation* at different levels of representation in the production system.

An important constraint on interpreting such findings and generalising them across languages, however, is the fact that languages vary considerably on a number of grammatical dimensions that can be relevant for incremental formulation. One of these dimensions is linear word order: a significant complication in interpreting accessibility effects like the ones described by Gleitman et al. (2007) for English is that the first-mentioned element in the sentence also happens to be the subject of the sentence. Thus in subject-initial languages like English, it is difficult to tease apart whether accessibility influences *linear word order* directly or

whether it influences *subject assignment* (Bock and Warren 1985; McDonald et al. 1993), and only indirectly word order. A strong or 'radical' version of linear incrementality (Gleitman et al. 2007) would hold that accessibility directly drives lexical encoding and that early formulation involves little grammatical encoding (assignment of grammatical functions); on this view, subject assignment *follows* from whichever element is lexically retrieved first. The alternative view would be that planning the first character and retrieving the first content word involves not only the lexical encoding of one message element, but also the assignment of that element to the subject function: subject assignment implies some advance planning of the relational structure of the event (who is doing what to whom) as well as some grammatical-level processing.

Studies of languages that allow word order scrambling and thus do not confound subject position and sentence-initial position provide support for both possibilities. Some work has found that accessible concepts are more likely to become subjects, rather than simply sentential starting points (Christianson and Ferreira 2005, for Odawa). Other work (both experimental and corpus-based) has shown that conceptual accessibility can directly affect word order, even when grammatical function (subjecthood) is controlled for (Branigan and Feleki 1999, for Greek; Ferreira and Yoshita 2003, for Japanese; Kempen and Harbusch 2004, for German; MacWhinney and Bates 1978, for Italian and Hungarian). There is also recent evidence to suggest that within a language, both word order *and* grammatical function assignment may be influenced by conceptual accessibility (Tanaka et al. 2011, for Japanese).

In contrast, other eye-tracking evidence from English is more compatible with the structure-driven view of formulation. Using an eye-tracked picture description task, Griffin and Bock (2000) found that English speakers did not preferentially fixate either character in the depicted events within the first 400 ms of picture onset. Only after 400 ms did they direct their gaze preferentially to the character they would mention first. The authors interpret this as evidence of an early prelinguistic 'gist apprehension' phase, in which speakers encode the relationship between event characters before beginning linguistic encoding of the first-mentioned character. On this account, early gist apprehension allows for the generation of a larger message representation and, on this basis, selection of a suitable structural frame; this information, in turn, controls the order in which speakers encode individual event characters linguistically. More generally, unlike word-driven formulation, the 'structure-driven' view predicts that visual or conceptual salience of individual characters plays a subordinate role to 'wholistic' message-planning processes: speakers look to the character they will mention first not because their attention was initially drawn to it (contrary to Gleitman et al. 2007), but because their eyes were guided there by the structural framework generated shortly after picture onset (Bock et al. 2004).

Further support for structure-driven formulation comes from a study by Lee et al. (2013), who examined the structure of advanced planning in English using a picture description task. Analysis of speech onset times and word durations showed that when producing relative clause constructions (*the student of*

the teacher who is raising her hand), structurally dependent lexical items were planned together, suggesting that formulation involved advance hierarchical planning of a sentence structure.

5.2.1 Flexibility

In sum, there is a range of evidence to support both the word-driven and the structure-driven view of sentence formulation. Recently, evidence has accumulated to suggest that mixed findings like these can also reflect the fact that the time-course of formulation is flexible: while speakers are *able* to prepare increments consisting of multiple referents before initiating overt production, they may plan either more or less information before initiating speech under different conditions. Under certain circumstances, speakers might begin formulation by encoding isolated bits of information; while under other circumstances, they may begin by encoding the entire relational structure of a message, with linguistic formulation accordingly affected by the nature of the initial conceptual planning.

Factors that contribute to reductions in the scope of advance planning include extra-linguistic variables like time pressure (Ferreira and Swets 2002), cognitive load (Wagner et al. 2010) and differences in working memory capacity (Swets et al. 2008)—all of which can constrain the amount of information that speakers can encode in parallel in a given time window. Another set of factors concerns production processes proper, for example, resource constraints affecting the coordination of lexical and structural processes (Konopka 2012), or the relative ease of formulating a message plan. Kuchinsky and Bock (2010) found, for example, that attentional cueing had an effect on first mention (replicating Gleitman et al. 2007) for events for which the relation between the characters was difficult to conceptualise or interpret. In other words, for hard-to-interpret events, directing attention to one character in the event resulted in early mention of this character in the sentence, consistent with word-driven formulation. For easily encodable events, by contrast, attentional cueing had no effect on first mention. This suggests that fast encoding of a rudimentary message structure, and not character accessibility, mediated subject selection. Thus, different formulation strategies may be induced by differences in how ‘hard’ or ‘easy’ it is to apprehend the relational content of a message.

More generally, evidence of flexibility in the incremental preparation of messages and sentences motivates one key conclusion and makes one important prediction for research on sentence formulation across languages. The conclusion is that differences across studies may be a natural outcome of differences in the way that speakers coordinate encoding or prioritise encoding of different types of information (individual elements of a message versus the message “as a whole”) when preparing their utterances. The prediction then is that details of this coordination should be sensitive to the order in which words must be ultimately produced in an utterance. Incrementality naturally assumes that some parts of a message or

sentence are encoded before other parts of a message or sentence undergo encoding; thus, to produce language efficiently, formulation can benefit from speakers' ability to prioritise encoding those parts of a message and sentence that must be expressed first. This implies that the time-course of formulation should vary across languages with different basic word orders, because language-specific constraints on word order should license allocation of resources to encoding different parts of a message or sentence at different points in time. In the next section, we review evidence suggesting that the order of encoding operations during sentence formulation may indeed depend on the grammatical properties of the target language—and that this may naturally result in planning patterns that resemble either word-driven or structure-driven planning.

5.3 Cross-Linguistic Differences

To what extent can reliance on different planning strategies be driven by *grammar*? To a certain degree, it may seem fairly unquestionable that the formulation process would be influenced by language-specific constraints, given that the target structures of linguistic encoding are language-specific. A key question for theories of incrementality, however, is where and how far up in the production system language-specific properties might be expected to exert an influence on formulation. We address this question by considering how grammar influences encoding of complex relationships between elements of a message. We first review a set of studies that suggest that the grammars of languages may differ in the extent to which they are compatible with word-driven formulation. Then we turn to the special case of verb-initial languages, and ask whether the sentential position of the verb, as well as its morphological properties, can exert an influence not just on the timing of linguistic-level formulation processes, but also on message-level formulation itself.

5.3.1 Different Grammars, Different Formulation Preferences

To date, research on the time-course of sentence formulation comes largely from work on English. Arguably, the grammar of English affords a high degree of flexibility in planning. Sentences typically begin with subjects, which are not morphologically dependent on any other element in the sentence. As such, speakers may begin by selecting and retrieving a single noun lemma, without engaging in any advance planning of the rest of the message or planning of a sentence frame (consistent with radical, word-driven incrementality). Alternatively, nothing in the grammar prevents a structure-driven formulation process either: speakers may, in principle, begin formulating their sentences by encoding some of the hierarchical relationships between message elements early in the formulation process.

However, while linear incrementality is compatible with the syntax of English, it is apparently more problematic for other types of grammars. Evidence for this conclusion comes from a set of cross-linguistic studies that have employed the attentional cueing paradigm to study the effects of perceptual accessibility on structure choice. Myachykov et al. (2010) compared the performance of English and Finnish speakers in a task modelled on Gleitman et al. (2007). While the data from English participants replicated earlier results (speakers were more likely to begin their sentence by mentioning the cued referent first), the authors failed to find any effect of perceptual salience in Finnish: Finnish speakers consistently produced transitive SVO sentences, regardless of the position of the cued referent. Notably, Finnish is a *case-marking* language: the authors suggest that early commitment to a case-marker requires a larger degree of pre-planning compared to a language like English that lacks case marking on nouns. In Finnish, then, reliance on word-driven formulation may be attenuated by its case-marking properties.

Further empirical support for this possibility comes from Korean, which, like Finnish, is a case-marking language. Hwang and Kaiser (2009) found for Korean that priming patient characters with semantic prime words (to increase their conceptual accessibility) or employing visual cues (to increase their perceptual accessibility) did not influence structure choice. Thus, just as in the case of Finnish, attentional salience did not affect speakers' choice of sentential starting points.

In sum, it seems, logically, that in order to produce an initial case-marked noun phrase, speakers need to have already engaged in more than simple word retrieval to begin production: the selection of the appropriate case-marker on a sentence-initial noun should necessitate the early selection of a grammatical function for this noun. This would require some advance planning of the relational structure of the target message and some grammatical-level processing. As a result, not all languages allowing early mention of sentence subjects may equally support a radically incremental, word-driven formulation process. While English speakers can and do engage in word-driven formulation, this tendency does not generalise across languages in similar experimental paradigms.

There is also evidence that languages may differ in the *extent* to which they rely on a given formulation strategy. Myachykov and Tomlin (2008) employed an explicit cueing procedure to study the effects of attentional salience on Russian structure choice. The cueing procedure they employed was modelled on an earlier study on English by Tomlin (1995), which made use of an animated programme referred to as the "Fish Film". In this task, participants viewed and described a series of animations of two differently coloured fish swimming towards each other, culminating with one (the agent) eating the other (the patient). In each trial, an arrow appeared above one of the fish, and participants were explicitly instructed to look at the cued fish, and then describe the scene however they liked. Tomlin found that English speakers very consistently began their sentences with the cued fish (producing nearly 100 % actives when the agent was cued, and passives when the patient was cued). Myachykov and Tomlin (2008) repeated the study with Russian speakers to test which structural preferences would be revealed in a language that allowed more structural choices: Russian has a passive alternation like

English, but it also allows scrambling (object-subject as well as subject-object word order). The authors found that speakers tended to produce active subject-first constructions when the agent was cued (i.e., allowing the cued character to be the sentence-initial element), and active object-first constructions when the patient was cued (again, allowing the cued character to be the sentence-initial element). Notably, this effect was much smaller (around 20 %) compared to the almost 100 % effect size observed for English speakers. Interestingly, the passive was produced very rarely (only around 2 % of the time).

Several conclusions can be drawn from these findings. First, the fact that speakers typically started their sentence with the cued character is evidence that, as in English, speakers of Russian can adopt a word-driven formulation strategy. Second, the fact that the effect size was much smaller in Russian suggests that different languages accommodate attentional effects differently. One possible explanation for the cross-linguistic difference in effect size may be that Russian, like Korean and Finnish, is a case-marking language. Alternatively (or additionally), it could stem from differences in the relative frequencies of the different structural options in the two languages. In Russian there was a greater overall tendency to rely on canonical active SVO sentence structures, regardless of the position of the cue: this may reflect the overall high frequency of active SVO structures in Russian, compared to other structural alternatives (especially compared to the passive, which is reportedly very rare; Myachykov et al. 2011).

In this vein, MacDonald (2013) suggests that some cross-linguistic differences in production strategies may emerge from differences in how strongly specific structures are favoured in a given language. It is robustly attested that speakers have a tendency to reuse recently produced structures. This tendency, referred to as structural persistence, or syntactic priming, is often assumed to be the result of long-term implicit learning of structure-building procedures (Bock and Griffin 2000; Bock et al. 2007; Chang et al. 2006; Jaeger and Snider 2013). On this view, speakers constantly learn from their own productions and the productions of others: the more often a structure is used or heard, the more likely it is to be used again. Such structural biases may, therefore, induce a structure-driven formulation strategy over a word-driven strategy, by facilitating the mapping between a message and an abstract structural representation. Evidence of this relationship already exists for English and Dutch (Konopka 2012; Konopka and Meyer 2014; Van de Velde et al. 2014). On the assumption that languages differ in terms of the strength (and direction) of their overall structural biases, it is possible that these differences may also give rise to cross-linguistic differences in the extent to which speakers' structural choices are driven by lexical availability (see also Gennari et al. 2012).[2]

In sum, evidence from attentional cueing studies suggests that structural choices are not affected by perceptual salience equally across languages. In some languages, such as English, entities that are made accessible via visual cueing exert an influence on structural choices, consistent with word-driven incremental

[2]This of course leaves open the interesting question of what gives rise to such cross-linguistic differences in frequency distributions to begin with.

formulation. Other languages appear to show little sensitivity to perceptual salience, suggesting that linearization in these languages is not affected by the accessibility of individual message entities in the same manner. Such cross-linguistic differences in reliance on word-driven formulation may be due to grammatical differences: as reviewed above, one set of languages that has, to date, been found to exhibit little or no effects of perceptual accessibility are case-marking languages. Plausibly, case-marking is not readily compatible with a strongly word-driven formulation process as it necessitates the early assignment of grammatical functions to arguments. Of course, the strong version of word-driven formulation is not tenable for longer utterances in languages like English either: English speakers may certainly *begin* utterances with accessible words, but word order in the rest of the utterance must obey certain grammatical constraints. However, the simple fact that there *are* reported differences across languages with respect to effects of perceptual salience on subject selection suggests that this aspect of sentence formulation may be modulated by language-specific properties.

5.3.2 A View from Verb-Initial Languages

The cross-linguistic studies discussed so far have all shared a common property: they all concern subject-initial languages. Subject-initial languages in fact make up the vast majority of languages of the world. According to the World Atlas of Language Structures, SOV and SVO languages together constitute around 76 % of the world's languages (41 and 35 % respectively). It is therefore unsurprising that, to date, most psycholinguistic studies have centred on this language type. Far rarer, and far more under-studied, are languages whose basic sentences do not start with subjects, or indeed, with nouns of any grammatical function: i.e., verb-initial languages. VSO and VOS languages together make up around 8 % of the world's languages (6 and 2 % respectively). They offer a particularly interesting test case for studying the effects of grammar on sentence formulation: in order to produce a verb-initial sentence, relational information presumably *must* be planned early in order to retrieve an appropriate sentence-initial verb. Comparing the time-course of sentence formulation for verb-initial and subject-initial languages provides a unique means of assessing the extent to which message and sentence formulation may be influenced by a language's basic word order (also see Hwang and Kaiser 2014, for evidence from a verb-final language).

As outlined above, it is of course self-evident that to a certain extent, formulation will be affected by linear word order. Eye-tracking studies have already established empirically, moreover, that *within* English, the order of words in sentences affects the order in which they are encoded linguistically. For example, when preparing to produce an active sentence, speakers first fixate the agent (the sentence-initial subject) and then the patient (Griffin and Bock 2000). This suggests that, at least in the context of simple picture description tasks, speakers lexically encode the noun phrases in their sentences in order of mention. Similar left-to-right order

effects are observed within noun phrases as well (Janssen et al. 2008). It seems reasonable to assume, then, that *across* languages, word order differences would also affect the order of encoding operations involved in linguistic formulation.

Could word order influence the time-course of *message* formulation? According to top-down models of sentence production, this possibility is in fact ruled out on theoretical grounds: given the principles of information encapsulation and unidirectionality, message preparation should unfold without recourse to information encapsulated at linguistic levels of formulation (Bock and Levelt 1994; Garrett 1980; Levelt 1989). This predicts that message-level ordering decisions should not vary as a consequence of language-specific word order constraints.

As we discussed above, there is already evidence at the level of noun phrase production to suggest, contrary to the predictions of top-down models, that even message-level encoding operations may be affected by word order (Brown-Schmidt and Konopka 2008). It remains an empirical question whether or not we should expect to find such effects in longer utterances. In fact, neither of the two theories of incrementality we have considered so far would predict an effect of word order on message-level processes. According to radical, word-driven incrementality, linearization is controlled by the relative availability of individual elements of a message: the resulting word order of an utterance is thus assumed to be constrained by the properties of the message, and not the other way around. According to structure-driven formulation, speakers begin formulation by generating a 'wholistic' message plan that triggers building of a structural sentence plan and then word retrieval. In Griffin and Bock's (2000) version of this theory, the initial phase of message formulation ('gist apprehension') is assumed to be isolated from subsequent linguistic processes. Thus, the time-course of message formulation itself would not be predicted to vary as a consequence of the linear word order of the target utterance.

5.3.3 Tzeltal

To address this question, Norcliffe et al. (in press) conducted an eye-tracked picture description task in Tzeltal, a Mayan language spoken in Mexico by over 400,000 speakers (Polian 2013). Tzeltal's basic word order is VOS. The language also optionally permits SVO word order, thus allowing a within-language contrast of how sentence formulation might vary as a consequence of the linear position of subjects and verbs. For a direct comparison with an SVO language, the same experiment was also carried out with speakers of Dutch. The methodology was modelled on previous picture description studies (Griffin and Bock 2000; Konopka and Meyer 2014): speakers described pictures of simple transitive events involving familiar actions and characters while their speech and gaze were tracked.

Speakers described simple pictured events (e.g., a woman chasing a chicken) eliciting transitive descriptions, of the type exemplified below. [1] shows an

example sentence with VOS word order, and [2] shows the alternative SVO word order (the inflected verb stem, together with the preceding aspect marker, is underlined in both examples):

[1] ya s-nuts-ø me'mut te ants=e
INC[3] 3S.ERG-chase-3S.ABS chicken the woman=CL
"The woman is chasing a chicken"
(VERB-OBJECT-SUBJECT)

[2] te ants=e ya s-nuts-ø me'mut
the woman=CL INC 3S.ERG-chase-3S.ABS chicken
"The woman is chasing a chicken"
(SUBJECT–VERB–OBJECT)

We hypothesised that if word order mediated the relationship between the uptake of visual information in an event and the formulation of a description of that event, then early verb placement would require earlier encoding of relational information. The degree to which speakers engage in encoding of relational information can be assessed in terms of patterns of *divergence* or *convergence* of fixations to event characters (agents and patients) before speech onset: while encoding of individual characters at the outset of formulation is reflected in preferential and sustained fixations on a single character (see Gleitman et al. 2007), relational encoding should be indexed by distributed fixations *between* the two characters (as relational information is presumably 'distributed' between characters in an event; see Griffin and Bock 2000).

Importantly, the aim was to test how early an effect of word order would arise. If formulation is modulated by linguistic structure from the outset of formulation, then gaze patterns in verb-initial and subject-initial sentences should reflect word order differences immediately after picture onset (0–400 ms). If, by contrast, word order does not influence early formulation, then word order should only shape the distribution of fixations after 400 ms, that is, in time windows associated primarily with linguistic encoding.

The results supported the first possibility, both within Tzeltal and in comparisons between Tzeltal and Dutch: time-course analyses revealed effects of verb placement on the time-course of formulation from the earliest time-windows until articulation. In both Tzeltal and in Dutch, subject-initial sentences were formulated in a similar way to English sentences with the same word order (Gleitman et al. 2007; Griffin and Bock 2000; Kuchinsky and Bock 2010): formulation began with a rapid divergence of fixations to the two characters immediately after picture onset, and speakers continued fixating the first-mentioned character until speech onset. This cross-linguistic similarity in the formulation of subject-initial

[3]The following abbreviations are used: INC = incompletive aspect, 3S = third person singular, ERG = ergative, ABS = absolutive, CL = clitic.

sentences demonstrates that when the linear order of words in event descriptions is the same across languages, so is the time-course of formulation. The formulation of Tzeltal verb-initial sentences was markedly different: speakers distributed their gaze between agents and patients across a very broad window before speech onset. Formulation of active sentences began with a short-lived spike of fixations to the agent within 300 ms of picture onset and was followed by a convergence of fixations between agents and patients until speech onset. Formulation of passive sentences showed a similar pattern (although speakers generally preferred fixating the agent over the patient, consistent with the robust finding that speakers generally attend to agents more than patients; see Cohn and Paczynski 2013, for a review).

Taken together, these results demonstrate that from a very early stage of formulation, the word order that was under production strongly influenced how speakers constructed their sentences online. This suggests that, very rapidly after picture onset, speakers are able to generate a rudimentary structural plan and this plan can then guide subsequent conceptual and linguistic encoding operations. Supporting the conclusion of fast message-level encoding, a number of recent studies have demonstrated that very brief presentations (40–300 ms) of event pictures are in fact sufficient for speakers to identify event categories, as well as the role and identity of characters in the event (Dobel et al. 2007; Hafri et al. 2013). Identification of such information is presumably sufficient to allow for the generation of a conceptual and linguistic structural frame, which can direct the eye to efficiently sample information from the scene as the structure calls for it. Such tight parallels between fixation patterns and linguistic structure from the earliest stages of conceptual formulation, suggest, ultimately, that there may be no strict separation between processes related to conceptualization and those related to linguistic formulation.

5.3.4 Tagalog

An interesting question is whether the nature of early relational encoding for verb-initial structures also differs across languages as a function of the properties of verbs themselves. In Tzeltal, the extensive prioritising of early relational encoding may be driven not only by the verb's placement, but also by its complex morphology, which specifies information about both participants in the event. Further evidence supporting the possibility that verbal morphology can affect the early stages of formulation in verb-initial languages comes from Tagalog, an Austronesian language spoken in the Philippines by 23 million speakers.

The preferred word order for transitive sentences in Tagalog is also verb-initial. Interestingly, Tagalog requires marking the semantic role of one of the arguments (the privileged syntactic argument, or PSA) on the verb. This is illustrated in the examples below (two sentences describing an event where a child is kicking a ball). In [3], the PSA (marked with the prefix *ang*) is the actor (the agent in the event), so the verb takes 'actor voice' (AV) marking. In [4], where the PSA is the undergoer (the patient), the verb takes 'undergoer voice' (UV) marking. The order of postverbal

arguments is not fixed, thus either the PSA or the non-PSA argument may occur directly after the verb (here we only show examples with the PSA in final position):

[3]	*s<um>isipa*	*ng=bola*	*ang=bata*
	<AV>kick[4]	NPSA=ball	PSA=child
	predicate	undergoer	actor
	"The child kicks the ball."		

[4]	*s<in>ispa*	*ng=bata*	*ang=bola*
	<UV>kick	NPSA=child	PSA=ball
	predicate	actor	undergoer
	"The child kicks the ball."		

From the perspective of formulation, this system of verbal marking implies that speakers must engage in fairly extensive encoding of relational information early in the formulation process. Besides having to select a suitable verb to begin producing a sentence (as in Tzeltal), Tagolog speakers must make a commitment to treat one of the arguments of the verb as the PSA. Crucially, they must do this at the outset of formulation, even though they do not need to encode the PSA character itself linguistically and produce it until after they have produced the verb (i.e., typically after speech onset).

Sauppe et al. (2013) tested whether this grammatical requirement results in different gaze patterns during formulation of sentences that mark either the agent or the patient in the event as the PSA. Tagalog speakers performed an eye-tracked picture description task similar to the tasks described above in English, Dutch, and Tzeltal. Comparisons of fixation patterns to agents and patients across different sentence types showed that speakers briefly fixated the PSA character within an early time window (0–600 ms) and then fixated the two characters in the order of mention. As in other languages, fixations in the early time window can be seen as reflecting processes involved in the apprehension of the gist of the to-be-described event and generation of a structural framework, and fixations in later time windows index the order of lexical retrieval operations. Thus an effect of PSA marking on early eye movements suggests an early effect of linguistic structure on formulation (specifically, on processes at the interface of message formulation and linguistic encoding).

5.4 Conclusion

Production research carried out in the last several decades has shown that message and sentence formulation proceed incrementally. Importantly, however, it has recently become obvious that the timecourse of incremental message and sentence

[4]The following abbreviations are used: AV = actor voice, NPSA = non-privileged syntactic argument, PSA = privileged syntactic argument, UV = undergoer voice.

formulation is relatively flexible. While flexibility can be observed *within* languages under different conditions, the most striking demonstrations of flexibility are arguably provided by cross-linguistic comparisons. In this respect, the studies reviewed in this chapter have shown, first, that incrementality is a general principle of production that applies cross-linguistically and, second, that incremental encoding can be controlled by those aspects of the language that are responsible for linearization—i.e., grammar. In short, differences in language-specific grammatical constraints on word order result in differences in the order of encoding operations performed to produce grammatically correct utterances. These effects are observed in utterances ranging from simple noun phrases, which express relatively short messages, to full sentences, which can convey a wider range of complex concepts and relations.

As noted, most of the support for these conclusions comes from eye-tracking studies. The value of this methodology is that eye movements 'illustrate' incrementality in ways that other dependent measures cannot: speakers look at the referents they will describe, so eye movements can 'track' the incremental assembly of messages and sentences in real time. Crucially, the relationship between *looking* and *speaking* appears to be constant (or stable) across languages: speakers of different languages look at referents in a display in the order in which they will mention them (there is, quite strikingly, not yet any documented exception to this pattern). The stability of this relationship allows interpretations of differences in fixation patterns during sentence production in different languages in terms of underlying cross-linguistic differences in the time-course of encoding proper.

As such, eye-tracking as a methodology has provided important insight into a number of questions relevant for formulation. First, eye movements can reveal aspects of language that matter for formulation. The results show that cross-linguistic differences at both syntactic and morphosyntactic levels influence the time-course of formulation. Examples include syntactic features such as verb placement (as shown in Tzeltal) as well as morphosyntactic features such as language-specific agreement-marking on verbs (as shown in Tagalog). Second, eye movements can reveal *when* a particular linguistic feature influences formulation. Again, the results show that the timing of encoding depends on linear order: speakers prioritise encoding of the information that must be expressed next in the sentence shortly before it must be articulated.

The significance of cross-linguistic differences for theories of formulation depends on what *type* of information the target language requires speakers to express early on in a sentence. As described, an important contribution of eye-tracking to studies of formulation concerns the theoretical distinction between word-driven and structure-driven formulation. Since eye movements can be influenced both by lower-level factors like salience and accessibility (people look at objects or referents that attract their attention) and by higher-level factors like communicative goals (people look at objects or referents that they want to say something about), eye-tracking provides some of the most detailed data needed to distinguish between accounts assuming lexical guidance and accounts assuming structural guidance in formulation. Eye movements thus reveal whether speakers talk about referents that they happen to fixate first in a display or whether speakers look at referents in a display in a goal-directed fashion—or, for present purposes, in a manner reflecting guidance from a

linguistic framework. The results from numerous studies converge on a similar conclusion: language, or more specifically linguistic structure, can exert a strong influence on what speakers encode with priority at various points during formulation. If the information that receives priority is character-specific information (e.g., the subject of a sentence), the time-course of encoding resembles word-driven formulation; if languages require expressing relational information early on (e.g., the sentence verb or agreement marking on the verb), the time-course of encoding shows evidence of early relational encoding. Thus languages may inherently differ in the extent to which they 'allow' word-driven or structure-driven formulation.

At the same time, we note that these conclusions are based on a relatively small set of cross-linguistic comparisons. In principle, there are also other theoretical possibilities to consider, e.g. encoding patterns that lie somewhere between word-driven and structure-driven formulation. There may be as many different ways of assembling sentences incrementally as there are language-specific instantiations of linear ordering and dependency relations. Language-specific grammatical constraints may apply to encoding at different levels in the production system: some may have to do with higher-level conceptual groupings, while other may apply to smaller, local dependencies. As in the case of most experimental methods, eye-tracking has some limitations (see Irwin 2004) which may or may not make it suitable to study some of these fine-grained distinctions (the first of these limitations is that, of course, we have to use visual stimuli to elicit speech spontaneously). Nevertheless, there are still numerous targeted comparisons of individual language contrasts that can plausibly be addressed with this methodology.

With the advent of portable eye-trackers, eye-tracking research is no longer restricted to laboratories on university campuses. We now have the means to conduct eye-tracking studies in the field, with different groups of speakers and languages. The value of visual-world eye-tracking as a 'field-friendly' method has been demonstrated already through its successful application in a small, but diverse sample of languages and cultures (e.g. Huettig et al. 2011; Mishra et al. 2012; Norcliffe et al. in press; Sauppe et al. 2013). For language production research, this is an important advancement. It allows us to extend the typological base of vision-based research, in order to assess the extent to which our current models of formulation can adequately allow for language-specific instantiations of general principles. Current research has, of course, only scraped the surface of the astonishing range of linguistic diversity exhibited by the world's languages (Evans and Levinson 2009), so this remains an exciting empirical question.

References

Arnold, J. E., Wasow, T., Losongco, T., & Ginstrom, R. (2000). Heaviness vs. Newness: The effects of structural complexity and discourse status on constituent ordering. *Language, 76*, 28–55.

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

Bock, J. K., Irwin, D. E., & Davidson, D. J. J. (2004). Putting first things first. In F. Ferreira & M. Henderson (Eds.), *The integration of language, vision, and action: Eye movements and the visual world* (pp. 249–278). New York: Psychology Press.

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

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

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

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

Branigan, H. P., & Feleki, E. (1999). Conceptual accessibility and serial order in Greek speech production. *Proceedings of the 21st Cognitive Science Society Conference.* Vancouver.

Brown-Schmidt, S., & Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition, 109*, 274–280.

Caramazza, A., Miozzo, M., Costa, A., Schiller, N., & Alario, F.-X. (2001a). The gender congruity effect: Evidence from Spanish and Catalan. *Language and Cognitive Processes, 14*, 381–391.

Caramazza, A., Miozzo, M., Costa, A., Schiller, N., & Alario, F.-X. (2001b). A cross-linguistic investigation of determiner production. In E. Dupoux (Ed.), *Language, brain and cognitive development: Essays in Honor of Jacques Mehler* (pp. 209–226). Cambridge, MA: MIT Press.

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

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

Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative. *Cognitive Psychology, 67*, 73–97.

Costa, A., Sebastian-Galles, N., Miozzo, M., & Caramazza, A. (1999). The gender congruity effect: Evidence from Spanish and Catalan. *Language and Cognitive Processes, 14*, 381–391.

Cutler, A. (1985). Cross-language psycholinguistics. *Linguistics, 23*, 659–667.

Cutler, A., Mehler, H., Norris, D. G., & Segui, J. (1983). A language-specific comprehension strategy. *Nature, 204*, 159–160.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. J*ournal of Memory and Language*, *25*, 385–400.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Nature, 340*, 229–230.

Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly scene. *Acta Psychologica, 125*, 129–143.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*(5), 429–492.

Ferreira, F., & Swets, B. (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., & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research, 32*, 669–692.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language*. New York: McGraw-Hill.

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

Gennari, S. P., Mirković, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology, 65*, 141–176.

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language, 57*, 544–569.

Glucksberg, S., & Danks, J. H. (1975). *Experimental psycholinguistics: An introduction*. Hillsdale, N. J.: Erlbaum.

Griffin, Z. M. (2004). Why look? Reasons for eye movements related to language production. In J. Henderson & F. Ferreira, (Eds.), *The integration of language, vision, and action: Eye movements and the visual world* (pp. 213–247). New York: Taylor and Francis.

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

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experiment Psychology: General, 142*, 880–905.

Henderson, J. M., & Ferreira, F. (Eds.). (2004). *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Huettig, F., Singh, N., & Mishra, R. K. (2011). Language-mediated visual orienting behavior in low and high literates. *Frontiers in Psychology, 2*, 285.

Hwang, H., & Kaiser, E. (2009). *The effects of lexical vs. perceptual primes on sentence production in Korean: An on-line investigation of event apprehension and sentence formulation*. Paper presented at the 22nd CUNY Conference on Human Sentence Processing, Davis, CA.

Hwang, H., & Kaiser, E. (2014). The role of the verb in grammatical function assignment in English and Korean. *Journal of Experimental Psychology: Learning, Memory and Cognition*. June 2 (Epub ahead of print).

Ibbotson, P., Lieven, E., & Tomasello, M. (2013). The attention-grammar interface: Eye-gaze cues structural choice in children and adults. *Cognitive Linguistics, 24*.

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 105–133). New York, NY: Psychology Press.

Jaeger, T. F., & Norcliffe, E. (2009). The cross-linguistic study of sentence production. *Language and Linguistic Compass, 3*(4), 866–887.

Jaeger, T. F. & Snider, N. (2013) Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*, 57–83.

Janssen, N., Alario, F.-X., & Caramazza, A. (2008). A word-order constraint on phonological activation. *Psychological Science, 19*, 216–220.

Kempen, G., & Harbusch, K. (2004). A Corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In T. Pechmann (Ed.), *Language Production* (pp. 173–181). Berlin: Mouton.

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

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language, 66*, 143–162.

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology, 73*, 1–40.

Kuchinsky, S. E., & Bock, K. (2010). *From seeing to saying: Perceiving, planning, producing.* Paper presented at the 23rd meeting of the CUNY Human Sentence Processing Conference, New York, NY.

Lee, E. K., Brown-Schmidt, S., & Watson, D. W. (2013). Ways of looking ahead: Incrementality in language production. *Cognition*, *129*, 544–562.

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

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 1–16.

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

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

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

Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20*, 298–305.

Mehler, J., Dupoux, E., Nazzi, T., & Dehaene-Lambertz, G. (1996). Coping with linguistic diversity: The infant's viewpoint. In J. L. Morgan & K. Demuth, et al. (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 101–116). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Mehler, J., Sebastian, N., Altmann, G., Dupoux, E., Christophe, A., & Pallier, C. (1993). Understanding compressed sentences: The role of rhythm and meaning. In P. Tallal, A. M. Galaburda et al. (Eds.), *Temporal information processing in the nervous system: Special reference to dyslexia and dysphasia. Annals of the New York Academy of Sciences* (Vol. 682, pp. 272–282). New York, NY: New York Academy of Sciences.

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

Miozzo, M., & Caramazza, A. (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*, 907–922.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, *5*(1), 3, 1–10.

Myachykov, A., & Tomlin, R. S. (2008). Perceptual priming and structural choice in Russian sentence production. *Journal of Cognitive Science, 6*, 31–48.

Myachykov, A., Garrod, S., & Scheepers, C. (2010). Perceptual priming of structural choice during english and finnish sentence production. In R. K. Mishra & N. Srinivasan (Eds.), *Language & cognition: State of the art* (pp. 54–72). Munich: Lincom Europa.

Myachykov, A., Thompson, D., Scheepers, C., & Garrod, S. (2011). Visual attention and structural choice in sentence production across languages. *Language and Linguistic Compass*, *5*, 95–107.

Norcliffe, E., Konopka, A. E., Brown, P., & Levinson S. C. (in press). Word order affects the time-course of sentence formulation in Tzeltal. *Language, Cognition, and Neuroscience.*

Paul, H. (1886/1970). The sentence as the expression of the combination of several ideas. In A. L. Blumenthal (Ed. & Trans.), Language and psychology: Historical aspects of psycholinguistics (pp. 34–37). New York: Wiley.

Polian, G. (2013). *Gramática del tseltal de Oxchuc*. Mexico, D.F.: Centro de Investigaciones y Estudios Superiores en Antropología Social.

Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin, R. D. Jr., & Levinson, S. C. (2013). Dependencies first: Eye-tracking evidence from sentence production in Tagalog. In: M. Knauff, M. Pauen, N. Sebanz, & E. Wachsmuth (Eds.), *Proceedings of the 35th annual meeting of the Cognitive Science Society* (pp. 1265–1270). Austin, Texas: Cognitive Science Society.

Schiller, N., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language, 48*, 169–194.

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 841–850.

Swets, B., Jacovina, M. E., & Gerrig, R. J. (2008). *Individual differences in the planning scope of language production*. Paper presented at the 49th meeting of the Psychonomic Society, Chicago, IL (November).

Tanaka, M. N., Branigan, H. P., McLean, J. F., & Pickering, M. J. (2011). Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language, 65*, 318–330.

Tomlin, R. (1995). Focal attention, voice, and word order: An experimental, cross-linguistic study. In M. Noonan & P. Downing (Eds.), *Word order in discourse* (pp. 521–558). Amsterdam: John Benjamins.

Tomlin, R. (1997). Mapping conceptual representations into linguistic representations: the role of attention in grammar. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 162–189). Cambridge: Cambridge University Press.

Van de Velde, M., Meyer, A. S., & Konopka, A. E. (2014). Message formulation and structural assembly: Describing "easy" and "hard" events with preferred and dispreferred structures. *Journal of Memory and Language, 71*, 124–144.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experiment Psychology: Learning, Memory, and Cognition, 36*, 423–440.

Wundt, W. (1900). *Völkerpsychologie: Eine Untersuchung der Entwicklungsgesetze von Sprache, Mythus und Sitte* (Vol. 1). *Die Sprache* [Language]. Leipzig: Kroner-Engelmann.