7 How to Use Psycholinguistic Methodologies for Comprehension and Production

Kim McDonough and Pavel Trofimovich

Background

Psycholinguistics is the study of the psychological processes involved in language, and its main goals are to understand how people comprehend and produce language. For a skilled language user, understanding and producing language seem deceptively simple. For a psycholinguist, however, language comprehension and production involve a complex interaction of various processing components, which include accessing the lexicon, building a syntactic structure, and encoding and decoding the sound patterns of a language, as well as interpreting and expressing intended pragmatic messages. By studying these various components, psycholinguists attempt to figure out what processes, mechanisms, or procedures underlie language use and learning.

Psycholinguists typically study language comprehension and production as separate sets of processes. One reason for this is that comprehension and production both pose distinct challenges to language users. For example, language comprehension involves extracting meaning from a speech signal or printed text whereas language production involves converting a preverbal message into speech or text using appropriate lexicon, grammar, and phonology or orthography. Another reason is that compared to comprehension, production appears to be much harder to study experimentally. This is because researchers often find it difficult to control input and elicit relevant output when studying language production. Indeed, it is much easier to manipulate words and sentences to be read or heard than to control the ideas and means of expression used for speaking and writing. It is not surprising, then, that even after decades of psycholinguistic research, comprehension tends to be studied more extensively than production (see Bock, 1996, for additional reasons).

Research Methods in Second Language Acquisition: A Practical Guide, First Edition.
Edited by Alison Mackey and Susan M. Gass.

Psycholinguists have developed a number of ingenious methods for studying how people comprehend and produce language. The goal of this chapter is to describe four of these methods, with a particular focus on their use in second language (L2) research. The methods we highlight have been used extensively in research with first language (L1) speakers. They allow for the study of language comprehension and production at different levels of analysis (words, sentences, longer discourse), and they are adaptable to L2 research contexts. In discussing each method, we first describe it, then contextualize it within its relevant theoretical framework, and finally raise some methodological considerations. Because space limitations do not allow for a complete inventory of psycholinguistic methods, we refer interested readers to other sources listed under 'Further Reading' near the end of the chapter.

Language Comprehension

Self-Paced Reading Task

Description of the task

In a self-paced reading task, participants are seated at computers with programs such as E-prime, PsyScope, or DMDX and read text from a computer screen. The text is segmented into words or short phrases, and participants press a keyboard key or a response button to display each consecutive segment until they reach the end. For example, the sentence *The bad boys | watched almost every day | were playing | in the park* could be presented to participants in four consecutive segments marked by vertical bars. This sentence, taken from Juff's (1998) study, illustrates that readers can encounter temporary ambiguity in interpreting some structures. The second segment above could be interpreted as both a reduced relative clause (with the relative pronoun and auxiliary verb omitted) or as a main clause with a past tense verb. The third segment is critical because it signals that *were playing* is the main verb and that *watched almost every day* is a reduced relative clause. The logic of self-paced reading is that the time needed to read each segment reflects the cognitive workload experienced in processing that segment. Thus, the measure of interest in self-paced reading is the amount of time between two successive button presses, especially for the ambiguous segments.

Several versions of the task have been developed. These versions are illustrated in figure 7.1. In the "stationary-window" version of the task (row A), each subsequent segment replaces the preceding segment, usually in the center of the screen. In the "moving-window" version (rows B and C), the text is masked by a pattern of dashes and spaces, and each consecutive segment appears on the screen sequentially, replacing the dashes. The "moving-window" version can be used with a cumulative or a non-cumulative presentation. The cumulative presentation (row B) reveals increasingly more segments with the previous segments remaining on screen. The non-cumulative presentation (row C) presents a subsequent segment but removes or covers up the previous segments. The cumulative moving-window presentation is less preferred because participants tend to reveal multiple segments in order to read them all at once.

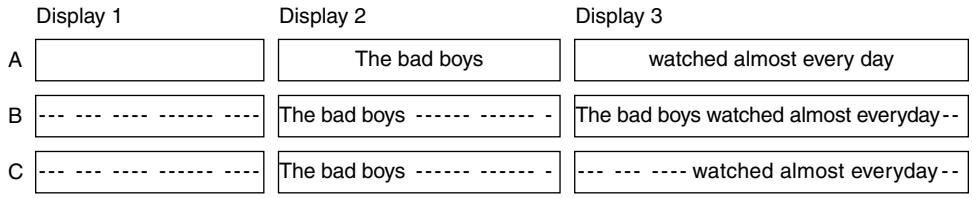

Figure 7.1 A schematic illustration of the first three display screens for the sentence *The bad boys watched almost every day were playing in the park* in stationary-window (A), cumulative moving-window (B), and non-cumulative moving-window (C) versions of the task.

Theoretical frameworks

L1 comprehension Self-paced reading is typically used in L1 research that investigates syntactic parsing (or simply parsing), which is the process by which people analyze the syntactic structure of a string of words to arrive at a correct interpretation. Skilled language users usually process syntactic information in print and speech rapidly, encountering little comprehension difficulty. Therefore, researchers use sentences that create difficulties for comprehension, such as "garden-path" sentences that are initially misinterpreted and must be reanalyzed. Structures typically manipulated in garden-path sentences include reduced relative clause–main verb ambiguity, relative clause attachment (Dussias, 2003), subject–object ambiguity (Hopp, 2006), and *wh*-extraction (Jackson & Bobb, 2009).

Because the comprehension process involves several levels of processing (e.g., grapheme/phoneme, word, sentence, text) and includes several subprocesses (van Gompel & Pickering, 2007), two key theoretical questions have been raised. Do people entertain only one interpretation of something they read or listen to, or do they consider several plausible interpretations simultaneously? How do people use the many sources of information available to them (e.g., syntax, semantics, discourse, prosody, etc.)? Self-paced reading has been used extensively to address these two questions, with a focus on the theoretical debate between interactive and modular theories of sentence comprehension. Interactive theories (constraint-based or one-stage theories) assume that people entertain several sources of information simultaneously (e.g., semantics, discourse constraints, frequency of use) in order to interpret a syntactic structure (e.g., Trueswell & Tanenhaus, 1994). In contrast, modular theories (two-stage theories) posit that people use only syntactic information to create an initial interpretation of a sentence and that other sources of information (e.g., semantics, context) are used later in the comprehension process (e.g., Frazier, 1990). This theoretical debate continues due to strong evidence in support of both views (see van Gompel & Pickering, 2007).

Bilingual/L2 comprehension Current bilingual and L2 research investigates the parsing strategies employed by L2 users. The central issue is whether L2 users develop syntactic parsing strategies that are specific to the L2 or whether they use L1-based parsing strategies, or perhaps strategies that are found in neither language (see Dussias & Piñar, 2009, for a recent summary). Briefly, the syntactic parsing of

L2 users is influenced by many factors. Some of these factors are linguistic in nature (i.e., specific to the lexical, semantic, or syntactic information available in sentences). Other factors are related to L2 users' experiences with a language (i.e., proficiency or type of exposure) and to individual differences (e.g., working memory). It appears that many of these factors interact in determining precisely how L2 users' parsing strategies in the L2 are similar to or different from those used in the L1. A sample self-paced reading study that investigates syntactic parsing in L2 users is summarized in study box 7.1.

Study Box 7.1

Rah, A., & Adone, D. (2010). Processing of the reduced relative clause versus main verb ambiguity in L2 learners at different proficiency levels. *Studies in Second Language Acquisition*, *32*, 79–109.

Purpose

Examine how L2 learners of different proficiency levels process the main verb versus reduced relative clause ambiguity (*The bad boys watched almost every day were playing in the park*).

Participants

22 intermediate- and 22 advanced-level German learners of English, all students at a German university; 22 native English speakers.

Task design

The 30 target sentences included (1) unambiguous sentences, (2) ambiguous sentences with a good cue that ruled out a transitive reading of the sentence, and (3) ambiguous sentences with a poor cue, which left the transitive interpretation plausible until the disambiguating region (italicized).

(1) The brown sparrow | seen | by the hungry cat | *pecked* | at an insect.
(2) The brown sparrow | noticed | on an upper branch | *pecked* | at an insect.
(3) The brown sparrow | noticed | almost every day | *pecked* | at an insect.

Procedure

A stationary (non-cumulative) centered self-paced presentation was used; 10 sentences were followed by a comprehension question.

Analysis

Data analyzed using 2 × 3 × 3 (group × sentence type × region) ANOVA followed by univariate ANOVAs and pairwise comparisons.

Results

Reading patterns for the learners and the native speakers were similar. The learners of different proficiency levels differed in overall speed of reading, but not in pattern of processing.

Methodological considerations

There are several methodological considerations to keep in mind when designing self-paced reading experiments. Because self-paced reading is used to test claims about comprehension processes, the task must encourage participants to comprehend the sentences. To accomplish this, researchers typically ask participants to answer a comprehension question (Rah & Adone, 2010), judge grammaticality (Juffs, 1998), decide whether another sentence conveys the same meaning (Hopp, 2006), or make a plausibility judgment (Williams, Möbius, & Kim, 2001). Usually researchers only analyze the reading times for the sentences which are responded to correctly. Before analyzing the data, researchers also often try to confirm that all participant groups are matched for comprehension rates on these tasks. This helps ensure that potential differences in reading times can be attributed to the kinds of materials being manipulated rather than to participants' processing strategies (e.g., reading for comprehension vs. skimming the text).

Another consideration involves the need to create target and baseline materials which are maximally comparable so that participants' performance can be interpreted appropriately. For example, it is important to match the target and baseline materials for word length and for lexical content, especially before the disambiguating segment, or the region where researchers expect to find reading time differences. This ensures that potential differences in reading times are not due to different lexical content across sentences. It is also important to match the length of disambiguating segments across all materials in terms of total character length and frequency. This helps ensure that these segments are equally "salient," given the fact that shorter and more frequent words are often read quickly or even skipped over in reading (see Ferreira & Clifton, 1986, for a statistical procedure to adjust reading times for segment length). Finally, the disambiguating segments should not be placed at the ends of sentences. This helps minimize sentence wrap-up effects, or the tendency for reading times to be slower and more variable at sentence and clause boundaries.

Researchers wishing to use self-paced reading to study language comprehension need to be aware of certain limitations. One concern is that self-paced reading is prone to spillover effects. This refers to the tendency for processing effects to "carry over" from one segment to the next. For example, when researchers present readers with the sentence *The bad boys* | *watched almost every day* | *were playing* | *in the park*, they expect to find longer reading times for the third segment (compared to the same segment in a baseline condition) because this segment will signal to readers that they have been led on a "garden path" and will need to reinterpret the sentence.

However, due to spillover effects, longer reading times associated with this segment could manifest themselves much later, for example, when readers process one of the following segments. To minimize spillover effects, researchers should avoid using word-by-word presentations (e.g., *The* | *bad* | *boys* | etc.) and should carefully match materials for number of words, especially for segments preceding the disambiguating segment (Mitchell, 1984). Another concern about self-paced reading is that processing might be affected by a particular text segmentation strategy. In other words, it is possible that researchers could find different results depending on how they segment their materials (*The bad boys* | *watched almost every day* | *were playing* | *in the park* vs. *The bad boys* | *watched* | *almost every day* | *were* | *playing* | *in the park*). Although "there is no solid evidence that researchers have ever been misled by segmentation biases in the self-paced reading task" (Mitchell, 2004, p. 26), researchers may wish to replicate their findings using a number of different segmentation strategies.

Self-Paced Listening Task

Description of the task

Self-paced listening (or the auditory moving-window technique) is the auditory equivalent of self-paced reading. Compared with self-paced reading, this task is relatively new. It was first described by Ferreira, Henderson, Anes, Weeks, and McFarlane (1996) in a study of lexical and syntactic processing with L1 speakers. In a self-paced listening task, listeners hear sentence segments presented one at a time, pressing a button to play the next segment. For example, ambiguous sentences like *When* | *Roger* | *leaves* | *the house* | *is* | *dark* and *When* | *Roger* | *leaves* | *the house* | *it's* | *dark* could be played to participants in several separate segments, demarcated here by vertical bars (Titone et al., 2006). The end of a sentence is usually marked by a tone, which is often followed by a comprehension question (or a related task) which ensures that participants process each sentence. The logic behind this task is identical to that of self-paced reading: the time participants take to listen to each segment reflects the processing load (and therefore, the processing difficulty) they experience in comprehending that segment. At least two temporal variables can be measured in this task: (a) inter-response time, or the time between each consecutive press of the button (Ferreira et al., 1996) and (b) pause duration, or the time between the offset of one segment and participants' press of the button to initiate the next one (Titone et al., 2006).

Theoretical frameworks

L1 comprehension In L1 research, self-paced listening has been used to address the same topics examined with self-paced reading tasks, such as how listeners interpret the structure of spoken utterances and what information they use to do so. For example, Ferreira et al. (1996) showed that listeners, just like readers, take longer to

process syntactically ambiguous than unambiguous sentences. In addition, these researchers also showed that sentence prosody (metrical and intonational structure of utterances) influences syntactic processing, such that pitch contours and pausing that are congruent with the syntactic structure make ambiguous sentences easier to process. This finding suggests that listeners use prosody (along with other lexical and contextual information) to comprehend spoken sentences.

Self-paced listening has been used in L1 research to compare child and adult processing of syntactically complex and simple utterances (Booth, MacWhinney, & Harasaki, 2000), compare younger and older adults' processing of spoken passages (Titone, Prentice, & Wingfield, 2000), test the relationship between speech processing and working memory (Waters & Caplan, 2004), and investigate the processing of spoken sentences by individuals with brain damage (Caplan & Waters, 2003) and children with language impairments (Marshall, Marinis, & van der Lely, 2007). In contrast, relatively little bilingual and L2 research to date has used self-paced listening, which we believe is a reflection of L2 researchers' relatively low familiarity with this task. We are aware of only one published study (de Jong, 2005) and a few conference presentations (e.g., Heredia, Stewart, & Cregut, 1997) that have used self-paced listening. Clearly, the advantages and limitations of this method need to be explored in future L2 research. In study box 7.2, a sample L1 self-paced listening study is summarized.

Study Box 7.2

Titone, D. A., Koh, C. K., Kjelgaard, M. M., Bruce, S., Speer, S. A., & Wingfield, A. (2006). Age-related impairments in the revision of syntactic effects of prosody. *Language and Speech*, *49*, 75–99 (Experiment 1).

Purpose

Investigate the role of prosody in younger and older adults' processing of syntactically ambiguous sentences.

Participants

24 younger (aged 18–28) and 24 older (aged 65–83) adults, all native speakers of English.

Task design

18 syntactically ambiguous sentence pairs patterning according to "early" and "late closure" (see example below), recorded in three prosodic conditions: when sentence contained no prosodic information (baseline prosody), when prosodic and syntactic boundaries coincided (cooperating prosody), and when prosodic and syntactic boundaries conflicted (conflicting prosody). | = segmentation locations, * = prosodic boundary, *italics* = ambiguous region.

Baseline prosody	
Late closure	When | Roger | *leaves* | *the house* | it's | dark
Early closure	When | Roger | *leaves* | *the house* | is | dark
Cooperating prosody	
Late closure	When | Roger | *leaves* | *the house** | it's | dark
Early closure	When | Roger | *leaves** | *the house* | is | dark
Conflicting prosody	
Late closure	When | Roger | *leaves** | *the house* | it's | dark
Early closure	When | Roger | *leaves* | *the house** | is | dark

Procedure

Participants listened to a sentence in a self-paced listening task, then repeated or paraphrased that sentence (to ensure that they processed it for comprehension and resolved syntactic ambiguity). Pause duration was recorded in the disambiguating segment *it's/is*.

Analysis

Data analyzed using a 2 × 2 × 3 (age × syntax × prosody) ANOVA, followed by planned pairwise comparisons.

Results

Younger and older adults were similar in their interpretation of sentences and in their use of prosody for sentence comprehension. However, younger adults appeared faster than older adults at using congruent prosodic patterns for sentence comprehension.

Methodological considerations

Because self-paced listening is modeled after self-paced reading, the methodological issues described in the previous section are important. However, additional issues stemming from the auditory nature of the task should also be addressed. One issue relates to preparation of task materials. Unlike printed text, speech does not have clear-cut boundaries between adjacent words, with sounds produced in an overlapping fashion. This makes it hard for researchers to divide a continuous speech stream into segments. To address this problem, some researchers record each individual word (or segment) separately and present their materials as sequences of individually recorded segments (Booth et al., 2000). Other researchers segment a continuous stream of speech, trying to ensure that the transitions between adjacent segments are smooth and that words at segment boundaries are intelligible (Ferreira et al., 1996).

Yet other researchers, especially those interested in processing of longer texts, segment their recordings where naturally occurring pauses are present, for example, at clause or sentence boundaries (Titone et al., 2000).

A related issue concerns the role of prosody in comprehension of speech. In spoken language, variations in pitch, pausing, and duration all signal important information about the structural configuration of an utterance. Put differently, prosodic and syntactic patterns of spoken utterances are normally closely aligned, and listeners rely on both syntax and prosody in comprehension of speech. This relationship between prosody and syntax could introduce an unwanted confound in a self-paced listening study if prosodic factors are not the focus of investigation. There are several ways of minimizing the unwanted effects of prosody. One solution is to record each word (or segment) of an utterance in isolation and out of sequence, for example, by reading them from a list. Sentences recorded in this fashion lack coarticulatory and prosodic cues (Booth et al., 2000). Another solution is to use a digital splicing technique whereby researchers take a single prosodic pattern (e.g., from a neutral, baseline utterance) and splice it into their target materials using speech editing software (Felser, Marinis, & Clahsen, 2003). This allows researchers to create materials that are comparable in their prosodic content across conditions (see Ferreira, Anes, & Horine, 1996, on prosody in self-paced listening).

Another consideration to take into account is the choice of the dependent measure. As noted earlier, researchers typically collect two principal measurements in self-paced listening. Using inter-reponse times as the primary measure in self-paced listening could be problematic if the segments of interest are not matched for duration across conditions. Because longer segments are processed more slowly than shorter segments, longer inter-response time may simply reflect processing time rather than comprehension difficulty. Pause duration (which is equivalent to the "difference time" measure proposed by Ferreira et al., 1996) at least in part avoids this problem because this measure excludes the duration of the segment.

Language Production

Picture–Word Interference Task

Description of the task

The picture–word interference task is often described as a Stroop-like task because it similarly directs speakers to articulate words while ignoring distracting information (see Dell'Acqua, Job, Peressotti, & Pascali, 2007, and van Maanen, van Rijn, & Borst, 2009, for similarities and differences between these tasks). In the Stroop task, speakers articulate the font color while ignoring the name of the color written in orthography (e.g., *green* written in purple font). In the picture–word interference task, speakers name objects depicted through pictures while ignoring aural or visual distracters (e.g., naming the picture of a dog while seeing the word *cat*). The picture–word interference

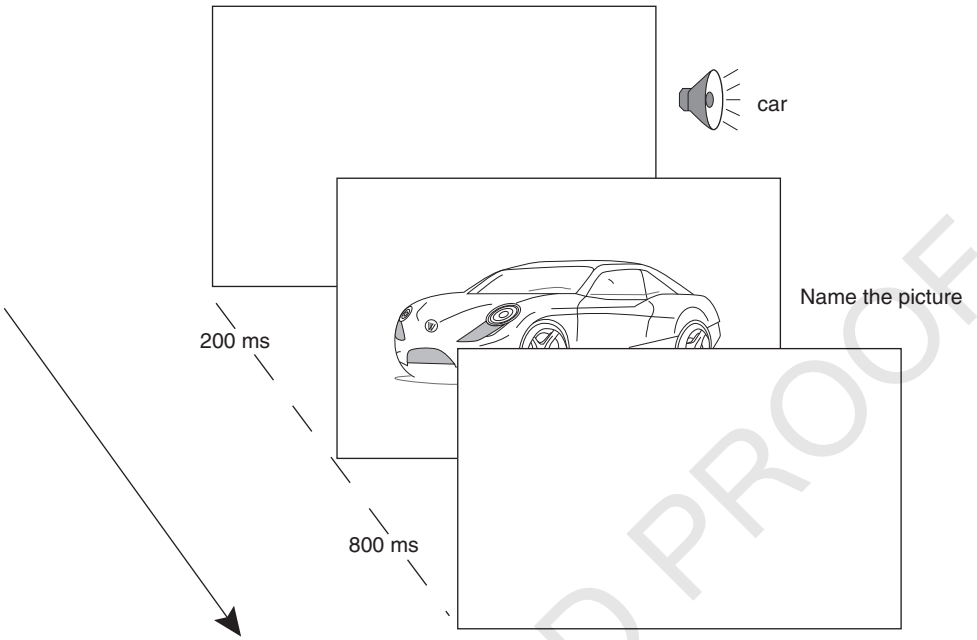

Figure 7.2 A schematic illustration of the picture-word interference task (based on Hantsch, Jescheniak, & Schriefers, 2009). This illustration depicts the following sequence of events for each trial: (1) a semantically related auditory distracter *car* (the picture's basic-level name) is presented 200 ms before the picture, (2) the target picture (Porsche) is presented for 800 ms, and (3) the participant names the picture (i.e., says *Porsche*) as quickly as possible.

task elicits reaction times in the form of naming latencies, which are believed to provide insight into the nature of lexical retrieval. The logic behind the task is that the distracters activate their corresponding lexical representations and this interferes with picture naming, making it slower and more difficult. Two variables are measured in this task: (a) naming latency, which is the amount of time required for speakers to name the picture, and (b) accuracy, which is whether the object names accurately represent the illustrations. Slower naming latencies (compared to a baseline condition) indicate that the distracters interfere with lexical selection of the target words, while faster naming latencies indicate that the distracters facilitate selection.

To carry out the picture–word interference task, individual speakers are seated at computers with programs such as E-prime, PsyScope, or DMDX and high-quality voice-activated recorders. The typical presentation of stimuli involves a fixation point which remains on screen briefly (e.g., 500 ms), after which a blank screen appears (e.g., 500 ms). Then the picture appears on screen with the distracter word written on or near the picture. Alternatively, distracters may be presented aurally instead of visually. Depending on the focus of the experiment, the distracter may be presented with the onset of the picture, shortly before, or shortly after the picture. The target picture remains on screen for about 2 seconds and the trial terminates immediately after the naming response is provided or after the 2 seconds elapse. There is usually a 1-second interval between items. A sample procedure for a picture–word interference task is illustrated in figure 7.2.

Theoretical frameworks

L1 production The picture–word interference task has been used in L1 speech production research to test claims about lexical selection. Contemporary speech production models are in agreement that the basic architecture of the production system consists of three stages (for an overview, see Griffin & Ferreira, 2006). These include conceptualization (preverbal message), formulation (lexical selection, grammatical processing, and phonological assembly), and articulation (physical articulation of overt speech). However, current models diverge in terms of specific claims about (a) the process of lexical selection (competition during activation or selection), (b) the association of grammatical properties with lexical selection (automatic or subsequent), and (c) the impact of phonological encoding on lexical selection (feed-forward or cascading).

Picture–word interference tasks have been used to investigate the timing of the lexical and phonological assembly phases within the formulation stage. These studies have shown that compared to unrelated distracters, semantically related distracters lead to slower naming latencies, which is the semantic interference effect (e.g., Hantsch et al., 2009). For example, hearing or seeing the word *cat* at the same time as or 400 ms before a picture of a dog is presented slows picture naming when compared to unrelated distracters, such as *table*. In contrast, phonologically similar but semantically unrelated distracters facilitate naming latencies when they are presented after the object to be named, which is the phonological facilitation effect (e.g., Bi, Xu, & Caramazza, 2009; Zhang, Chen, Weekes, & Yang, 2009). For example, hearing or reading the word *doll* milliseconds after a picture of a dog facilitates naming latencies compared to a phonologically unrelated distracter, such as *book*.

Bilingual/L2 production One main question for bilingual speech production is the extent to which the language not in use, referred to as the non-response language, affects production in the response language. For example, if a Spanish–English bilingual is asked to name a picture of a dog in English, do the corresponding lexical items in each language, dog and perro, become activated? Contemporary models of bilingual speech production assume that the conceptual system is shared by the two languages and that activation from the conceptual system is not specific to the response language. This implies that the relevant lexical items in both languages get activated (for review, see Costa, 2004). To explain how bilinguals manage to produce the appropriate lexical item in the appropriate language, one account posits an inhibitory process that suppresses activation of the words in the non-response language, while alternative accounts assume the existence of a selection mechanism that ignores activation of words in the non-response language.

Picture–word interference tasks have been used to determine whether activation of the non-response language interferes with lexical selection in the response language. Costa and his colleagues have shown that distracters that are translation equivalents of the target words facilitate naming latencies compared to unrelated distracters (Costa & Caramazza, 1999; Costa, Miozzo, & Caramazza, 1999). For example, a Spanish–English bilingual would name a picture of a dog in Spanish (would say *perro*) faster if the distracter were the English word *dog* as opposed to a non-translation

equivalent, such as *table*. These results have been used to support the lexical selection mechanism, as opposed to inhibitory processes. Costa and his colleagues have also shown that non-target-language distracters that are phonologically related to the targets can impact naming latencies (Costa et al., 1999), which suggests that both lexical and phonological activation of the non-target language occurs. A sample study using a modification of the picture–word interference task, in which the distracters are pictures rather than words, is summarized in study box 7.3.

Study Box 7.3

Colomé, A., & Miozzo, M. (2010). Which words are activated during bilingual word production? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 96–109 (Experiment 2).

Purpose

Determine whether distracters that are phonologically related to the non-response language affect naming latencies in the response language. It was predicted that phonologically related distracters would facilitate naming latencies.

Participants

24 Catalan–Spanish bilinguals with high proficiency in both languages who acquired their second language before age six.

Task design

Target pictures were 21 objects to be named in Catalan. They were paired with a distracter picture that was a cross-language similar name or a phonologically and semantically unrelated name. The target pictures were green, while the distracter pictures were red. Composite images were created so that the target picture and distracter picture were partially overlapping. For example, a picture of a vest to be named in Catalan (*armilla)* was paired with a picture of a squirrel, which is *ardilla* in Spanish, or a picture of a beak, which is *pico* in Spanish. None of the target and distracter picture pairs was semantically related to any other. The 21 target pictures were also paired with another set of semantically and phonological unrelated distracters that served as fillers and were not analyzed. The target pictures were presented in three blocks, with each block consisting of six pictures from each distracter type (phonologically related, unrelated, or fillers).

Procedure

Individual testing in soundproof booths at a computer with voice key recording. Preliminary activities included a familiarization phase for the pictures and

object names with feedback, practice picture naming, instructions to name the green picture while ignoring the red picture, and practice items. Each trial began with a fixation point (500 ms), followed by a blank screen (500 ms) and the composite picture (400 ms). Recording of response latencies began when the picture appeared, and continued for up to 2 seconds with a 1-second blank interval separating the trials.

Analysis

No responses were discarded for recording failures or fast responses. Responses in which participants produced unintended names or had verbal dysfluencies (stuttering, repairs, non-verbal sounds) were coded as errors (3.2%) and entered into error analyses, which revealed no significant results.

Results

A t-test indicated that naming latencies were significantly faster for pictures that had related versus non-related distracters.

Methodological considerations

Because the picture–word interference task uses naming latencies and accuracy rates as evidence of interference or facilitation, the experimental materials manipulate the semantic and phonological relationships between the distracters and the objects to be named. The distracters should have varying relationships with the objects to be named, such as semantically related, phonologically related, or unrelated controls, and should be matched for length and frequency. When cognate distracters are used in bilingual experiments, they should share a significantly larger number of phonemes in identical word position than non-cognates. Most studies typically have small object sets and use a single object with multiple distracters, which can artificially decrease picture naming latency due to repetition practice. This potential problem can be avoided by using a unique picture for each trial (Knupsky & Amrhein, 2007), which increases the breadth of the stimuli and improves the ecological validity of the task.

In terms of the location of the distracter word, it is usually presented in the center of the image. However, its position has been manipulated in order to disambiguate recognition of the distracter, as reflected in eye movements to it, from interference of the distracter, as reflected by naming latencies (Kaushanskaya & Marian, 2007). Another modification is to present two images in the visual array and then present the distracters aurally. This allows for greater testing of relationships as some distracters can relate to the target picture, while others can be associated with the picture to be ignored (Oppermann, Jescheniak, & Schriefers, 2008). And in the picture–picture interference task, one picture is the distracter while the other picture is the object to be named. In this version of the task, color

is typically used to indicate which picture should be named (Navarrete & Costa, 2005; Colomé & Miozzo, 2010).

All participants must be familiar with the pictures and the objects to be named prior to carrying out the experimental tasks. During the familiarization phase, feedback is generally provided to ensure that the participants recognize the objects in the pictures and use the appropriate words to name them. For example, the familiarization phase would help ensure that participants name the object previously shown in figure 7.2 using the word *Porsche* instead of *car*. Practice trials help ensure that the participants understand the instructions, which is particularly important in the picture–picture interference task where color indicates which picture should be named (e.g., "Name the object in the green picture but ignore the red picture"). In terms of analyzing responses, not only unintended object names but also no responses, self-corrections, or dysfluencies are typically coded as errors, but responses that were not detected due to technical problems are generally excluded. Outliers are responses that deviate more than two standard deviations from a mean (the participant or item mean), and those scores are either excluded or replaced with the cell mean.

Sentence Preamble Task

Description of the task

The sentence preamble task is most often used to elicit errors in subject–verb number agreement, although pronoun agreement has also been tested with this task. Speakers are given a phrase that serves as the subject of a sentence and are asked to generate a complete sentence. The subject phrases, or sentence preambles, contain two noun phrases. The first noun phrase is the grammatical subject of the sentence, and the second noun phrase is embedded in a prepositional phrase or a relative clause. The grammatical number of the first noun (the head noun) and the second noun (the local noun) are manipulated in order to increase or reduce number conflict. For example, the preamble *the key to the cabinets* has number conflict between *key* and *cabinets*, while the preamble *the key to the cabinet* does not. Speakers are asked to generate full sentences using the preambles as the subjects (e.g., *The key to the cabinets is in my desk drawer*), and their predicates are analyzed for subject–verb agreement errors. The logic of the task is that agreement errors should be greater when the head and local noun phrases mismatch in number as opposed to when they match.

The sentence preamble task (illustrated schematically in figure 7.3) is presented through experimental software such as PsyScope or DMDX with high-quality voice-activated recorders. The typical presentation of stimuli begins with a fixation point, such as a dot, an X, or a cross, which remains on screen for approximately 800 ms. Next, the preamble is presented either aurally or visually, with on-screen presentation time for visual preambles adjusted by their length. When the preamble ends, a visual cue appears, such as an exclamation point, at which point speakers repeat and complete the preamble. If the experimental design involves manipulation of semantic relationships or plausibility, then the task is adjusted by specifying a word to be

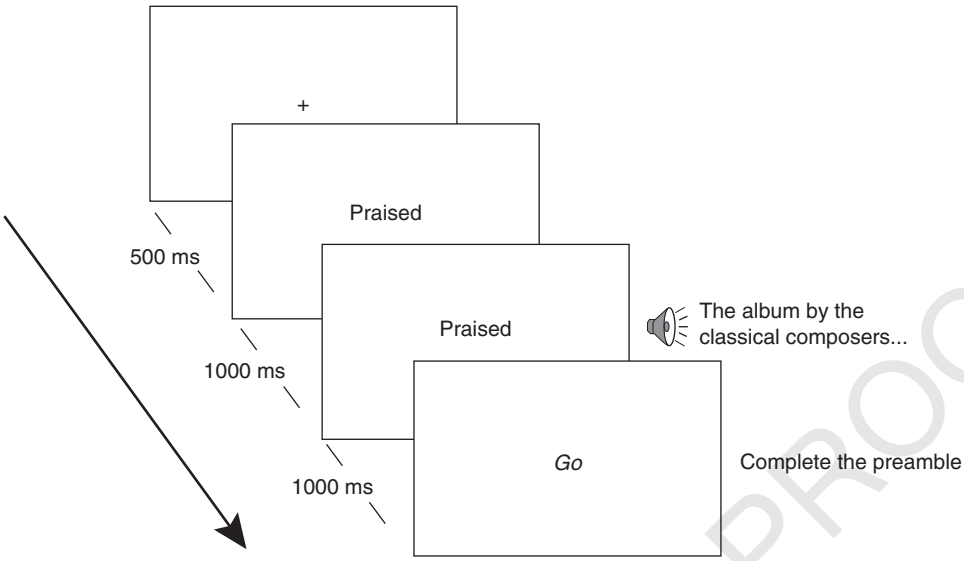

Figure 7.3 A schematic illustration of the sentence preamble task (based on Thornton & MacDonald, 2003). This illustration depicts the following sequence of events for each trial: (1) a fixation point (+) is presented; (2) the verb to be used in preamble completion is shown after 500 ms; (3) a preamble is played after 1000 ms, and the verb stays on screen for the duration of the preamble; (4) after a 1000-ms delay, *Go* appears on screen as the cue for participants to repeat the preamble and complete the sentence in the passive voice.

included in the predicate, such as a verb or an adjective. In this modification, this additional word is typically presented after the initial fixation point and remains on screen for about 600 ms. The sentence-preamble task is usually self-paced with the participant pressing the space bar to move to the next preamble.

Theoretical frameworks

L1 production Within the three-stage speech production model described previously, grammatical encoding occurs in two phases within the formulation stage. First, functional processing assigns grammatical functions (i.e., subject or object) and morphological information to lexical representations, which creates hierarchical relationships. Next, the output of the functional processing stage is subject to positional processing, which imposes serial word order on the elements within phrases and between phrases within the utterance, which is referred to as linearization (see Ferreira & Engelhardt, 2006, for an overview of syntax in production). Researchers have used sentence preamble tasks initially to test claims that non-syntactic information has no impact on grammatical encoding. However, subsequent studies have shown that non-syntactic information, such as the notional number of the head noun (i.e., its semantic properties as opposed to grammatical number), also impacts grammatical agreement (Eberhard, 1999; Barker, Nicol, & Garret, 2001; Thornton & MacDonald, 2003; Humphreys & Bock, 2005). For example, the head

noun in the preamble the label on the wine bottles is grammatically singular but notionally plural as it can be interpreted as referring to multiple labels. A sample study is summarized in study box 7.4.

Study Box 7.4

Humphreys, K., & Bock, K. (2005). Notional number agreement in English. *Psychonomic Bulletin and Review, 12*, 689–695.

Purpose

Determine whether notional number influences subject–verb agreement in both the visual and auditory modalities. It was predicted that, if notional number plays a role in agreement, plural verbs will be more frequent after spatially distributed phrases (*the gang near the motorcycles*) than spatially collective phrases (*the gang on the motorcycles*). It was predicted that, if modality influences subject–verb agreement, plural verbs will be more frequent when preambles are read rather than heard.

Participants

144 college-age native English speakers.

Task design

18 sentence preambles were created with a head noun and prepositional phrase modifier. Six versions of each preamble were created by manipulating the head noun (collective or plural), the local noun in the prepositional phrase (singular or plural), and the prepositions (distributed or collective construal). Six lists were created consisting of three versions of each preamble and 50 fillers.

Procedure

Participants carried out the task individually using Macintosh computers with PsyScope experimental software. Half of the participants heard the preambles over external speakers, while the other half read the preambles on screen. When cued by the program, they repeated and completed the preamble.

Analysis

Responses were transcribed and verb number was classified into four categories: (1) singular, (2) plural, (3) uninflected (such as past tense verbs), and (4) miscellaneous (incorrect, unintelligible, or dysfluent preambles; no verb produced, preambles not used as the subject).

Results

ANOVAs were carried out with the percentage of plural verbs as the dependent variable and head noun, local number/construal, and modality as the factors, followed by planned pairwise comparisons. The main findings were (a) more plural verbs occurred after plural local nouns than singular local nouns, (b) more plural verbs occurred after distributed construals than collective construals, and (c) modality had little impact on agreement performance.

Bilingual/L2 production In terms of grammatical processing, bilingual speech production researchers have questioned the relationship between the grammatical properties of the two languages and how it impacts production of the response language. Previous studies have shown that speakers of Spanish, Italian, French, and Dutch are more sensitive to notional number than are English speakers. In other words, unlike English speakers, they make more agreement errors when the head noun, while grammatically singular, can be interpreted as conceptually plural. Bilingual studies have used the sentence preamble task to explore whether bilingual speakers' agreement errors are similar to L1 or L2 patterns. For example, while low-proficiency Spanish L2 speakers appear to carry over their English agreement errors when speaking Spanish, Spanish-dominant bilinguals with native proficiency in English show Spanish error patterns when speaking English (Nicol, Teller, & Greth, 2001).

Methodological considerations

In terms of task design, the most important consideration is the design of the sentence preambles. A set of basic sentence preambles is created first, and then manipulated so that there is one version of each preamble for every condition being tested. The post-modifying constituent with the local noun is typically a prepositional phrase, as studies which included relative clauses found that they elicit fewer subject–verb agreement errors (Bock & Cutting, 1992). The constituents that modify the head noun need to be matched in length. The filler preambles should avoid the relationships being targeted in the experimental conditions, but should have similar length and complexity. Some studies provide a verb to be used when completing the preamble in order to have greater control over the plausibility relations among the head noun, local noun, and the verb (Thornton & MacDonald, 2003). When the sentence preambles manipulate semantic factors, pilot testing is needed to control ratings of plausibility or distributed versus collective interpretations across the experimental conditions (Thornton & MacDonald, 2003; Humphreys & Bock, 2005).

For the analysis, the preamble completions are transcribed and then classified based on subject–verb agreement. Typically four categories are used to capture the logical possibilities. Correct completions are generally defined as correct repetition of the preamble and completion using a verb with correctly marked agreement. Agreement errors are defined as correct repetition of the preamble with incorrectly marked agreement, such as using a plural verb after a grammatically singular head

noun due to the presence of a plural local noun. Uninflected responses are defined as a correct repetition of the preamble along with a completion with a verb that lacks number inflection or occurs in past tense. The final category is often termed "other" or "miscellaneous" and includes situations in which the preamble was not repeated or was incorrectly repeated, or there was no response, or the preamble was not used as the subject in the completion.

Concluding Remarks

As this brief overview has made apparent, psycholinguistic methodologies typically involve individual speakers who are asked to comprehend or produce words or sentences in isolation. Garnham, Garrod, and Sanford (2006) acknowledged that this focus on individual speakers persists despite the fact that dialogue is the most natural and basic form of language use, and expressed hope that future studies would investigate language processing during dialogue. We too are hopeful that L2 researchers interested in psycholinguistics will expand their research methodologies to explore language processing during situated language use. Of course, not all of the tasks described here could be adapted for conversation, but it is possible to design communicative tasks that have semantically and phonologically related distracters or present pictures with potentially facilitating or inhibiting images.

We also share Sears's (1986) concern that the inferences made from research carried out with a narrow database, specifically college students tested in a laboratory with academic-like tasks, may skew a field's understanding of basic processes. Like Sears's field of social psychology, L2 research also over-represents late adolescents in academic settings and under-represents other age groups in more varied contexts. And certainly Grosjean's (2008) advice concerning the need for caution when using experimental tasks designed for monolinguals with bilingual speakers warrants reiteration. Issues such as language proficiency and dominance, language mode, and the comparability of stimuli across languages can influence the extent to which research findings represent bilingual and L2 processing generally as opposed to task-specific processing. Finally, we would like to encourage L2 researchers to also consider using these comprehension and production methodologies as a tool for exploring the processes involved in language acquisition.

Project Ideas and Resources

Further Reading

Clahsen, H. (2008). Behavioral methods for investigating morphological and syntactic processing in children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 1–27). Amsterdam, Netherlands: John Benjamins.

Ferreira, F., & Anes, M. (1994). Why study spoken language? In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp.33–56). New York, NY: Academic Press.

Heredia, R., & Stewart, M. T. (2002). On-line methods in bilingual spoken language research. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 7–28). Amsterdam, Netherlands: Elsevier.

McDonough, K., & Trofimovich, P. (2008). *Using priming methods in second language research*. New York, NY: Routledge.

Study Questions

1. In applied linguistics research, are comprehension and production typically studied separately or together? What are some areas of research within applied linguistics that either integrate or separate the study of comprehension and production?
2. In the self-paced reading task, researchers typically ask participants to do a secondary task, such as answer a comprehension question or make a plausibility judgment, in order to increase the likelihood that the sentences are comprehended. In what situations might the secondary task influence the participants' performance?
3. Why might a researcher decide to use inter-response time instead of pause duration as the dependent measure for the self-paced listening task?
4. Considering their similarities, in what situations might it be more appropriate to use the picture–word interference task instead of the Stroop task?
5. All of the tasks described in this chapter are administered to individuals seated at computers that have been programmed with experimental software. How could these tasks be adapted for use in low-tech environments or for group administration, such as in L2 classrooms?

References

Barker, J., Nicol, J., & Garrett, M. (2001). Semantic factors in the production of number agreement. *Journal of Psycholinguistic Research, 30*, 91–114.

Bi, Y., Xu, Y., & Caramazza, A. (2009). Orthographic and phonological effects in the picture–word interference paradigm: Evidence from a logographic language. *Applied Psycholinguistics, 30*, 637–658.

Bock, J. K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin and Review*, *3*, 395–421.

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

Booth, J. R., MacWhinney, B., & Harasaki, Y. (2000). Developmental differences in visual and auditory processing of complex sentences. *Child Development*, *71*, 981–1003.

Caplan, D., & Waters, G. (2003). On-line syntactic processing in aphasia: Studies with auditory moving window presentation. *Brain and Language*, *84*, 222–249.

Colomé, A., & Miozzo, M. (2010). Which words are activated during bilingual word production? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 96–109.

Costa, A. (2004). Speech production in bilinguals. In T. Bhatia & W. Ritchie (Eds.), *The handbook of bilingualism* (pp. 201–223). Malden, MA: Blackwell.

Costa, A., & Caramazza, A. (1999). Is lexical selection language specific? Further evidence from Spanish–English bilinguals. *Bilingualism: Language and Cognition, 2*, 231–244.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41*, 365–397.

De Jong, N. (2005). Can second language grammar be learned through listening? An experimental study. *Studies in Second Language Acquisition*, *27*, 205–234.

Dell'Acqua, R., Job, R., Peressotti, F., & Pascali, A. (2007). The picture–word interference effect is not a Stroop effect. *Psychonomic Bulletin and Review, 14*, 717–722.

Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition*, *25*, 529–557.

Dussias, P. E., & Piñar, P. (2009). Sentence parsing in L2 learners: Linguistic and experience-based factors. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 295–317). Bingley, England: Emerald.

Eberhard, K. (1999). The accessibility of conceptual number to the processes of subject–verb agreement in English. *Journal of Memory and Language, 41*, 560–578.

Felser, C., Marinis, T., & Clahsen, H. (2003). Children's processing of ambiguous sentences: A study of relative clause attachment. *Language Acquisition*, *11*, 127–163.

Ferreira, F., Anes, M. D., & Horine, M. D. (1996). Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research*, *25*, 273–290.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.

Ferreira, F., & Engelhardt, P. (2006). Syntax and production. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 61–91). Amsterdam, Netherlands: Elsevier.

Ferreira, F., Henderson, J. M., Anes, M., Weeks, P. A., & McFarlane, D. K. (1996). Effects of lexical frequency and syntactic complexity in spoken-language comprehension: Evidence from the auditory-moving window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 324–335.

Frazier, L. (1990). Exploring the architecture of the language-processing system. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 383–408). Cambridge, MA: MIT Press.

Garnham, A., Garrod, S., & Sanford, A. (2006). Observations on the past and future of psycholinguistics. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 1–18). Amsterdam, Netherlands: Elsevier.

Griffin, Z., & Ferreira, V. (2006). Properties of spoken language production. In M. Traxler & M Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 21–60). Amsterdam, Netherlands: Elsevier.

Grosjean, F. (2008). *Studying bilinguals*. Oxford, England: Oxford University Press.

Hantsch, A., Jescheniak, J., & Schriefers, H. (2009). Distractor modality can turn semantic interference into semantic facilitation in the picture–word interference task: Implications for theories of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1443–1453.

Heredia, R. R., Stewart, M. T., & Cregut, I. (1997, November). Bilingual online sentence processing: Frequency and context effects in code switching. Poster presented at the 38th annual meeting of the Psychonomic Society, Philadelphia, PA.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, *22*, 369–397.

Humphreys, K., & Bock, K. (2005). Notional number agreement in English. *Psychonomic Bulletin and Review, 12*, 689–695.

Jackson, C. N., & Bobb, S. C. (2009). The processing and comprehension of *wh*-questions among second language speakers of German. *Applied Psycholinguistics*, *30*, 603–636.

Juffs, A. (1998). Main verb versus reduced relative clause ambiguity resolution in L2 sentence processing. *Language Learning*, *48*, 107–147.

Kaushanskaya, M., & Marian, V. (2007). Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning, 57*, 119–163.

Knupsky, A., & Amrhein, P. (2007). Phonological facilitation through translation in a bilingual picture naming task. *Bilingualism: Language and Cognition, 10*, 211–223.

Marshall, C., Marinis, T., & van der Lely, H. (2007). Passive verb morphology: The effect of phonotactics on passive comprehension in typically developing and Grammatical-SLI children. *Lingua*, *117*, 1434–1447.

Mitchell, D. C. (1984). An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 69–89). Hillsdale, NJ: Lawrence Erlbaum.

Mitchell, D. C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 15–32). New York, NY: Psychology Press.

Navarrete, E., & Costa, A. (2005). Phonological activation of ignored pictures: Further evidence for a cascade model of lexical access. *Journal of Memory and Language, 53*, 359–377.

Nicol, J., Teller, M., & Greth, D. (2001). The production of verb agreement in monolingual, bilingual, and second language speakers. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 117–133). Oxford, England: Blackwell.

Oppermann, F., Jescheniak, J., & Schriefers, H. (2008). Conceptual coherence affect phonological activation of context objects during object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 587–601.

Rah, A., & Adone, D. (2010). Processing of the reduced relative clause versus main verb ambiguity in L2 learners at different proficiency levels. *Studies in Second Language Acquisition*, *32*, 79–109.

Sears, D. (1986). College sophomores in the laboratory: Influences of a narrow data base on psychology's view of human nature. *Journal of Personality and Social Psychology, 51*, 513–530.

Thornton, R., & MacDonald, M. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language, 48*, 740–759.

Titone, D. A., Koh, C. K., Kjelgaard, M. M., Bruce, S., Speer, S. A., & Wingfield, A. (2006). Age-related impairments in the revision of syntactic effects of prosody. *Language and Speech*, *49*, 75–99.

Titone, D., Prentice, K. J., & Winfield, A. (2000). Resource allocation during spoken discourse processing: Effects of age and passage difficulty as revealed by self-paced listening. *Memory and Cognition*, *28*, 1029–1040.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton & L. Frazier (Eds.), *Perspectives on sentence processing* (pp. 155–179). Mahwah, NJ: Lawrence Erlbaum.

van Gompel, R. P. G., & Pickering, M. J. (2007). Syntactic parsing. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 289–308). Oxford, England: Oxford University Press.

van Maanen, L, van Rijn, H., & Borst, J. (2009). Stroop and picture–word interference are two sides of the same coin. *Psychonomic Bulletin and Review, 16*, 987–999.

Waters, G. S., & Caplan, D. (2004). Verbal working memory and on-line syntactic processing: Evidence from self-paced listening. *Quarterly Journal of Experimental Psychology*, *57A*, 129–163.

Williams, J. N., Möbius, P., & Kim, C. (2001). Native and non-native processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics*, *22*, 509–540.

Zhang, Q., Chen, H., Weekes, B., & Yang, Y. (2009). Independent effects of orthographic and phonological facilitation in spoken word production in Mandarin. *Language and Speech, 52*, 113–126.