



Observing the what and when of language production for different age groups by monitoring speakers' eye movements

Zenzi M. Griffin *, Daniel H. Spieler

Georgia Institute of Technology, School of Psychology, 654 Cherry St., Atlanta, GA 30332-0170, USA

Accepted 8 August 2005

Available online 10 November 2005

Abstract

Research on adult age differences in language production has traditionally focused on either the production of single words or the properties of language samples. Older adults are more prone to word retrieval failures than are younger adults (e.g., Burke, MacKay, Worthley, & Wade, 1991). Older adults also tend to produce fewer ideas per utterance and fewer left-branching syntactic structures (e.g., Kemper, Greiner, Marquis, Prenovost, & Mitzner, 2001). The use of eye movement monitoring in the study of language production allows researchers to examine word production processes in the context of multiword utterances, bridging the gap between behavior in word production studies and spontaneous speech samples. This paper outlines one view of how speakers plan and produce utterances, summarizes the literature on age-related changes in production, presents an overview of the published research on speakers' gaze during picture description, and recaps a study using eye movement monitoring to explore age-related changes in language production.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Language production; Object naming; Age differences; Aging; Verbal ability; Vocabulary; Lexical access; Word selection; Codability; Word frequency; Syntax; Disfluency; Timing; Speech errors; Older adults; Eye movements; Spontaneous speech; Picture description

1. Introduction

Although intuitions often tell us otherwise, speaking is complicated and effortful. Speakers determine the communicative goals that they use language to achieve and from these goals, they choose an idea to express in each utterance. Within an utterance, the elements of an idea are sequenced for grammatical expression and paired with words. Each meaningful word is selected from a working lexicon containing tens of thousands of words for an average adult. These words are assembled as a rhythmic pattern of sounds, articulated via the coordination of over a hundred muscles (Lenneberg, 1967).

Most impressively, this occurs with enough speed to accommodate a normal conversational speech rate of two to three words per second or about 5.65 syllables per second (see e.g., Bock, 1982; Garrett, 1988; Levelt, 1989; MacKay, 1982; for theories and reviews).

The overarching goal of speakers is usually to produce relatively fluent speech that achieves their communicative goals in a timely manner (e.g., Clark, 2002). For uninterrupted speech, speakers need to coordinate the sequence and timing of processes such that their results become available when the speaker requires them. In other words, what speakers want needs to be ready when they want it. Depending on the process at hand, the *what* refers to ideas, concepts, words, or motor programs that express the speaker's intentions. Failing to produce the *what* disrupts the communicative goals of speaking. The *when* concerns the sequence and timing of processes that operate on the *what*. Failure to coordinate the *when* of

* Corresponding author.

E-mail addresses: zenzi.griffin@psych.gatech.edu (Z.M. Griffin), daniel.spieler@psych.gatech.edu (D.H. Spieler).

production has different consequences depending on whether processing time is over- or underestimated.

A common assumption is that the conceptual content expressed in an utterance is not inherently ordered although its expression in speech is (Bock, 1982; Levelt, 1981; Wundt, 1900/1970). Speakers parcel and order the content to be expressed over time. When the encoding of content in words and sounds lags too far behind articulation, the fluency of speech suffers. While completing preparation of a word that should have been ready to articulate, a speaker may repeat or prolong the preceding words, utter fillers such as *uh* or *um*, or pause (see e.g., Goldman-Eisler, 1968; Jurafsky, Bell, Gregory, & Raymond, 2001; Shriberg, 1994). Although such disfluencies disrupt the flow of speech, they do not necessarily impede its understanding (e.g., Arnold, Fagnano, & Tanenhaus, 2003; Fox Tree, 1995). Nonetheless, they may exact a cost for the speaker because listeners associate such disfluencies with uncertainty and even incompetence (e.g., Christenfeld, 1995). In contrast, the larger the lag between the preparation of words and their articulation, the greater the demands on working memory from buffered articulatory plans. Although such buffering typically results in fluent speech (Griffin & Bock, 2000; Wheeldon & Lahiri, 1997), studies of short-term memory suggest that effort increases with the number of items buffered (e.g., Kahneman & Beatty, 1966; Peavler, 1974).

These are important but loose constraints on production. Even when speech is error-free and flows smoothly, there can be no single solution to the logistical problem of getting what the speaker wants when they want it. That is, fluent and errorless speech may be the result of considerable advance planning with buffering or successful last-second word retrieval. Understanding the various ways in which speakers may get what they want when they want it is one of the primary questions in language production.

Across their lifespans, speakers follow a developmental trajectory in which there are significant changes in the underlying cognitive processes upon which language processing depends. Of course, the most studied portion of this trajectory corresponds to childhood and early stages of language acquisition (see e.g., Ingram, 1989). Yet, as an individual speaker moves from their 20s to their 50s and on into their 70s, there are also significant changes to the speed and efficiency of information processing (e.g., Cerella, 1985; Hale & Myerson, 1995, 1996). Evidence suggests that younger and older speakers may not make the same decisions about the what and when (the content and timing) of language production (Griffin & Spieler, 2000; Kemper, Herman, & Lian, 2003). The fact that the vast majority of individuals continue to communicate effectively well into old age suggests that speakers are able to make adjustments to processing that allow them to continue to ful-

fill the combined constraints on timing and content that are present in every communicative situation. What adjustments do speakers make and what consequences do these adjustments have for the production of spoken language?

In what follows, we provide a brief overview of language production processes and highlight the relationship between what and when, or content and timing, in production processes. We then review some basic findings relevant to understanding age differences in language production and discuss results from our laboratory that deal specifically with how young and older speakers solve, and in some cases, do not solve the what and when problem in language production. We end by noting what the study of age differences offers to understanding language production in general.

2. Language production overview

2.1. The what

Language production is divided into multiple stages of processing that begin with a communicative intent and end with the execution of a motor program (see e.g., Fromkin, 1968; Levelt, 1989). In the first stage, a speaker creates a communicative intent or goal for a speech act. This contains the information that the speaker wishes to convey (e.g., an anecdote) and the goal of communication (e.g., amusing others). This communicative intent may require multiple utterances and the speaker must decide how to parcel the information within this communicative intent into individual utterances. This stage we refer to as message planning and its product is a chunk of conceptual information (i.e., a *message*) that can be translated into a single utterance,¹ roughly the size of a single clause or sentence. While this message requires linguistic processing before it is in a form that may be spoken, it contains all of the conceptual and pragmatic information for an utterance. For example, a message that is eventually output as *She gave the dog a bone* will contain information indicating that the topic of the assertion is female, identifiable in the current context, who of her own volition transferred a meatless skeletal remnant to an identifiable canine pet at an earlier point(s) in time (see Jackendoff, 1990, for one theory of message structure).

The next stage determines the words used and their order in the eventual utterance. Many accounts view the syntactic structure of an utterance as primarily driven by

¹ Although sentences are easily identified in written language, they cannot be identified using the same or similar criteria in spoken language (e.g., Miller & Weinert, 1998). Instead, it is common to talk about utterances instead. These correspond roughly to a single response, uttered with a continuous intonational contour.

word choice (e.g., Bock & Levelt, 1994; Ferreira, 2000; Levelt, 1989; Pickering & Branigan, 1998). Alternatively, syntactic structure may be viewed as a function of properties of message elements and relationships between them, in combination with experience- and availability-based decisions about the order in which to express the message elements (Chang, 2002; Chang, Dell, Bock, & Griffin, 2000; Dell, Chang, & Griffin, 1999; Gordon & Dell, 2002; Griffin & Weinstein-Tull, 2003). Utterances often begin with a reference to a message element that is already known to the addressee or that is highly available due to other factors such as animacy, vividness, or its role in the message, but not perceptual salience (for review, see Bock, 1982; Bock, Irwin, & Davidson, 2004; Dell et al., 1999). In English, this element usually functions as the subject of the sentence. In *She gave the dog a bone*, the message element functioning as grammatical subject is a female who is identifiable in the current context and performed the action the speaker mentions in the utterance. The speaker selects a word (or words) to express this message element in a noun phrase and assembles the sounds of the word(s). A single message element will usually correspond to a noun phrase, which in turn may be comprised of multiple words and phrases (e.g., *the abnormally tall woman with long hair*), or a single noun (e.g., *woman*) or pronoun (*she*). In our example, the discourse context allows the speaker to use a pronoun, *she*, to encode this first part of the message, but leads the speaker to choose a more specific noun phrase, *the dog*, for the recipient of the woman's action.

The sounds that express a message element often form a single phonological word, containing one stressed syllable and a number of unstressed syllables (e.g., *the-Woman*). However, they may form units smaller than a phonological word when expressed with pronouns such as *she* or *it*. Phonological words are often considered the units of phonological encoding because they define the boundaries within which sounds exert a strong influence each other's pronunciations (i.e., coarticulation; see Wheeldon, 2000). Also, when speaking extemporaneously in the laboratory, speakers seem to prepare approximately one phonological word before beginning to speak (Wheeldon & Lahiri, 1997; but see Costa & Caramazza, 2002; Griffin, 2003). In contrast, greater advanced planning seems to involve the buffering of units that correspond to phonological words (Sternberg, Knoll, Monsell, & Wright, 1988; Wheeldon & Lahiri, 1997). Articulation of an utterance is typically considered to involve transcoding of syllables from a phonological word into motor programs or goals and executing these (see Levelt, Roelofs, & Meyer, 1999). Curiously, although preparation time increases with the number of phonological words planned, other increases in the number of syllables, segments, or duration of an utterance do not increase preparation time (see Bacchoud-Levi, Dupoux, Cohen, & Mehler, 1998; Griffin,

2003; Meyer, Roelofs, & Levelt, 2003; Sternberg et al., 1988; Wheeldon & Lahiri, 1997, 2002).

To summarize, in terms of content, speakers create messages in which they focus on a single message element at a time, select words to express the element, retrieve the sounds of the word(s), assemble a phonological word, which is articulated as a sequence of syllable-sized motor programs. Across speakers of all ages, the same considerations are likely to apply. However, at each level in processing, speakers have some flexibility. They may parcel their message into larger or smaller packages, resulting in either longer or shorter utterances with more or less complex grammatical structures (e.g., Kemper et al., 2003). Likewise, speakers may prefer more or less specific lexical labels for the contents of their utterances. These decisions about the *what* of processing have consequences for the *when*.

2.2. The when

It is common to talk about the timing of language production as incremental (e.g., Kempen & Hoenkamp, 1987; Levelt, 1989). That is, processing at one stage may work on an increments (rather than whole units) output by an earlier stage, such that processing occurs at multiple stages simultaneously. For instance, speakers usually begin articulating words of an utterance before they finish retrieving all the sounds of the utterance's words. In this sense, all current theories of language production are incremental. The implications of incrementality vary with the level of processing one considers and the units one believes are involved. Furthermore, there is an important distinction to be made between a processing stage that obligatorily operates on any available increment as opposed to one that may act on increments or wait (Ferreira & Swets, 2002). For instance, in a strategically incremental system, a word could be selected to express part of a message before an entire message was created. In an architecturally incremental system, a word selection might inevitably begin as soon as part of the message was possible to lexically encode. The implications of incrementality also depend on whether buffering is possible and at which stages (compare models in Dell et al., 1999, & Martin & Freedman, 2001). For these reasons, it is important to bear in mind the generality of the term in discussing the timing of production and to provide more precise characterizations of the timing of processing stages and the flexibility of the timing.

The scope or unit of encoding at each processing level is typically smaller than the scope or unit of encoding at the immediately prior level. Theories agree that processes below the message level deal in units smaller than complete clauses (Dell, 1986; Garrett, 1975; Kempen & Hoenkamp, 1987; Levelt, 1989). There is good evidence that speakers often create a representation that corresponds to a proposition or clause at the message level

before beginning to produce an utterance (see Bock & Levelt, 1994, for review), although the details of all message elements may not initially be specified (e.g., Griffin & Oppenheimer, 2003). Most theories also hold that words are usually phonologically encoded shortly before they are said (e.g., Butterworth, 1989; Ferreira, 2000). The controversy lies primarily in the extent to which speakers select words before articulation of an utterance begins and the extent to which multiple content words may be selected in parallel (see Levelt & Meyer, 2000; Martin, Miller, & Vu, 2004; Smith & Wheeldon, 1999).

In the theory of language production presented here, lexical-grammatical encoding operates on one message element at a time, which typically corresponds to a single simple or complex noun phrase. Speakers focus on a message element until the words to express it have been phonologically encoded (Meyer, Sleiderink, & Levelt, 1998). Timing evidence suggests that only one noun may be lexically and phonologically processed at a time (Griffin, 2003; Meyer et al., 2003),² but parallel preparation of words from different grammatical classes to express a single message element or within idioms may be possible (see e.g., Levelt & Meyer, 2000; Martin & Freedman, 2001).

The words that have already been prepared combine with the unexpressed message content and the grammatical and lexical constraints of the speaker's language to determine the order in which the remaining message elements are lexically encoded and which function morphemes should be added (Chang, 2002; Chang et al., 2000; Dell et al., 1999; Gordon & Dell, 2002; Griffin & Weinstein-Tull, 2003). For example, having encoded a female agent as a grammatical subject, *she*, an English speaker has little choice but to express the action next in the form of a verb. In contrast, a speaker of a language like Russian would have more flexibility about what part of the message to express next and a speaker of Korean must express the other nominal message elements in a clause (the dog and the bone) before producing the verb. Although English speakers have little flexibility in when the action corresponding to the verb *give* is expressed relative to the other message elements, they have flexibility in the ordering of expressions for dog and bone. An English speaker may say *She gave a bone to the dog* as well as *She gave the dog a bone*. The speaker's grammatical knowledge would signal inclusion of the word *to* before the noun phrase expressing the goal (dog) when it was encoded after the theme (bone), but omit *to* when the goal was encoded before a simple theme.

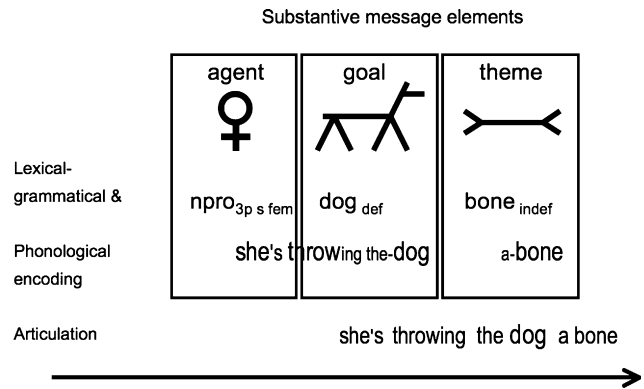


Fig. 1. Illustration of the time course and sequence for focusing on message elements to lexically and grammatically encode them relative to articulating an utterance. The focus and lexicalization time for each substantive element is contained in a square. The relationship between these squares and the utterance should also roughly reflect the sequence and relative timing of eye movements to the referents in a described scene when there is minimal indecision about order of mention.

So, the speaker has a proposition-sized representation of their message that guides sequential linguistic processing (Wundt, 1900/1970). Within the message, the speaker concentrates on a single message element at a time, selects words to express it, retrieves and assembles sounds of the words to create a phonological-prosodic plan. Fig. 1 illustrates a time course for planning the utterance *She's throwing the dog a bone*.³ Via coordinated articulatory gestures, speakers articulate phonological words. Speakers may begin articulation soon after a phonological word has been encoded, but unless it requires a long time to articulate or the speaker articulates it slowly, this is likely to lead to later disfluencies when articulation proceeds faster than phonological words are created. Relatively fluent speech seems to require about one second of unspoken material or one phonological word to be buffered. Thus, speakers spend more time preparing a second word before beginning to speak when the first word is short like *she* as opposed to a longer word like *woman* (Griffin, 2003). Although a one-second buffer suffices for young adults who usually prepare nouns in less than a second and rarely suffer from word retrieval problems, it may not work for older speakers. Investigating this question requires a sensitive measure for tracking the timing of word preparation in multiword utterances. Next, we relate these production processes to speakers' eye movements.

² It is necessary to assume that delays in speech onset for naming pairs of semantically related objects (Freedman, Martin, & Biegler, 2004) are due to difficulty in selecting a name for the first object when there is a similar concept in the message rather than attempts to select names for both objects simultaneously.

³ We do not think there is sufficient evidence to indicate when verbs are typically selected. Only the phonological encoding of the verb *throwing* is illustrated in the figure.

3. Eye movements and the what and when of production⁴

Probably the most common method of specifying the messages that speakers use to produce language is to ask them to label line-drawn objects one at a time (Bock, 1996). Researchers measure both the time that this object naming requires and its accuracy. When researchers wish to study the production of multiword utterances, they often use line drawings of simple actions such as a woman handing a bone to a dog. The visual information in these displays is a crude attempt to specify and externalize message content for each utterance. Researchers may analyze a speaker's choice of syntactic structure for the resulting description, its fluency, or the speed with which speech is initiated, but prior to the use of eye movement monitoring, methods of tracing when speakers prepared individual words in their utterances were relatively indirect and required many questionable assumptions. A strength of using eye movement monitoring in the study of language production is the ability it provides to make inferences about when a speaker prepares words, particularly nouns (see Griffin, 2004; Meyer & Lethaus, 2004, for more on the strengths and weakness of eye movement monitoring in language production).

People acquire detailed information about the visual world from only a small part of it at a time (for review see Henderson & Ferreira, 2004; Irwin, 2004; Rayner, 1998). This is the part of the world that a person is *foveating* or *fixating*. The limited size of this area can readily be experienced by staring at one word on a page and trying to recognize other words without moving one's eyes. Although it is obvious that text fills the page, it should be impossible to identify words that are not adjacent to the fixated word without moving one's eyes. Beyond the region of the world that one views with high acuity is a greater region experienced with lower acuity. This lower acuity vision allows viewers to see how far text on a page extends and, in general, to take in enough information to decide where potentially interesting visual information is located that they might want to fixate. The degree to which objects and words can be recognized without fixating on them is a function of their size on the retina (how far away the viewer is and how large they are) and what is around them. Words tend to be placed tightly together in text, which makes it difficult to identify words that are far from fixation. In contrast, objects in scenes are often relatively large and may be spaced further apart, making it easier to recognize ones that do not lie close to fixation. Thus, much detail related to an

object or scene can be extracted within a single fixation (e.g., Potter, 1975), but few words may be read (see Rayner, 1998).

Unsurprisingly, these properties of the human visual system have major consequences for using eye movements as a reflection of cognitive processing. Just because a speaker has not fixated on an object does not mean that it has not been recognized. For instance, when describing line-drawn scenes that contained several objects, people did not fixate all of them to find the object being acted on or the one to mention first in an utterance (Griffin & Bock, 2000). Nonetheless, agency, relative humanness, and animacy were used to decide which object to mention first, implying that these important properties were recognized for many objects without fixating them. However, degrading the quality of visual forms (Meyer et al., 1998) or creating more visually complex scenes (Griffin, 1998) increases the likelihood that objects may be fixated before they are recognized and incorporated in messages.

Despite the ability to recognize some objects prior to fixation, people tend to fixate objects while they perform related cognitive operations so the time spent fixating an object tends to reflect the time spent in related processes (e.g., Just & Carpenter, 1976). However, people often fixate multiple parts of an object (Buswell, 1935) and the duration of individual fixations on an object is relatively uninformative (Griffin, 1998). Although people do not acquire visual information while moving their eyes between parts of an object, the durations of these saccades may be devoted to non-visual cognitive operations related to the object (see Irwin, 2004). Therefore, production researchers typically combine multiple concurrent fixations on an object in what they call a gaze, a viewing, or an inspection, which begins when an object is first fixated and ends when the eye moves to fixate a different object. What is considered an object or region of interest depends on how researchers conceptualize the visual displays that they use. For instance, when speakers are shown an analog clock to use in telling time (e.g., "twelve twenty five"), regions of interest may be quadrants of the clock (Bock, Irwin, Davidson, & Levelt, 2003). In contrast, when speakers describe actions in scenes, like a child giving an apple to a teacher, a clock in the scene may only be considered as part of the background such that gazes on it are not individuated from fixations on the wall or floor (Griffin, 1998).

When describing visual scenes, the time spent gazing at an object is highly correlated with the time required to select a name for it and assemble its sounds (Griffin & Bock, 2000). Lexically encoding a message element is fastest when it corresponds to something a speaker has recently referred to and the speaker uses a pronoun like *it* or recently produced word (e.g., Lachman, Shaffer, & Hennrikus, 1974). Thus, speakers spend less time gazing at objects when naming them for a second time, using

⁴ The use of eye movement monitoring in language production owes its existence to the creation of the visual world paradigm, in which the time course of spoken language interpretation is inferred from the eye movements of listeners as they hear instructions (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

either a pronoun or a common noun (Van Der Meulen, Meyer, & Levelt, 2001). The amount of time it takes to select a content word increases as the number of context-appropriate choices increases. For example, a television may be called *television* or *TV* and it takes longer to produce either of those names relative to objects with a single dominant name like *table* (Lachman, 1973; Lachman et al., 1974). This difference in object naming latencies is also reflected in the time speakers spend gazing at objects before naming them (Griffin, 2001). Speakers take less time to retrieve words that they frequently use such as *baby* relative to uncommon words like *button* (Oldfield & Wingfield, 1965). This frequency effect is directly reflected in the time speakers spend gazing at objects before referring to them (Meyer et al., 1998). When factors such as frequency and number of context-appropriate names are controlled, it does not take any more time to prepare long words like *chandelier* compared to short words like *chef* (Bachoud-Levi et al., 1998). Likewise, the length of its name does not affect the time spent gazing at an object presented in mixed lists (Griffin, 2003; Meyer et al., 2003).

The high correlation between object naming latencies and the time spent gazing at them before saying their names makes eye movement monitoring redundant with naming latencies when speakers name single objects. The benefit of eye movement monitoring comes from the information it provides about preparing words when there are multiple objects mentioned in an utterance. In particular, monitoring eye position allows researchers to make inferences about when a word is prepared without altering the basic task of picture description. Earlier methods of probing the time course of word production involved presenting distracting words (Meyer, 1996, 1997), switching tasks or stimuli across trials (Dell & O'Seaghdha, 1992), or assuming that picture preview sped up response latencies in increments that correspond to completed processing at a stage of word production (e.g., Martin et al., 2004; Smith & Wheeldon, 1999). Researchers have also argued for various time courses for word preparation based on patterns in speech errors and disfluencies (Garrett, 1975; Goldman-Eisler, 1968). Although these methods are all informative, eye movement monitoring provides a far richer basis for evaluating the timing of word preparation.⁵

⁵ A recent study by Morgan and Meyer (2005) found a preview benefit from presenting a picture of a homophone (e.g., flying bat) in the position of a to-be-named object (e.g., baseball bat) before the speaker fixated the to-be-named object (replicating Pollatsek, Rayner, & Collins, 1984). While this result could be interpreted as showing retrieval or activation of object names before fixation, it is also plausible that activation of the homophone's concept only contributed to the activation of the target name via resonating activation after the homophonic word form was retrieved by the to-be-named object. The absence of interference from categorically related preview objects further suggests that previewed objects did not activate object name representations prior to speakers fixating to-be-named objects.

Moreover, eye movement monitoring may be informative about more than just the timing of word preparation. The order in which message elements are lexicalized is related to syntactic structure. For example, in producing dative sentences, lexicalizing a theme before a goal creates a prepositional dative such as *The woman gave a bone to the dog*, whereas lexicalizing the goal before the theme creates a double object, *The woman gave the dog a bone*, or heavy NP shift dative, *The man gave to the dog an absolutely tremendous bone* (Chang et al., 2000). Indecision and uncertainty are readily detected in eye movements where observers tend to look back and forth between candidates that they consider (Russo & Rosen, 1975). Likewise, indecision about the order in which to lexicalize message elements may be reflected in eye movements between objects in a scene. Such shifts imply that speakers have not decided word order and syntactic structure for an utterance. Thus, the time course for structuring utterances may be indirectly studied via eye movement monitoring (Griffin & Garton, 2003; Griffin & Mouzon, 2004). Furthermore, using more complex pictures, it may be possible to track the speaker's decisions about what to include in each message (Holsánová, 2001) and from which similar objects a target must be linguistically distinguished (Gregory, Joshi, & Sedivy, 2003).

It is primarily the ability to trace the time course of word production processes within the production of multiword utterances that makes eye movement monitoring such a promising paradigm for studying age-related changes in language production. In the next section, we briefly review the literature on language production in older adults to explain why this is the case.

4. The what and when of production in older adults

4.1. Word production

Across the life span, speakers learn new words, expanding their vocabularies. Indeed, most samples of convenience used to compare the performance of 20 year olds with that of adults over the age of 60, show higher vocabulary scores and better definition naming performance for the older group (e.g., Dahlgren, 1998; Rastle & Burke, 1996; Schroeder & Salthouse, 2004). These age-related vocabulary differences combine differences associated with aging and presumably exposure with differences due to cohort membership (Alwin & McCammon, 2001; Hertzog & Schaie, 1988). Among individuals in their 80s and 90s, vocabulary knowledge appears to decline but not as dramatically as other abilities such as perceptual speed (e.g., Lindenberger & Baltes, 1997).

Increasingly with age, speakers report a strong sense of knowing the word that they intend to say despite

being unable to retrieve all of its sounds (e.g., Burke et al., 1991; but see Brown & Gollan, 2003). A speaker who experiences this is in a *tip-of-the-tongue state* (TOT) (see Brown, 1991; Brown & McNeill, 1966). For example, a diary study showed that the older adults encountered an average of 6.56 TOT states during the course of the month compared to 3.92 TOTs for the younger adults (Burke et al., 1991).

Studies of TOT states in young adults and aphasics indicate that speakers in TOT states are often able to identify word-specific grammatical information such as count-mass noun distinctions in English (e.g., *a noodle* vs. *some spaghetti*; Vigliocco, Martin, & Garrett, 1999) and grammatical gender in Italian (Badecker, Miozzo, & Zanuttini, 1995; Vigliocco, Garrett, & Antonini, 1997). In many cases, speakers can report the first sound of the target word and its number of syllables (see Brown, 1991, for review). Thus, it is argued that speakers in TOT states have successfully selected a word but have failed to retrieve all of the word's phonological information. However, when older adults are asked to report characteristics of a TOT word, they recall less phonological information than younger adults do (e.g., Dahlgren, 1998; Maylor, 1990; Rastle & Burke, 1996). The increased susceptibility of older adults to TOT has been ascribed to a general weakening of the connections between representations, which in the Node Structure Theory of language production most affects the one-to-many connections between lexical representations and phonological representations such as syllables (see Burke et al., 1991; MacKay & Burke, 1990). The particularly weak activation of phonological information in older adults accounts for their reduced ability to report partial information about TOT words. This transmission deficit hypothesis receives further support from the finding that TOTs can be prevented or resolved by priming speakers with phonologically related words (James & Burke, 2000). This even holds for proper names, which are more prone to TOT states than other words (Burke, Locantore, Austin, & Chae, 2004).

One issue is whether age-related increases in vocabulary knowledge may account for some or all of the age-related increase in TOTs. Indeed there exists another class of speakers with large vocabularies who also shows a greater susceptibility to TOT states relative to less verbally proficient controls. Specifically, bilinguals experience more TOTs than do monolinguals when the sought-after words do not share forms across their languages (Gollan & Acenas, 2004). Furthermore, in a TOT study of young, middle-aged, and older adults, Dahlgren (1998) found no age differences in the frequency of experimentally elicited TOTs after age differences in vocabulary were taken into account. Gollan and Acenas (2004) posit that the effect of vocabulary on bilingual TOTs may be due to differences in usage leading to weaker connection strengths in bilinguals. Alternatively, Dahlgren (1998) suggests that a greater vocabulary may

lead to diffusion of conceptual activation among more potential expressions as in fan effects. Either of these mechanisms could lead to increases in word finding difficulty without invoking any effect of aging per se. More work needs to be done to investigate the extent to which these factors contribute to TOT states, but picture naming studies offer some support for the idea that older adults consider more alternatives responses than younger adults do and that this results in difficulty.

Asking speakers to label isolated pictures of objects or actions is the preferred method for evaluating the speed and accuracy of word production processes. Older adults often show a decrease in naming accuracy in both normative studies (Borod, Goodglass, & Kaplan, 1987; Kaplan, Goodglass, & Weintraub, 1983; Nicholas, Brookshire, MacLennan, Schumacher, & Porrazzo, 1989; Van Gorp, Satz, Kiersch, & Henry, 1986) and experimental studies (Au, Joung, Nicholas, & Obler, 1995; Bowles, Obler, & Albert, 1987). For both object and action naming, this decrease in accuracy appears most reliably in subjects over the age of 70, while the decrease is absent or less pronounced for adults in their 50s and 60s (Barresi, Nicholas, Connor, Obler, & Albert, 2000; Feyereisen, 1997; Goulet, Ska, & Kahn, 1994; Nicholas, Obler, Albert, & Goodglass, 1985). In studies that also measure the speed of picture naming, older adults may or may not be slower to provide a correct label to the picture, depending on the types of objects presented and the factors controlled (see Bowles, Obler, & Poon, 1989, and Mortensen, Meyer, & Humphreys, in press, for reviews).

The most common type of error made by unimpaired speakers in picture naming is semantically related word substitutions (e.g., *goat* for *sheep*; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997). This is also true of picture naming errors in older adults (Goulet et al., 1994; Schmitter-Edgecombe, Vesneski, & Jones, 2000). These errors are attributed to difficulty in word selection, where semantically related words are thought to compete for selection (e.g., Butterworth, 1989; Levelt, 1989). The age-related increase in error rates is not entirely attributable to visual confusions although researchers may observe a small increase in mixed perceptual or visual errors (see Goulet et al., 1994, for discussion).

Differences in vocabulary create problems in comparing picture-naming performance across age groups. Vocabulary differences alter the number and relative strength of candidate names speakers consider, which means that the difficulty of selecting the same word for the same picture may be very different for younger and older adults. For example, a young adult may consider *blimp* to be the only possible label for a picture of a blimp, whereas an older adult may consider *dirigible*, *zeppelin*, and *Hindenburg* reasonable candidates as well. The diversity and relative strength of candidate names exerts a strong influence on object naming latencies (e.g., Lachman, 1973; Lachman et al., 1974). Because picture-naming studies do

not traditionally take into account differences in the number of alternatives speakers of different age groups consider, they may conflate age differences with codability differences. It is even possible for vocabulary or codability differences across groups to favor older adults (e.g., Schmitter-Edgecombe et al., 2000).

The results from picture naming and TOT studies suggest that with age (particularly after turning 70), speakers are slower and less successful in retrieving words. It is not clear whether this age difference reflects some deficiency of processing (e.g., weakening connects in NST) or if it reflects differences in the structure of knowledge for older adults that in turn present different challenges during word retrieval for older compared to younger speakers. Regardless of the true source of these differences, they suggest that there often is an age-related difference in the *what* and the *when* of word production. Next we review age-associated differences in speaking more generally.

4.2. Production of multiword utterances

A number of studies have examined speech samples elicited from young and older adults. The measures and analyses of these language samples differ widely, but a few general conclusions appear relatively consistent across studies. Content measures reflect how much information younger and older speakers put into a single utterance. Here, an utterance is typically (although not consistently) defined either as a conversational turn, or a unit of speech that roughly maps onto a main clause and any accompanying subordinate clauses, or a single intonation contour. It is tempting to conclude that older adults tend to produce utterances that express fewer propositions and simpler syntax than younger adults (Kemper & Rash, 1988; Kynette & Kemper, 1986). However, many studies find that older adults produce the same number of words, if not more, per utterance than young adults and include as much, if not more, information (Cooper, 1990; Griffin & Spieler, 2004; Kynette & Kemper, 1986; Pasupathi, Henry, & Carstensen, 2002; Schmitter-Edgecombe et al., 2000). The primary age difference in syntax is that older adults appear less likely to produce left branching sentences in English, but these are uncommon and disdained structures in general (see Kemper, 1993). That said, longitudinal studies are also consistent with a decline in left branching structures and propositions per utterance when speakers reach their mid-seventies (Kemper, Greiner, et al., 2001; Kemper, Thompson, & Marquis, 2001).

Turning to the timing of speech in discourse settings, the results are more variable. At least some of this variability is likely to reflect the very different contexts in which speech is elicited (conversation, question answering, picture description, etc.). It is also likely to reflect the strategies speakers may use to speak in a timely manner.

Spierer, Horton, and Shriberg (2004) analyzed transcriptions and timing data from more than 1500 10 min telephone conversations drawn from 534 individual speakers. Across the speakers contributing to the corpus, the slowing of speech rate from 20 to 68 years of age in this data set was only 6%. This 6% age difference in speech rate is in striking contrast to the 30–50% often found in age comparative studies of other speeded tasks (e.g., Hale & Myerson, 1995, 1996). Yet, these conversation results agree with other studies reporting only small but significant changes in speech rate using tasks such as picture description (Cooper, 1990). Furthermore, Spieler and colleagues found that their older speakers tended to use more uncommon words than their younger speakers did. Thus again, differences in word choice may inflate or create age-associated differences in speed.

Spierer and colleagues' (2004) analyses of speech disfluencies, specifically fillers (*uh* or *um*), repetitions (*I... I went to the store*), and false starts (*I... you should go to the store*; known as deletions in the speech processing literature), only revealed a significant increase in the frequency of fillers with age (see also Bortfeld, Leon, Bloom, Schober, & Brennan, 2001). Fillers predict an impending delay in speaking (Clark & Fox Tree, 2002), which is consistent with the idea that word production may be slower in older adults. However, fillers are also used to mark uncertainty (Bortfeld et al., 2001; Smith & Clark, 1993), which could be related to age differences in familiarity with conversational topics or psychology laboratories, depending on the study. In other contexts however, older adults do appear to be more disfluent (Cooper, 1990; Kemper, 1992; Schmitter-Edgecombe et al., 2000)⁶ and this includes the use of silent pauses that are most associated with low transition probabilities and, in turn, with word selection difficulty (see Goldman-Eisler, 1968).

Notably, no age differences are typically found in latencies to speak or fluency measures for sentence construction tasks (Altmann & Kemper, in press; Davidson, Zacks, & Ferreira, 2003). These tasks reduce or eliminate differences in message and lexical content by providing speakers with content words to use to create grammatical sentences. In two experiments, Davidson and colleagues (2003) tested construction of dative sentences such as *I gave the dog a bone* and found identical results for younger and older speakers. Similarly, Altmann and Kemper (in press) found no age effects in the time younger and older adults took to initiate constructed sentences.

⁶ Complicating the review of this literature is that the definition of disfluency (insofar as it is clearly defined at all) varies radically across studies and can include discourse markers such as "well" and "anyway" (e.g., Kemper et al., 2003) that may serve a specific communicative intent (Clark & Fox Tree, 2002).

In summary, the results from studies of multiword utterances are not as consistent as one would like. Several suggest that older adults spontaneously produce simpler utterances, perhaps with slightly slower speed and more disfluencies. Other studies find little or no age difference in these aspects of multiword utterances. Like word retrieval, there are multiple differences between younger and older speakers that may contribute to differences in sentence production. It is often argued that older adults produce simpler utterances because age-related decreases in working memory capacity hinder them from producing more complex ones (Kemper, 1993). Indeed, creating and maintaining a message representation for grammatical encoding is probably the aspect of language production that makes greatest demands on working memory (Bock, 1982). In the next section, we consider a few other factors.

4.3. Other factors affecting the *what* and *when* of sentence production

There are a number of other factors aside from age per se that could contribute to differences in utterance content and complexity across age groups. One factor that is likely to be particularly important is exposure to complex utterances. Exposure to syntactic structures influences how likely speakers are to produce them (see Bock, 1986; Bock & Griffin, 2000; Tomasello, 2000; Weiner & Labov, 1983). Undergraduates are heavily exposed to the complex syntactic structures of academics in lectures and textbooks. After completing their educations, adults primarily use and experience simpler sentences in spoken and written language (see Miller & Weinert, 1998). Thus, the younger adults, typically college students, who participate in studies live in an environment in which written and spoken language is unusually complex. Furthermore, older adults report that younger interlocutors often simplify the speech that they address to them (e.g., Giles, Fox, & Smith, 1993; O'Connor & St. Pierre, 2004). This simplified speech or “elderspeak” shares many characteristics with baby talk or infant-directed speech, in that it involves simplified vocabulary, shorter and less complex grammatical structures, and exaggerated intonation (see e.g., Giles et al., 1993). Thus, age-related differences in syntactic complexity may be due in part to differences in the frequency and recency with which adults have been exposed to complex syntactic structures. Moreover, consistent with the idea that exposure to some syntactic structures varies with age group, older speakers may show greater effects of syntactic priming than undergraduates do for structures that occur more often in academic writing than elsewhere. For example, older speakers more reliably described a scene with a passive sentence after exposure to a passive sentence than younger ones did (Altmann, Kemper, Mathews, & Mullin, 2004). So, syntactic priming studies

and information about differences in grammatical complexity across populations and registers suggest that the grammatical complexity of older adults' utterances may largely reflect the grammatical complexity of the language directed to them.

Older speakers may also elect to simplify the content of their speech in order to facilitate its timing and avoid disrupting the flow of speech when they have word retrieval difficulties. Speakers can strategically alter the *what* of their speech to facilitate the *when*. Evidence comes from a study of language production in speech-only and dual task situations (Kemper et al., 2003). When simultaneously performing a second task such as finger tapping or playing a simple videogame, younger adults simplified the content of their speech but spoke as rapidly as they did when performing no secondary task. The change in content resulted in fewer ideas per utterance and simplified syntactic structure. That is, younger adults altered their message content to maintain a high speech rate. The speech of younger adults performing a secondary task resembled the speech of older speakers with no secondary task. In contrast, when performing a second task, older speakers did not simplify the content of their messages but spoke much more slowly than when speaking was their only task. One possibility is that the older adults may already have simplified their messages in response to age-related changes in production and further simplification was not feasible when a second task was added (Kemper et al., 2003). Instead, the older adults slowed their speech rate but maintained the same basic content level of their utterances.

In fact, several lines of evidence suggest that speakers can vary their speech rate and that such variations in timing have predictable influences on properties such as the fluency of the resultant speech (e.g., Griffin, 2003; Oomen & Postma, 2001b) or their ability to carry out a secondary task (Oomen & Postma, 2001a; Oomen & Postma, 2002). Specifically, by delaying or slowing speech, speakers can provide a cushion between the time when words are prepared and when articulation actually occurs. That is, the span between what part of the utterance is being prepared and the part that is currently being articulated. Increasing this “mind to mouth” span (Bock, 1995) provides a time cushion during which speakers can experience momentary word retrieval failures or correct speech errors without the necessity of an overt sign of that disruption such as a disfluency.

Across individuals, for example, fast speakers regardless of age are less fluent than slow speakers (Shriberg, 1994; Spieler et al., 2004). This pattern should occur because, assuming similar word preparation speeds, fast speakers will tend to allow less time between preparation of a word and its articulation, and thus have less of a time cushion when word preparation takes longer than anticipated. Second, when speakers in an experiment are compelled to complete preparation of an utterance prior

to articulating it, they take more time to begin speaking than extemporaneous speakers do, but their utterances are articulated more fluently, with shorter pauses (Griffin & Bock, 2000). Speech rate during articulation can be quicker because word preparation occurs prior to speech. Also, faster speech rates are strongly associated with increased speech errors (see e.g., Dell, 1986; MacKay, 1982; Oomen & Postma, 2001b), which greatly disrupt the fluency of speech when they are detected and repaired.

Another source of evidence relating timing and fluency is a study of older and younger speakers by Griffin and Spieler (2000). We discuss this in more detail because it demonstrates both the general idea behind variations in planning prior to speaking and it also supports the notion that older adults may, under some circumstances, prefer to do more planning of an utterance prior to speaking. The experimental task required speakers to insert the name of a pictured object into a sentence frame such as “They saw the object.” We manipulated the codability⁷ and frequency of the object name, two factors that have large and consistent effects on the speed to name an object in isolation (e.g., Bonin, Charlard, Meot, & Fayol, 2002; Goodglass, Theurkauf, & Wingfield, 1984; Lachman et al., 1974). Variations in when speakers prepared the final word of the utterance were reflected in effects of codability and frequency on speech onsets, durations, and disfluencies. Specifically, the more that speakers prepared object names prior to speech onset, the more likely that speech onset would show effects of codability and frequency. In contrast, the more speakers delayed preparing object names until after speech began, the more likely codability and frequency effects were to appear in speech durations and disfluencies rather than speech onset.

In the absence of any time pressure, both younger and older speakers prepared object names fully before speaking. While this seems surprising given how little advanced preparation speakers appear to do in other speaking tasks (e.g., Griffin, 2001; Griffin & Bock, 2000; Kempen & Huijbers, 1983; Smith & Wheeldon, 1999), recall that only one content word was generated for each sentence in this task, unlike most other sentence produc-

tion studies. With time pressure, younger adults began speaking much sooner after an object was displayed. Their speech onsets were independent of the objects’ codability and frequency, while there were large effects for these variables on their speech durations and disfluency rates. Basically, they started speaking before they prepared the single novel word in the sentence. In contrast, older adults continued to show large effects of codability and frequency on speech onset, but no effects of these factors on speech durations. Importantly, young and older adults were equivalent in overall fluency of the resultant utterance. Thus, the older adults showed a preference for preparing names prior to speech even under time pressure. This resulted in delayed speech onset for older compared to younger speakers, but once they began to speak, the utterances of older adults were as fluent as younger adults. Thus, in at least some circumstances, older adults prefer to engage in more preparation of their upcoming speech than do younger speakers (Griffin & Spieler, 2000).

So, speakers can and do modulate the content and timing of their speech and such variations impact its resultant fluency (Kemper et al., 2003). Moreover, we have preliminary evidence that there are circumstances in which older adults engage in more advance planning than younger adults do (Griffin & Spieler, 2000). However, advanced planning requires buffering prepared words until they can be articulated. When only one novel content word is produced in each utterance, such buffering should be minimal and very easy. As the number of novel words increases though, the buffering of prepared words should become more demanding (e.g., Peavler, 1974). But the demand on memory capacity only increases dramatically if speakers prepare all of the words in an utterance prior to beginning to articulate it. Buffering a second and a half of upcoming speech instead of one second could provide a cushion of time without greatly increasing the demands on working memory (see Griffin, 2003, for more on buffering).

In the present framework, we posit that greater vocabulary and naming specificity preferences may result in slower and less consistently successful word retrieval in older adults than in younger ones. Furthermore, the relatively faster word preparation of younger adults may allow them to engage in last-second name preparation during the production of messages that may be relatively complex. Slower and slightly more error prone word production processes in older adults may lead them to produce less complex messages with the benefit that this simplifies the timing of the resultant utterance. Differences in exposure to complex language may also create age-related differences in message content and fluency.

Any of these reasons could lead younger and older adults to differ in their modal solutions to the *what* and

⁷ Codability refers to the number of names available identifying a particular object or concept. It is estimated by observing the names that speakers give an object or picture in norming studies. In the usage here, a highly codable object has a single dominant name in norming studies whereas a medium codable object has multiple names, none of which is provided by 80% of normers. However, codability is not the same as name agreement. Often, a picture will have low name agreement because the object depicted is difficult to identify. For example, a picture meant to depict a cloud (Snodgrass & Vanderwart, 1980) is sometimes labeled as *mashed potatoes*. To avoid this, when manipulating codability, we attempt to choose pictures that elicit synonyms or names that may designate the same object, and match high and medium codable pictures on measures of object identification speed.

when of language production. The general goal for speakers remains the same, to convey messages reasonably clearly and fluently. In the remaining portion of this paper, we focus on the hypothesis that older adults may plan their words slightly further in advance relative to younger speakers in order to achieve fairly fluent speech. Importantly, the use of eye movement monitoring in addition to measures of speech timing allows researchers to distinguish between long latencies that are associated with slow word preparation and those that are due to preparing more words prior to speaking.

5. An eye-tracking study of younger and older speakers

Now we will review an eye tracking experiment in which we examined the influence of aging on the production of words within simple sentences. In doing so, we can observe how age-related changes revealed in experimental studies of single word production exert an influence in more complex production situations. During normal speech, speakers produce words that vary in the ease with which they can be selected, phonologically encoded, and articulated. There are at least three things that speakers may do to avoid producing an overt disfluency during speech as a result of retrieval difficulty. First, speakers may plan their speech far enough in advance to allow time to recover from momentary difficulties in retrieval. Second, speakers may modulate their speech rate, slowing articulation when it becomes apparent that upcoming words are not yet ready for production (e.g., Bell et al., 2003). Third, speakers may insert unnecessary but easily retrieved words (Ferreira & Dell, 2000; Ferreira & Firato, 2002). All of these adaptations allow more time for the completion of word retrieval and reduce the probability of overt disruptions in speech without sacrificing the specificity of the content conveyed.

In one study, we examined the extent to which the speech of young and older adults was responsive to the difficulty of upcoming words when content could not be adjusted (Spieler & Griffin, *in press*). Such external control over content made this production situation very different from natural speech. However, we assumed that the influence of planning and word retrieval on speech onset and timing would primarily reflect general language production processes rather than processes entirely idiosyncratic to the particular task. Similar word preparation patterns appear in eye movement studies of scene descriptions (Griffin & Bock, 2000) and card matching dialogues (Griffin & Garton, 2003; Horton, Metzling, & Gerrig, 2002), which indicate that the behaviors occur when speakers have more control over the content of their utterances, when the utterances are more complex, and when they speak to someone else to achieve a goal.

The younger adults who participated in the experiment were undergraduates at Stanford University

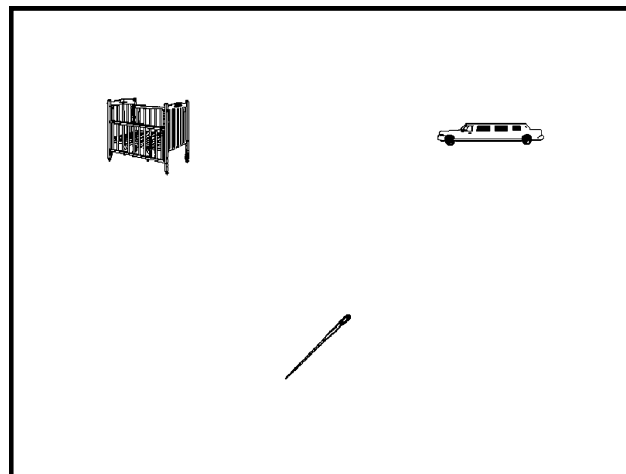


Fig. 2. An example of a stimulus from Spieler and Griffin (*in press*), which could be described as The crib and the limousine are above the needle. The limousine (limo) is a medium codable, low frequency item.

whereas the older adults were alumni and former staff. As a result, the groups were similar along many dimensions (see Spieler & Griffin, *in press*).⁸ Eye movements were monitored with an ISCAN 400 remote tracker that sat on a table between each participant and the computer monitor that presented stimuli. Participants leaned their foreheads against a forehead rest to prevent movements in depth that would reduce the ability of the eye tracker to report where they were looking. Speakers viewed an array of three objects as shown in Fig. 2 and described the array by inserting the object names into the sentence frame, “The A and the B are above the C.” One of the three objects was always the *critical object*, in which we manipulated the difficulty of word production along two dimensions. Difficulty was manipulated by varying the codability (high vs. medium codable) and the frequency (high vs. low frequency) of the object label. Both of these factors influence the speed with which these object labels are produced (Lachman, 1973; Oldfield & Wingfield, 1965), but not necessarily the time to identify objects (Johnson, 1992; Wingfield, 1967, 1968).

To examine the scope of planning prior to speech, we manipulated the position of the critical object. On half of the trials, the critical object appeared in the second (B) position while in the other half of the trials it appeared in the third (C) position. The critical object switched positions with a highly codable and repeatedly presented object. If speakers spoke as soon as the first object name was prepared, characteristics of the second object would have no influence on speech onset latencies. An

⁸ There were 15 adults in each group. The mean age of the younger group was 19.9 years and 74.5 for the older group. All were native speakers of American English, in good health, and with normal or corrected-to-normal vision. They did not differ significantly in WAIS vocabulary scores, but the younger participants showed better performance on the WAIS Digit-Symbol test.

intermediate amount of planning (e.g., the full subject noun phrase; Martin & Freedman, 2001; Smith & Wheeldon, 1999) would entail speech beginning once the first and the second object names were prepared. If so, speech onsets would reflect characteristics of the critical object when it was in the B position but not in C. If speakers prepared all the words in the utterance before speaking, the difficulty in naming the critical object would affect speech onset even when it occupied the C position.

The previous predictions assume that speakers begin speaking once they have completely prepared a speech unit of some size (content word, subject noun phrase, or clause). However, preparation of an object's name may begin before speech onset without speech onset being contingent on the complete preparation of the name. By monitoring speakers' eye movements before and during speech, we could roughly assess the preparation time allocated to the B and C objects independent of speech onset. If speakers only gazed at object B or C after starting to speak, it would suggest that they did not prepare these objects names before speaking. Observing codability and frequency effects only on gaze measures after speech began would strengthen this inference. In effect, the use of eye movement monitoring in language production constrains inferences about which processes speakers may carry out before and during speech.

We first examined what speakers did before the start of the first object's name. Older speakers took about 170 ms more time than younger adults to begin naming the first object.⁹ Analysis of the amount of time speakers spent gazing at first objects before naming them revealed a similar pattern. Older adults spent about 150 ms longer gazing at first objects than young adults, consistent with slightly slower word preparation times. In addition, older adults produced first object names fluently¹⁰ on fewer trials than the younger adults did. The only property of the first object that was manipulated was the frequency of its name. The frequency of the first object's name should have influenced the time to begin speaking regardless of the scope of planning other words, and it did. Both groups showed large and numerically similar effects of the frequency of the first object's name on its onset time and the amount of time spent gazing at it. Speakers produced high frequency first names more fluently than low frequency ones. This pattern of results for timing and eye movements for the first object suggest that older speakers took longer to prepare names for first objects than younger speakers did.

However, before concluding that old speakers are slower and less fluent than younger ones, we must note

that there was a small but significant difference in the extent to which younger and older speakers agreed on the names for these first objects. Young adults produced the object's dominant names a mean 94.6% of trials, whereas older adults produced the dominant names a mean 88.3% of trials. Although this 6.3% difference may seem minor, it completely accounted for the age differences in the timing of first object names and the time spent gazing at first objects. Specifically, when differences in name agreement were taken into account, the difference in when younger and older speakers began articulating the names of first objects was a non-significant 84 ms and gaze times, a non-significant 42 ms. Interestingly, the older adults remained marginally less fluent in producing the names. This example highlights the complexity of testing for age-related differences in language production even when the content of speech is relatively controlled via picture description tasks. In the following analyses for the critical object, age differences remained significant when name agreement was taken into account.

The important issue was whether the onset of speaking and thereby the onset of the first object's name was also influenced by characteristics of the second or third object's name. It was not, suggesting that speakers did not delay beginning their utterance in order to complete preparation of names for second or third objects. However, the fluency of the first object's name was affected by whether the second object was a critical object rather than a repeated one. Also, the frequency of the first object's name and the position of the critical object affected the fluency of older speakers more than that of younger speakers. This suggests that speakers were beginning to identify the second object while articulating the determiner *the* before the first object name. To more closely address the issue of preparation of the second and third object names, we examined speakers' gazes on the objects prior to the first object's name.

When the critical object was in the second position, young and old speakers gazed at it for less than 150 ms before starting to name the first object. This amount of time is enough for object identification (e.g., Potter, 1975), but it is unlikely to suffice for word selection, especially for medium codable objects. Speakers spent essentially no time on the third object, less than 10 ms on average. The codability and frequency of the critical object's name did not modulate the time speakers spent gazing at them before uttering first object names and there was no difference in gaze times between younger and older adults. Altogether these results suggest that both young and older speakers began identifying second objects immediately before producing first object names, hesitating more often when second objects were unfamiliar (cf. Morgan & Meyer, 2005). Speakers began preparing names for second objects only as they began to articulate first object names. Particularly important is

⁹ As in the literature review earlier, all differences that are discussed were statistically significant unless otherwise indicated.

¹⁰ A name was considered disfluent if it was preceded by a noticeable pause (over 200 ms), a filler ("um"), a false start, a stressed article (see Fox Tree & Clark, 1997), or if the noun was corrected.

the observation that, older adults began preparing second object names no earlier than younger speakers did relative to the onset of first object names. That is, older speakers did not prepare their speech further in advance than younger speakers did.

Analyses of speech timing and eye movements after the first noun support the conclusion that second and third object names were prepared during speech. Older speakers tended to take more time from the beginning of the first noun to the beginning of the critical name and spent 318 ms longer than younger ones gazing at critical objects. Because older speakers carried out no greater preparation before first nouns than younger speakers did and their speech content was the same, these longer intervals between first and critical nouns for older speakers had to be occupied either with words articulated more slowly or disfluencies such as pauses. Indeed, the age groups differed significantly in the proportion of fluent critical names they uttered. Younger speakers uttered 79% of their critical nouns fluently whereas older adults uttered only 65% fluently.

The lag between nouns and the time during speech gazing at critical objects was influenced by codability and frequency for younger and older speakers. Both measures were particularly long when the critical object was medium codable with a low frequency name. However, none of these stimulus effects interacted with age. In other words, the age groups showed comparable effects of the difficulty of preparing critical objects on the timing of their speech. However, older speakers were particularly disfluent when critical objects were medium codable and low frequency. For easily prepared words, older speakers were roughly 10% slower and less fluent, but the difference increased to 25% when words were difficult to prepare due to lower codability and frequency. Having already produced the first nouns, no measures were affected by first noun frequency.

The primary finding is the very strong consistency in the global aspects of timing for the two groups despite the somewhat slower speed and lower fluency for the older adults. To illustrate this consistency, we have plotted for every 8 ms the proportion of trials when speakers gazed at the critical object in position B relative to the onset of the first object's name and its name in the speakers' utterances (Fig. 3). Relative to first object names, older speakers did not begin gazing at critical objects any earlier than the younger adults did. In other words, older speakers planned and buffered their words no more than younger speakers did. However, they did take more time to prepare the names of critical objects, so relative to the onset of their names, older adults began gazing at the objects consistently earlier than the younger adults did. The longer it took to prepare the name of the critical object, the more likely speakers were to be disfluent. As a result, older speakers were often disfluent; they did not have the words they wanted to say ready when they wanted them.

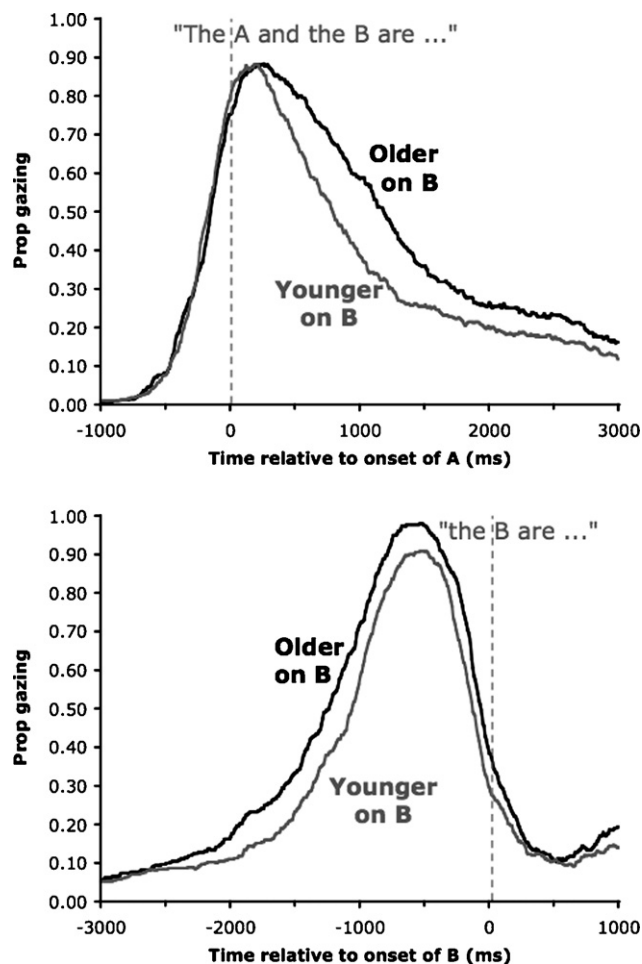


Fig. 3. Grand proportion of trials for which older and younger speakers were fixated on the critical object in B position for every 8 ms relative to the onset of the first object's name, A, (top) and relative to the critical object's name, B, (bottom). Data from Spieler and Griffin (in press).

Recall that we initially suggested that variations in advanced word preparation or message content might allow older adults to minimize the influence of age-related changes in word production speed and success. Here, however, the content of the utterances of younger and older adults was highly constrained, allowing no opportunity to re-structure utterances. Also, unlike the earlier experiment in which speakers named single objects in each utterance (Griffin & Spieler, 2000), an increased degree of planning would require speakers to buffer unrelated content words, which is far more effortful than buffering repeated function words or thematically related content words (Stine & Wingfield, 1987). Older adults in the present study did not prepare words further in advance in order to maintain a high level of fluency. In effect, the results of the experiment suggest that making no adjustments to the what and when of word preparation results in disfluent speech in older speakers. What remains to be shown is whether the reductions in utterance content and complexity that are

sometimes observed in the speech of older adults may serve to reduce these differences in fluency. The existing literature on speech samples and preliminary results from additional studies in our laboratories suggest that such adjustments of the *what* of production help the *when*. That is, altering content allows speakers get the words they want when they want them.

6. Looking back and ahead

Increases in word retrieval difficulty with age do not dramatically slow the rate of word production in spontaneous speech (e.g., Spieler et al., 2004). Indeed, young and older adults have considerable control over their speech rate and can adaptively vary it depending on context and other factors (e.g., Griffin & Spieler, 2000; Kemper et al., 2003). The scope of word preparation prior to utterance onset influences the speed and the fluency of speech. Furthermore, vocabulary and name specificity preferences may make word retrieval slower and less often successful for older adults, regardless of any effects of aging per se. Likewise, social goals and exposure to academic language also varies with age. All of these considerations make the study of age-related changes in language production particularly challenging and foil any attempt to account for them with a single factor.

The use of eye movement monitoring in the study of language production, with its ability to track the preparation of words, can reveal processing strategies such as slightly increased scope of preparation that may allow older adults to alleviate the influence of word retrieval differences on production. Likewise, with stimuli that permit variations in the message content of each utterance, eye movement data may provide insight into why content varies with age. However, eye movement monitoring is not a silver bullet with the capability to slay the complexities of research in language production and aging. As with any method, eye movement monitoring has its limitations that production researchers are beginning to explore. With respect to age, there is a striking one. We know that aging is associated with changes to vision and, critically, useful field of view decreases with age (see Irwin, 2004). As a result, older speakers may be less able to extract information about the identity of objects before moving their eyes to fixate on them. Thus, age-related differences for objects that are initially viewed in the periphery such as the critical objects in Spieler and Griffin (in press) may be accounted for by differences in parafoveal preview. A comprehensive account of aging and language production is only possible through converging methods and consideration of the many factors that influence language production.

Aside from informing the literature on cognitive aging, age-related changes in language production provide an ideal forum within which to explore intra- and

inter-individual, voluntary and involuntary differences in language production processes. Bringing insights from language production research to bear on questions of age-related changes promises to illuminate both areas of research. In particular, such comparisons highlight the ability of speakers to vary the *what* and *when* of their speech in ways that models of language production have barely begun to address.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.bandl.2005.08.003](https://doi.org/10.1016/j.bandl.2005.08.003).

References

- Altmann, L. J. P., & Kemper, S. (in press). Effects of age, animacy and activation order on sentence production. *Language and Cognitive Processes*.
- Altmann, L. J. P., Kemper, S., Mathews, A., & Mullin, D. A. (2004, April). Syntactic priming in older adults. In *Poster presented at the 10th cognitive aging conference, Atlanta GA*.
- Alwin, D. F., & McCammon, R. J. (2001). Aging, cohorts, and verbal ability. *Journals of Gerontology: Series B: Psychological Sciences & Social Sciences*, 56, S151–S161.
- Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal thee, um, new information. *Journal of Psycholinguistic Research*, 32, 25–36.
- Au, R., Joungh, P., Nicholas, M., & Obler, L. K. (1995). Naming ability across the adult life span. *Aging & Cognition*, 2, 300–311.
- Bachoud-Levi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, 39, 331–346.
- Badecker, W., Miozzo, M., & Zanuttini, R. (1995). The two stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, 57, 193–216.
- Barresi, B. A., Nicholas, M. T., Connor, L., Obler, L. K., & Albert, M. L. (2000). Semantic degradation and lexical access in age-related naming failures. *Aging, Neuropsychology, & Cognition*, 7, 169–178.
- Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America*, 113, 1001–1024.
- Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1–47.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.
- Bock, J. K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin and Review*, 34, 395–421.
- Bock, K. (1995). Sentence production: From mind to mouth. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (Vol. 11). *Handbook of perception and cognition* (pp. 181–216). Orlando, FL: Academic Press.
- Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego: Academic Press.
- Bock, K., Irwin, D. E., & Davidson, D. J. (2004). Putting first things first. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 249–278). New York: Psychology Press.

- Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, 48, 653–685.
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129, 177–192.
- Bonin, P., Chalard, M., Meot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, 93, 89–114.
- Borod, J. C., Goodglass, H., & Kaplan, E. (1987). Normative data on the Boston diagnostic aphasia examination, parietal lobe battery, and the Boston naming test. *Journal of Clinical Neuropsychology*, 2, 209–215.
- Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, 44, 123–147.
- Bowles, N. L., Obler, L. K., & Albert, M. L. (1987). Naming errors in healthy aging and dementia of the Alzheimer type. *Cortex*, 23, 519–524.
- Bowles, N. L., Obler, L. K., & Poon, L. W. (1989). Aging and word retrieval: Naturalistic, clinical, and laboratory data. In L. W. R. Poon & C. David (Eds.), *Everyday cognition in adulthood and late life* (pp. 244–264). New York, NY, US: Cambridge University Press.
- Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Review*, 109, 204–223.
- Brown, A. S., & Gollan, T. H. (November, 2003) Does aging really increase TOTs? It depends on the measure. *Poster presented at the 44th annual meeting of the Psychonomic Society*, Vancouver, Canada.
- Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325–337.
- Burke, D. M., Locantore, J. K., Austin, A. A., & Chae, B. (2004). Cherry pit primes Brad Pitt: Homophone priming effects on young and older adults’ production of proper names. *Psychological Science*, 15, 164–170.
- Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory & Language*, 30, 542–579.
- Buswell, G. T. (1935). *How people look at pictures*. Chicago, IL: University of Chicago Press.
- Butterworth, B. (1989). Lexical access in speech production. In W. D. Marslen-Wilson (Ed.), *Lexical Representation and Process* (pp. 108–135). Cambridge, MA: MIT Press.
- Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, 98, 67–83.
- Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive Science*, 93, 1–43.
- Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, 29, 217–229.
- Christenfeld, N. (1995). Does it hurt to say um? *Journal of Nonverbal Behavior*, 19, 171–186.
- Clark, H. H. (2002). Speaking in time. *Speech Communication*, 36, 5–13.
- Clark, H. H., & Fox Tree, J. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84, 73–111.
- Cooper, P. V. (1990). Discourse production and normal aging: Performance on oral picture description tasks. *Journal of Gerontology*, 45, 210–214.
- Costa, A., & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding in speech production. *Journal of Memory and Language*, 46, 178–198.
- Dahlgren, D. J. (1998). Impact of knowledge and age on tip-of-the-tongue rates. *Experimental Aging Research*, 24, 139–153.
- Davidson, D. J., Zacks, R. T., & Ferreira, F. (2003). Age preservation of the syntactic processor in production. *Journal of Psycholinguistic Research*, 32, 541–566.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, 23, 517–542.
- Dell, G. S., & O’Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42, 287–314.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in normal and aphasic speakers. *Psychological Review*, 104, 801–838.
- Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. Wheeldon (Ed.), *Aspects of Language Production* (pp. 291–330). London: Psychology Press.
- Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40, 296–340.
- Ferreira, V. S., & Firato, C. E. (2002). Proactive interference effects on sentence production. *Psychonomic Bulletin and Review*, 9, 795–800.
- 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.
- Feyereisen, P. (1997). A meta-analytic procedure shows an age-related decline in picture naming: Comments on Goulet, Ska, and Kahn (1994). *Journal of Speech & Hearing Research*, 40, 1328–1333.
- Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, 34, 709–738.
- Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing ‘the’ as ‘thee’ to signal problems in speaking. *Cognition*, 62, 151–167.
- Freedman, M. L., Martin, R. C., & Biegler, K. (2004). Semantic relatedness effects in conjoined noun phrase production: Implications for the role of short-term memory. *Cognitive Neuropsychology*, 21, 245–265.
- Fromkin, V. (1968). Speculations on performance models. *Journal of Linguistics*, 4, 1–152.
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 133–177). New York: Academic Press.
- Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: the Cambridge survey, III: Language: Psychological and biological aspects* (pp. 69–96). Cambridge: Cambridge University Press.
- Giles, H., Fox, S., & Smith, E. (1993). Patronizing the elderly: Intergenerational evaluations. *Research on Language & Social Interaction*, 26, 129.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.
- Gollan, T. H., & Acenas, L.-A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish–English and Tagalog–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30, 246–269.
- Goodglass, H., Theurkauf, J. C., & Wingfield, A. (1984). Naming latencies as evidence for two modes of lexical retrieval. *Applied Psycholinguistics*, 5, 293–303 (or 135–145).
- Gordon, J., & Dell, G. (2002). Learning to divide the labor between syntax and semantics: A connectionist account of deficits in light and heavy verb production. *Brain and Cognition*, 48, 376–381.
- Goulet, P., Ska, B., & Kahn, H. J. (1994). Is there a decline in picture naming with advancing age? *Journal of Speech & Hearing Research*, 37, 629–644.
- Gregory, M., Joshi, A., & Sedivy, J. C. (2003, March). Adjectives and processing effort in production: So, uh, what are we doing during disfluencies? In *Paper presented at the 16th annual CUNY conference on human sentence processing*, Cambridge MA.
- Griffin, Z. M. (1998). *What the eye says about sentence planning*. Unpublished Dissertation, University of Illinois at Urbana-Champaign, Champaign.

- Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, 82, B1–B14.
- Griffin, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin and Review*, 10, 603–609.
- Griffin, Z. M. (2004). Why look? Reasons for eye movements related to language production. In J. M. Henderson & F. Ferreira (Eds.), *The Interface of Language, Vision, and Action: Eye movements and the visual world* (pp. 213–248). New York: Psychology Press.
- Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11, 274–279.
- Griffin, Z. M., & Garton, K. L. (2003, March). Procrastination in speaking: Ordering arguments during speech. In *Paper presented at the 16th annual CUNY conference on human sentence processing*, Cambridge, MA.
- Griffin, Z. M., & Mouzon, S. S., (2004, March). Can speakers order a sentence's arguments while saying it? In *Poster presented at the 17th annual CUNY conference on human sentence processing*, College Park MD.
- Griffin, Z. M., & Oppenheimer, D. M. (2003, August). Looking and lying: Speakers' gazes reflect locus of attention rather than speech content. In *Paper presented at the 12th European conference on eye movements*, Dundee Scotland.
- Griffin, Z. M., & Spieler, D. H. (2000, November). Speech planning in younger and older adults. In *Poster presented at the 41st annual meeting of the Psychonomic Society*, New Orleans LA.
- Griffin, Z. M., & Spieler, D. H. (2004, April). Sentence complexity in event descriptions of younger and older speakers. In *Poster presented at the 10th cognitive aging conference*, Atlanta GA.
- Griffin, Z. M., & Weinstein-Tull, J. (2003). Conceptual structure modulates structural priming in the production of complex sentences. *Journal of Memory & Language*, 49, 537–555.
- Hale, S., & Myerson, J. (1995). Fifty years older, fifty percent slower? Meta-analytic regression models and semantic context effects. *Aging & Cognition*, 2, 132–145.
- Hale, S., & Myerson, J. (1996). Experimental evidence for differential slowing in the lexical and nonlexical domains. *Aging, Neuropsychology, & Cognition*, 3, 154–165.
- Henderson, J., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.
- Hertzog, C., & Schaie, K. W. (1988). Stability and change in adult intelligence: 2. Simultaneous analysis of longitudinal means and covariance structures. *Psychology and Aging*, 3, 122–130.
- Holsánová, J. (2001). *Picture viewing and picture description*. Unpublished Dissertation, Lund University, Lund.
- Horton, W. S., & Metzing, C. A., Gerrig, R. J. (2002, Nov.). Tracking speakers' use of internal and external information during referential communication. In *Paper presented at the 43rd annual meeting of the Psychonomic Society*, Kansas City, MO.
- Ingram, D. (1989). *First Language Acquisition: Method, description, and explanation*. New York: Cambridge University Press.
- Irwin, D. (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. 105–134). New York: Psychology Press.
- Jackendoff, R. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
- James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 26, 1378–1391.
- Johnson, C. J. (1992). Cognitive components of naming in children: Effects of referential uncertainty and stimulus realism. *Journal of Experimental Child Psychology*, 53, 24–44.
- Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: John Benjamins.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8, 441–480.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154, 1583–1585.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston naming test*. Philadelphia: Lea & Febiger.
- Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201–258.
- Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, 14, 185–209.
- Kemper, S. (1992). Language and aging. In F. I. M. Craik & T. A. Salt-house (Eds.), *The handbook of aging and cognition* (pp. 213–270). Hillsdale, NJ, England: Lawrence Erlbaum Associates.
- Kemper, S. (1993). Geriatric psycholinguistics: Syntactic limitations of oral and written language. In L. L. Light & D. M. Burke (Eds.), *Language, memory, and aging* (pp. 58–76). New York, NY, US: Cambridge University Press.
- Kemper, S., & Rash, S. J. (1988). Speech and writing across the life-span. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Clinical and educational implications* (Vol. 2). *Practical aspects of memory: Current research and issues* (pp. 107–112). Chichester: John Wiley.
- Kemper, S., Greiner, L. H., Marquis, J. G., Prenovost, K., & Mitzner, T. L. (2001). Language decline across the life span: Findings from the nun study. *Psychology and Aging*, 16, 227–239.
- Kemper, S., Herman, R. E., & Lian, C. H. T. (2003). The costs of doing two things at once for young and older adults: Talking while walking, finger tapping, and ignoring speech of noise. *Psychology & Aging*, 18, 181–192.
- Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and propositional content. *Psychology & Aging*, 16, 600–614.
- Kynette, D., & Kemper, S. (1986). Aging and the loss of grammatical forms: A cross-sectional study of language performance. *Language & Communication*, 6, 65–72.
- Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology*, 99, 199–208.
- Lachman, R., Shaffer, J. P., & Hennrikus, D. (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction time. *Journal of Verbal Learning and Verbal Behavior*, 13, 613–625.
- Lenneberg, E. H. (1967). *Biological foundations of language* (1st ed.). New York: Wiley.
- Levelt, W. J. M. (1981). The speaker's linearization problem. *Philosophical Transactions of the Royal Society of London B*, 295, 305–315.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology*, 12, 433–452.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Science*, 22, 1–45.
- Lindenberger, U., & Baltes, P. B. (1997). Intellectual functioning in old and very old age: Cross-sectional results from the Berlin Aging Study. *Psychology and Aging*, 12, 410–432.
- MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, 89, 483–506.
- MacKay, D. G., & Burke, D. M. (1990). Cognition and aging: A theory of new learning and the use of old connections. In T. M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 213–263). Oxford, England: North-Holland.
- Martin, R. C., & Freedman, M. L. (2001). Short-term retention of lexical-semantic representations: Implications for speech production. *Memory*, 9, 261–280.

- Martin, R. C., Miller, M., & Vu, H. (2004). Lexical–semantic retention and speech production: Further evidence from normal and brain-damaged participants for a phrasal scope of planning. *Cognitive Neuropsychology*, 21, 625–644.
- Maylor, E. A. (1990). Age, blocking and the tip of the tongue state. *British Journal of Psychology*, 81, 123–134.
- Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture–word interference experiments. *Journal of Memory and Language*, 35, 477–496.
- Meyer, A. S. (1997). Conceptual influences on grammatical planning units. *Language and Cognitive Processes*, 12, 859–863.
- Meyer, A. S., & Lethaus, F. (2004). The use of eye tracking in studies of sentence generation. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 191–212). New York: Psychology Press.
- Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, 48, 131–147.
- 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.
- Morgan, J. L., & Meyer, A. S. (2005). Processing of extrafoveal objects during multiple-object naming. *Journal of Experimental Psychology: Learning, Memory, And Cognition*, 31(3), 428–442.
- Mortensen, L., Meyer, A. S., & Humphreys, G. W. (in press). Age-related effects on speech production: A review. *Language and Cognitive Processes*.
- Miller, J., & Weinert, R. (1998). *Spontaneous Spoken Language: Syntax and discourse*. Oxford: Clarendon Press.
- Nicholas, M., Obler, L., Albert, M., & Goodglass, H. (1985). Lexical retrieval in healthy aging. *Cortex*, 21, 595–606.
- Nicholas, L. E., Brookshire, R. H., MacLennan, D. L., Schumacher, J. G., & Porrazzo, S. A. (1989). The Boston naming test: Revised administration and scoring procedures and normative information for non-brain-damaged adults. In T. E. Prescott (Ed.), *Clinical aphasiology* (Vol. 18, pp. 103–115). Austin TX: Pro-Ed.
- O'Connor, B. P., & St. Pierre, E. S. (2004). Older persons' perceptions of the frequency and meaning of elderspeak from family, friends, and service workers. *International Journal of Aging & Human Development*, 58, 197.
- Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17, 273–281.
- Oomen, C. C. E., & Postma, A. (2001a). Effects of divided attention on the production of filled pauses and repetitions. *Journal of Speech, Language, & Hearing Research*, 44, 997–1004.
- Oomen, C. C. E., & Postma, A. (2001b). Effects of time pressure on mechanisms of speech production and self-monitoring. *Journal of Psycholinguistic Research*, 30, 163–184.
- Oomen, C. C. E., & Postma, A. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes*, 17, 163–184.
- Pasupathi, M., Henry, R. M., & Carstensen, L. L. (2002). Age and ethnicity differences in storytelling to young children: Emotionality, relationality and socialization. *Psychology & Aging*, 17, 610–621.
- Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology*, 11, 559–566.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39, 633–651.
- Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General*, 113, 426–442.
- Potter, M. C. (1975). Meaning in visual search. *Science*, 187, 965–966.
- Rastle, K. G., & Burke, D. M. (1996). Priming the tip of the tongue: Effects of prior processing on word retrieval in young and older adults. *Journal of Memory & Language*, 35, 585–605.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Russo, J. E., & Rosen, L. D. (1975). An eye fixation analysis of multialternative choice. *Memory & Cognition*, 3, 267–276.
- Schmitter-Edgecombe, M., Vesneski, M., & Jones, D. W. R. (2000). Aging and word-finding: A comparison of spontaneous and constrained naming tests. *Archives of Clinical Neuropsychology*, 15, 479–493.
- Schroeder, D. H., & Salthouse, T. A. (2004). Age-related effects on cognition between 20 and 50 years of age. *Personality & Individual Differences*, 36, 393–404.
- Shriberg, E. E. (1994). Preliminaries to a theory of speech disfluencies. Unpublished Dissertation, University of California Berkeley, Berkeley CA.
- Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, 73, 205–246.
- Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language*, 32, 25–38.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.
- Spierer, D. H., & Griffin, Z. M. (in press). The influence of age on the time course of word preparation in multiword utterances. *Language and Cognitive Processes*.
- Spierer, D. H., Horton, W. S., Shriberg, E. E. (2004, April). Language use in the wild: Analyses of conversational speech across the lifespan. In *Paper presented at the 10th Cognitive Aging Conference, Atlanta GA*.
- Sternberg, S., Knoll, R. L., Monsell, S., & Wright, C. E. (1988). Motor programs and hierarchical organization in the control of rapid speech. *Phonetica*, 45, 175–197.
- Stine, E. L., & Wingfield, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology and Aging*, 2, 272–279.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634.
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74, 209–253.
- Van Der Meulen, F. F., Meyer, A. S., & Levelt, W. J. M. (2001). Eye movements during the production of nouns and pronouns. *Memory & Cognition*, 29, 512–521.
- Van Gorp, W. G., Satz, P., Kiersch, M. E., & Henry, R. (1986). Normative data on the Boston Naming Test for a group of normal older adults. *Journal of Clinical & Experimental Neuropsychology*, 8, 702–705.
- Vigliocco, G., Garrett, M. F., & Antonini, T. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8, 314–317.
- Vigliocco, G., Martin, R. C., & Garrett, M. F. (1999). Is “count” and “mass” information available when the noun is not? An investigation of tip of the tongue states and anomia. *Journal of Memory and Language*, 40, 534–558.
- Weiner, E. J., & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics*, 19, 29–58.
- Wheeldon, L. (2000). Generating prosodic structures. In L. R. Wheeldon (Ed.), *Aspects of Language Production* (pp. 249–274). London: Psychology Press.
- Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, 37, 356–381.
- Wheeldon, L. R., & Lahiri, A. (2002). The minimal unit of phonological encoding: Prosodic or lexical word. *Cognition*, 85, B31–B41.
- Wingfield, A. (1967). Perceptual and response hierarchies in object identification. *Acta Psychologica*, 26, 216–226.
- Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *American Journal of Psychology*, 81, 226–234.
- Wundt, W. (1900/1970). Die sprache. In A. L. Blumenthal (Ed.), *Language and psychology: Historical aspects of psycholinguistics*. New York: Wiley (first published 1900 ed., Vol. 1).