LANGUAGE AND COGNITIVE PROCESSES
2006, 21 (1/2/3), 291–321

Psychology Press
Taylor & Francis Group

The influence of age on the time course of word preparation in multiword utterances

Daniel H. Spieler and Zenzi M. Griffin
Georgia Institute of Technology, Atlanta, GA and Stanford University, Stanford, CA, USA

Sentence production requires speakers to co-ordinate the preparation of words so that they are ready for articulation when they are needed. Ageing appears to influence both the speed and likelihood of successful word retrieval. We examine how age differences in word production might influence the production of larger units of speech such as sentences. Speakers described displays containing three objects of systematically varied naming difficulty. The latency, duration, content, and fluency of speech in addition to its co-ordination with eye movements indicated that both young and older adults prepared their words immediately before uttering them. As a consequence, older adults were also significantly less fluent in their utterances than were younger adults.

The ease of producing speech belies the daunting complexity of the processes involved. Words must be selected in a grammatical sequence from a working lexicon for an adult of tens of thousands of words. The sounds of these words must be assembled into prosodic units, and the motor programs for articulation must be executed quickly and accurately. Most impressively, this must occur with enough speed to accommodate a

Correspondence should be addressed to either author at the School of Psychology, 654 Cherry Street, Georgia Institute of Technology, Atlanta, GA 30332-0170, USA.
Email: daniel.spieler@psych.gatech.edu or zenzi.griffin@psych.gatech.edu.

This research was supported by National Institute on Aging Grant RO1 AG17024. We would like to acknowledge Amit Mookerjee for assistance in software development, to Ellen Hamilton, Edna Chavira, and Nadine de Lassus for the substantial effort involved in data collection, and to Justin Storbeck, Jordan Wang, and Russell Smith for transcriptions and speech measurements. We would also like to thank Elisa Lawler and Tracy Mitzner for comments on an earlier version of this manuscript and Demetria Mitchell for assistance in manuscript preparation. Thanks also to the Cognitive Aging community at Georgia Tech for invaluable discussion.

http://www.tandf.co.uk/journals/pp/01690965.html DOI: 10.1080/01690960400002133

normal conversational speech rate of 2–3 words per second or about 5.65 syllables per second (see Bock, 1995, for review).

Along with the complexity of producing speech comes a certain amount of flexibility. We are particularly interested in adaptations that speakers may take to maximise the fluency of their utterances. That speakers can make such adaptations is beyond doubt and should be apparent to anyone with experience in answering questions in public speaking settings (e.g., those after a colloquium). In such a case, one often has the sense of engaging in far more extensive planning of the upcoming utterance than would be the case if the same question were asked while both speaker and questioner were at a table in a pub. To ensure the desired level of fluency, speakers must be able to make adjustments to the language production process in response to variations in the ease of the production task and the desired level of fluency.

While speakers may be able to adjust production in a number of ways, we focus here on the timing of word production processes. In particular, there are two adjustments that speakers might make that would allow for the maintenance of fluency in the face of variations in the ease of word retrieval. First, speakers may elect to modulate the number of words prepared prior to the onset of speech. Increasing the amount of an utterance that a speaker prepares prior to speaking provides a cushion to continue producing speech in the face of a momentary difficulty in the retrieval of unprepared words. Second, once speech has begun, speakers may make online and opportunistic adjustments of speech rate during an utterance, slowing speech in the moments prior to the production of words that are particularly difficult to retrieve (e.g., Jurafsky, Bell, Gregory, & Raymond, 2001). When a speaker fails to make such adjustments, the speaker becomes disfluent, emitting either filled pauses (''uh'' or ''um''), silent pauses, or other forms of disfluency (word repetitions, restarts, etc.) while completing word retrieval.

One possible reason that speakers may need to make adjustments is to accommodate not only momentary problems in production but also to accommodate the changes that occur to component language production processes as a result of ageing. Below we review evidence that some aspects of the component processes involved in language production change across the lifespan. In the present experiment, we compare the timing of word retrieval in the sentence production of young and older adults. Ageing provides a context where a general adaptation such as a modulation in the amount of word preparation may be an adaptive strategy to accommodate to age-related changes in the speed and reliability of component production processes. Because speakers may also modify the timing of speech in response to local variations in word retrieval, we also manipulated two stimulus properties that influence its difficulty to observe how young and older adults make such online modifications to the timing of production.

In the following sections, we first provide a basic outline of the language production system. We then review evidence for age-related changes to the specific processes of production and discuss how such changes might be evidenced in more complex production situations such as sentence production. Finally, we discuss the specific methods that we use for examining the scope of word preparation in speech and modulations in production in response to the difficulty of the production process.

LANGUAGE PRODUCTION PROCESSES

Language production is divided into multiple levels of processing. Production begins with an abstract message and ends with the execution of a motor program. Evidence for multiple levels in-between comes from analyses of speech error distributions and experimental studies of production (see Dell, 1995, and Levelt, 1999, for reviews). Speakers begin by formulating an intention to communicate, which is called a message (Fromkin, 1968). This message contains the conceptual and pragmatic information that the eventual utterance should express. Next, speakers select lexical representations (lemmas) that fit the semantic and pragmatic specifications of message elements. The general view is that these lemmas do not contain any phonological information (Kempen & Huijbers, 1983; but see Caramazza, 1997). The speed of lemma selection depends on the strength of the constraints guiding it and the number of lemmas competing to express the message element (e.g., Griffin & Bock, 1998; Lachman, 1973). For example, the presence of multiple possible names for an object (e.g., *television* or *TV*) will generally slow the lemma selection process compared with an object with a single appropriate name (e.g., *apple*).

Lemma selection is followed by phonological encoding. Phonological encoding involves associating the metrical structure (e.g., the verb lemma for *export* goes with a two-syllable frame with stress on the second syllable) with the word's phonemes. Finally, the resulting representations are used to retrieve motor programs for articulating syllables (see Levelt, Roelofs, & Meyer, 1999, for a theory of phonological encoding).

Speakers may vary the number of content words in an utterance that are prepared prior to speech (e.g., Griffin & Bock, 2000; Wheeldon & Lahiri, 1997). That is, the completion of phonological encoding of a word does not require the speaker to immediately say it. Rather, speakers appear able to vary the lag between the completion of phonological encoding and speech. In some cases, articulating a word may follow closely on the heels of the phonological encoding of it (e.g., Griffin, 2001) and in others, speakers may buffer phonological representations, preparing speech further in advance

(Ferreira & Swets, 2002; Griffin, 2003; Wheeldon & Lahiri, 1997). Note that the riskier strategy is one in which articulation closely follows phonological encoding because any delay in retrieving the next word may leave the speakers without anything new to say. A safer strategy would involve phonologically encoding as much as an entire sentence prior to speech onset. Many older theories of sentence production held that words in fluent utterances were selected in advance of speaking (e.g., Butterworth, 1989). This was based on word exchange errors and disfluency patterns, both of which have viable alternative accounts assuming last-second word preparation (see e.g., Dell, Chang, & Griffin, 1999; Griffin, 2001). Nonetheless, increasing the number of words retrieved before speaking is associated with later speech onsets and faster and more fluent production of the final utterance compared with a more incremental production strategy (Griffin & Bock, 2000).

Variations in advance word retrieval may be extremely useful for reducing the impact of age-related changes in specific processes in language production (Griffin & Spieler, 2000). When younger and older adults produced sentences under time pressure that each contained a single novel content word, speech latencies showed that the older adults engaged in more preparation of the content words prior to speech. In this study, they were no more disfluent than younger speakers. Below we review why older adults may elect to begin preparing words further in advance of saying them and what age differences in language production might look like depending on whether such an adaptation is selected or not.

AGE DIFFERENCES IN LANGUAGE PRODUCTION

Experimental studies that have attempted to isolate component processes of production have revealed a few consistent patterns of age-related change. Picture naming is a task that requires the selection of an appropriate picture name (i.e., lemma selection), phonological encoding, and articulation, in addition to the visual processing and identification of the object. Several studies have shown a small but consistent age-related increase in the time to produce picture names (Bowles, 1993; Ramsay, Nicholas, Au, Obler, & Albert, 1999) and in the probability of errors (Borod, Goodglass, & Kaplan, 1987; Feyereisen, 1997; Kaplan, Goodglass, & Weintraub, 1983; Nicholas, Brookshire, MacLennan, Schumacher, & Porrazzo, 1989; Van Gorp, Satz, & Kiersch, 1986; but see Goulet, Ska, & Kahn, 1994). Studies by Burke and colleagues also show an age-related increase in tip of the tongue (TOT) states (e.g., Burke, MacKay, Worthley, & Wade, 1991; James & Burke, 2000). In TOT states, speakers report a strong subjective sense of knowing the word, but they are unable to say it (Brown, 1991; Brown & McNeill, 1966). They are often able to report characteristics of the word form such as its initial phoneme and number of syllables, but they are unable to access all of its sounds. The partial information speakers retrieve suggests that lemma selection has occurred in TOT states, but not full access to the phonological form. Burke and colleagues have shown that older adults are consistently more likely to enter such TOT states. Their studies of phonological cueing support the argument that age-related increases in TOTs result from the weakening of connections between the lemma level and phonological information (Burke, Locantore, Austin, & Chae, 2004; James & Burke, 2000).

PRESENT STUDY

The goal of the present study is to begin to examine the influence of ageing on the production of simple sentences. In doing so, we can observe how age-related changes revealed in experimental studies of word retrieval may 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, encoded, and articulated. To avoid the inferential difficulties that observational studies encounter (e.g., changes in content that alone might result in changes in timing), the experimenters rather than the speakers determined the content of the utterances. Clearly such external control over content makes this production situation different from natural speech. However, we assume that the influence of word retrieval on speech onset and timing primarily reflect general language production processes rather than processes entirely idiosyncratic to the particular task. Consistent with this view, similar evidence for preparing words in the second before uttering them appears in scene descriptions (Griffin & Bock, 2000) and in card matching dialogues (Horton, Metzing, & Gerrig, 2002).

In this experiment, an array of three objects such as shown in Figure 1 was presented. Speakers were instructed to use a single sentence frame, inserting the object names into the sentence ‘‘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 lemma selection and phonological encoding. Difficulty was manipulated by varying the name agreement (codability[1] hereafter) and the frequency of the object name. Both of these factors influence the speed with which these object names are retrieved (Goodglass, Theurkauf, & Wingfield, 1984; Lachman, 1973; Oldfield & Wingfield, 1965; Paivio, Clark, Digdon, & Bons, 1989), but not the time to identify

[1] The term codability is used to highlight the fact that the objects are readily identified but differ in the number of names they elicit. Moderate name agreement alone can be due to the same object having multiple appropriate names or being very difficult to identify.

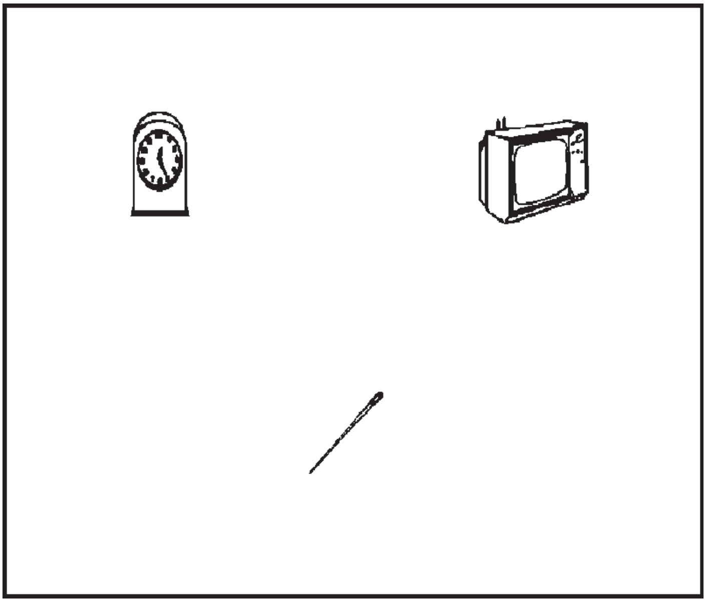

Figure 1. Sample array of three objects from which a subject might say "The clock and the television are above the needle".

objects (Johnson, 1992; Wingfield, 1967, 1968). In particular, lower word frequency is associated with phonological encoding errors (Dell, 1990; Kelly, 1986) and TOT states (Harley & Bown, 1998). Based on the idea that connections between lemmas and phonological information are particularly impaired with age (Burke et al., 1991), one would expect older adults to show greater difficulty with low frequency names relative to younger adults. Moreover, when a selected lemma suffers competition from other appropriate names (e.g., when objects are relatively lower in codability), the frequency of its form may play a still greater role in successful word production (Griffin & Bock, 1998).[2] We should note that the locus or loci of word frequency effects is a highly controversial topic in word production (see Caramazza, Costa, Miozzo, & Bi, 2001; Jescheniak, Meyer, & Levelt, 2003).

To examine the scope of preparation prior to and during speech, we manipulated the position of the critical object. On half of the trials, the

[2] An interaction between codability and frequency would be predicted by most interactive activation theories under these assumptions. However, it is unclear under which conditions activation from earlier levels of representation may facilitate the retrieval of phonological information in Node Structure Theory. For example, Rastle and Burke (1996) predicted and found that earlier semantic processing of words did not affect TOT rates, although one would expect extra activation of semantic representations to be passed on to phonological ones.

critical object appeared in the second (B) position while in the other half of the trials it appeared in the third (C) position. Varying the position of the object makes it possible to test whether speakers modulate the time to initiate speech or their speech rate depending on when the difficulty is encountered. If advanced word preparation is minimal, characteristics of the B object will have no influence on the time to initiate speech because speech onset may begin as soon as the first object name (hereafter referred to as object A) is available. An intermediate amount of word preparation might entail speech beginning once the first and the second object names (Objects A and B) both become available. If so, speech onsets will reflect characteristics of the critical object when in the B position but not in the third (Object C). Finally, the maximal amount of advanced word preparation would result in speech onset occurring only after all object names became available, which would entail effects of the critical object on speech onset even when it occupied the C position.

Note that this line of reasoning treats word preparation as all or nothing. If preparation includes Object B, speech onset will be influenced by the properties of B's name. If it is not influenced by B's name, then one might infer that preparation has not extended to the object B. However, it is entirely possible that speakers begin preparation of B and even C object names prior to speech onset, but that speech onset is ultimately determined only when the A object name becomes available. In this case, pre-speech preparation includes upcoming objects but only the availability of the object A's name delays speech onset. What is needed is a method for assessing the amount of processing allocated to the B and C objects independent of speech onset. We did this by monitoring speakers' eye movements over the object array in the time leading up to and during speech. In doing so, we took advantage of findings from several studies that show a close temporal relationship between eye movements over a picture and speakers' verbal descriptions of these pictures (Griffin, 2001; Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998; Meyer & Van der Meulen, 2000).[3] These experiments with young college age speakers typically show

[3] The degree of incrementality observed in these studies differs from that argued for in other similar non-eye tracking experiments such as Smith and Wheeldon (1999) and Martin, Miller, and Vu (2004). Those studies concluded that entire subject noun phrases were encoded prior to speech onset based on faster onset latencies for sentences such as ''The A moves above the B and C'' relative to ''The A and B move above the C''. While such results could be considered evidence that two nouns were encoded before speech onset in the latter sentence but not in the former, the small latency difference (under 100 ms) and the lack of any evidence for word specific processing of B object names suggest instead that some other aspect of preparing complex subject noun phrases delayed speech onset slightly or that preparation of B was begun but not completed (as in Griffin, 2001, 2003).

that the eyes land on an object that will be mentioned in the description about 1 second prior to the production of the object name in the description. The amount of time speakers spend gazing at an object that they name reflects the amount of time it takes to prepare a name for it in single object naming. These results suggest a preference for minimal advanced preparation of words by these speakers. The present study extends these results and asks whether age-related changes in the language production processes lead older speakers to prepare different numbers of words prior to and during speech. A portion of the results from the younger adults can be found in Griffin (2001).

Methods

Participants. Seventeen younger adults aged 18–23 years ($M = 19.9$, $SD = 1.8$) were recruited from the undergraduate student population at Stanford University. Seventeen adults aged 60–80 ($M = 74.5$, $SD = 6.7$) were recruited from the Palo Alto community. Not surprisingly, because the younger adults had not yet completed their schooling, younger adults had 13.9 (1.2) years of education whereas the older adults had 16.2 (1.5) years of education. The older adult participants included a high proportion of Stanford University alumni. All participants were native speakers of American English, in good health, with normal or corrected-to-normal vision. The two groups were very similar in WAIS Vocabulary, 52.8 (6.5) and 51.1 (7.7) for young and older adults, respectively. As is often found, the two groups differ in WAIS Digit Symbol performance, 74.3 (10.8) and 55.1 (11.9) for young and old. Because of the failed eye tracking for one older adult and mismatching stimulus lists for two young adults and one older adult, we limited our analyses to 15 list-matched speakers in each group. Data from 16 younger adults were reported in Griffin (2001). The matching of stimulus lists with the older adults means that the present results include fewer young participants and that the means for the dependent variables differ slightly.

Apparatus. Eye movements were monitored with a remote video-based pupil/corneal reflection system, an ISCAN ETL-400 with a high-speed upgrade sampling at 120 Hz. A ViewSonic P815 21-inch monitor displayed stimuli. Updating of the displays was synchronised with the vertical retrace of the monitor operating at 60 Hz. One computer processed eye image data, sending uncalibrated data to another computer, which was responsible for timing, presenting stimuli, digitally recording speech, calculating and recording calibrated eye position. Speech was recorded at 12 kHz via a SoundBlaster card, using a LabTec LVA 7330 headset microphone. Participants placed their foreheads against a rest to prevent movements in depth and to keep their eyes about 81 cm from the surface of the monitor. Displayed objects subtended a maximum of 5.5° of visual angle horizontally. Source code for all experiment software is freely available (Mookerjee, Spieler, & Griffin, 2000).

Materials and design. Displays were created with three line drawings of objects (see Figure 1). Each display contained one of 48 objects in position A, one of 48 *critical* objects in position B or C, and one of 3 *repeated* objects in position B or C. Pictured objects were line drawings from Snodgrass and Vanderwart (1980), the Philadelphia naming test (Roach, Schwartz, Martin, Grewal, & Brecher, 1996), and Huitema (1996).

The 48 A objects varied in name frequency. Half of the objects had high frequency names; their noun lemmas had a mean spoken frequency of 70.6 ($SE = 21$) per million words in the Celex database (Baayen, Piepenbrock, & Gulikers, 1995). The other half had low frequency names, with a mean Celex frequency of 5.5 (1.2). All of the A objects were highly codable in spoken picture naming norms (Griffin & Huitema, 1999); mean name agreement from norms with college-aged adults was 87.7% for high frequency names and 86.9% for low. Note that while these values for codability are based on younger adults, there is considerable consistency between younger and older adults based on naming responses given for the present stimuli. We return to this issue in the Results section. Another three objects (ruler, ladder, needle) were repeated across trials to simplify the counterbalancing of picture combinations. These pictures were chosen to be visually similar to one another so that participants would fixate them on every presentation to identify them despite their repetition. The repeated objects were all highly codable, with 86.6% agreement and had low frequency names, 7.7 per million. The 48 critical objects varied in mean codability and frequency, as shown in Table 1. Naming latencies from norming studies (Griffin & Huitema, 1999) or isolated picture naming (Griffin & Bock, 1998, Experiment 1) were available for all but one of the objects. Combining data from these two sources allowed for analysis of

TABLE 1
Means (and standard errors) for properties of critical objects based on norming data from undergraduates

Codability	*Frequency*	*Name agreement*	*Spoken frequency/million*	*Segments*	*Naming latency*	*Size decision latency*
High	High	94 (1.7)	107 (25.2)	4.4 (0.32)	928 (36)	978 (57)
High	Low	90 (3.2)	10 (1.8)	4.8 (0.52)	1111 (50)	876 (44)
Medium	High	42 (6.1)	120 (59.9)	5.3 (0.71)	1197 (62)	935 (51)
Medium	Low	42 (5.8)	5 (1.6)	5.4 (0.60)	1387 (73)	969 (49)

item means for naming latencies. These results showed significant, additive effects of codability, $F(1, 43) = 22.81$, $MSE = 38157$, and frequency, $F(1, 43) = 10.70$, $MSE = 38157$.

Data from another norming study further suggested that differences in naming latencies across conditions were not due to differences in the time needed to identify objects. Twenty Stanford students were asked to judge quickly and accurately whether each picture represented an object that was typically larger than an 8×11 inch sheet of paper. As expected, mean item latencies to make size judgements correlated highly with mean naming latencies, $r(n = 48) = .42$, $p < .005$. However, there were no main effects or interactions of codability and frequency, *minF*'s < 1.0 (Griffin, 2001).

Each participant saw 48 displays containing three objects each. Critical objects formed 12 quadruples containing one object from each cell of the codability by frequency crossing. Eight stimulus lists were created to counterbalance the frequency and identity of the A object for each member in a quadruple across lists, and the position of the critical object (B or C). Across lists, two high and two low frequency A objects rotated in a Latin Square through displays for a quadruple. Members of a quadruple always appeared with the same repeated object (e.g., four quadruples always appeared with the needle). Each list contained an equal number of items in each condition. A list of critical object names appears in the Appendix.

Within-subjects factors were the position, codability, and name frequency of the critical objects and the name frequency of the A objects. The codability and frequency of the critical objects were between-items factors. The primary dependent measures were latency from picture onset to the onset of the A object name, latency from the onset of A to the onset of the critical noun (either the B or C object name), time spent gazing at objects during these time periods, and the fluency of the object names. Gazes extend from the onset of the first fixation in the region of an object to the first saccade leaving the region.

Procedure. An experimenter tested participants one at a time in a large room. They were asked to describe displays with the frame, *The A and the B are above the C*, naming the objects from left to right, top to bottom. They were led through a 9-point calibration and validation routine, 3 practice trials, 1 warm-up trial, and then the 48 experimental trials. Each trial began with a validation point in the top centre of the monitor, about 4° horizontally from positions A and B. As soon as the participants fixated within about 2.3° of this point for 800 ms, the computer would display a stimulus and begin recording data. When a participant completed a sentence, the experimenter ended the trial. There was no explicit time pressure nor was fluency mentioned.

Data treatment and analysis. Each utterance was transcribed by one of two transcribers, one of whom also used sound editing software to measure the onset times for A object names and critical B or C object names.[4] The transcribers were blind to item properties. They agreed on the nouns produced in 144 of 144 cases compared. The following analyses are based on 1420 trials (98.6%). Twenty trials were excluded because speakers made extended comments about the objects or could not identify them at all. Three of the excluded trials came from younger adults.

Analyses include dominant and non-dominant names for objects, uttered fluently and disfluently as in normal speech. Non-dominant responses to medium codable objects have similar or lower name agreement, so the contrast between codability conditions was maintained. For most of the materials, non-dominant responses had similar word frequencies as dominant names, so the frequency contrast was maintained when non-dominant names were said. Noticeable pauses over 200 ms, filled pauses, false starts, corrections, and stressed articles (e.g., ‘‘thee’’ for ‘‘the’’ before a consonant; Fox Tree & Clark, 1997) were considered disfluencies.

Saccades were identified by finding samples where the change in eye position exceeded 60 degrees per second. The onset and offset of the saccades were then identified by a lower criterion of 30 degrees per second. All other samples were assumed to be fixations. Between the end of one saccade and the start of the next saccade, screen x and y coordinates were averaged for the location of that fixation. If fixations fell within a $7.2^\circ \times 7.4^\circ$ rectangle containing an object, they were categorised as on the object. Fixations on an object were collapsed into gazes under the assumption that speakers were still devoting attention to an object while moving their eyes to different parts of the same object.

The pictured objects and their names were originally selected and categorised for codability and name frequency based on the responses of college-aged adults. To minimise the likelihood that category assignments were more representative for one group than another, objects and their names were re-categorised based on the responses for each age group in the present experiment. Note that qualitatively similar results are obtained when the original categorisations were used. Eighty per cent name agreement was used as a cutoff for high versus medium codability. In all, 12 out of 48 items were categorised differently for the age groups. Four

[4] Transcriptions and soundfiles were also run through a forced alignment program, Fasttalk™, to measure noun onsets. Fasttalk yielded noun onsets with confidence values greater than zero for 81% of the first nouns and 89% of critical nouns in analysed trials. The agreement between the forced alignments and hand measurements was generally high with a mean difference of 19 ms for A names and 28 ms for critical nouns. Only the hand measurements are used here.

high codability items for young speakers were medium codable for older speakers (bomb, moon, maze, and plug). Six objects that were medium codable for the young were highly codable for the old (frying pan, palette, sled, stove, vise, and waffle iron). For three items, frequency categories differed across age groups, because *barbell* is less frequent than *weights*, *swimming pool* is less frequent than *pool*, and *frying pan* is less frequent than *pan* (using frequencies for compound words for compounds). Because items were categorised differently across groups, codability and frequency were treated as between- rather than within-items factors in the item analyses. In the subject analyses, Age was the between-subjects factor and A Object Name Frequency, Critical Object Position, Critical Object Codability, and Critical Object Frequency were within-subjects factors. In the item analyses, Critical Object Codability, Critical Object Frequency, and Age were between items and the remaining factors were within items. Two participants were missing means in one cell each and these were replaced with values calculated according to Winer, Brown, and Michels (1991). All proportional means were arc-sine transformed prior to analysis as recommended by Winer et al. All plotted and reported means are calculated by subjects and standard errors (reported in parentheses) are point estimates from subject analyses calculated for the specific condition mean. An alpha value of .05 is used for statistical significance. In the text below, we discuss effects that reached or approached significance in both subject and item analyses, but all effects that reached significance in either analysis are listed in the tables of inferential statistics.

Results and discussion

We first examine the preparation of the A object's name and whether preparation of other object names preceded the onset of A's name as indexed by gazes to the objects. The only property of the A object that was manipulated was its name frequency. Because object A provided the first content word in the utterance, its name frequency should influence its name onset regardless of the scope of preparation for other words. The important issue is whether the onset of A's name is influenced by characteristics of the B or C object. If A's onset is unaffected by the B and C object properties, then any advance processing of these objects does not delay the beginning of the utterance. To address the issue of *any* preparation prior to A's onset, we then examine speakers' gazes on the objects. If speakers only gaze at object B or C after starting to name A, this suggests a highly incremental strategy for word preparation. Such an inference would be strengthened by observing lexical effects (e.g., codability and name frequency) only on gaze durations for these objects after A's name began. Alternatively, if speakers spend time looking at the B or C objects before saying A, this would indicate some advance preparation.

Before ''A'' onset

Time for onset of ''A'' in speech. As sometimes observed in isolated object naming, older speakers took significantly more time than younger adults to begin saying the name of the first object, A, 1459 and 1316 ms respectively (see Figure 2 and Table 2). Both groups showed large effects of the frequency of the A object's name on A's onset time. The effect size for older adults was 120 ms and for younger adults, 159 ms, resulting in a main effect of frequency with no interaction with age. Manipulation of the critical object's position, codability, or frequency had no effect or interactions that approached significance in both analyses of A onsets.

Gaze on A. Analysis of the amount of time speakers spent gazing at A objects before naming them revealed a pattern very similar to name onset times. Older adults spent 872 (22) and young adults 729 (17) ms gazing at A objects, resulting in a significant main effect of age. The frequency effect for older adults was 146 ms and for younger adults, 123 ms, producing a main effect of A object name frequency, but no interaction with age. No other effects approached significance.

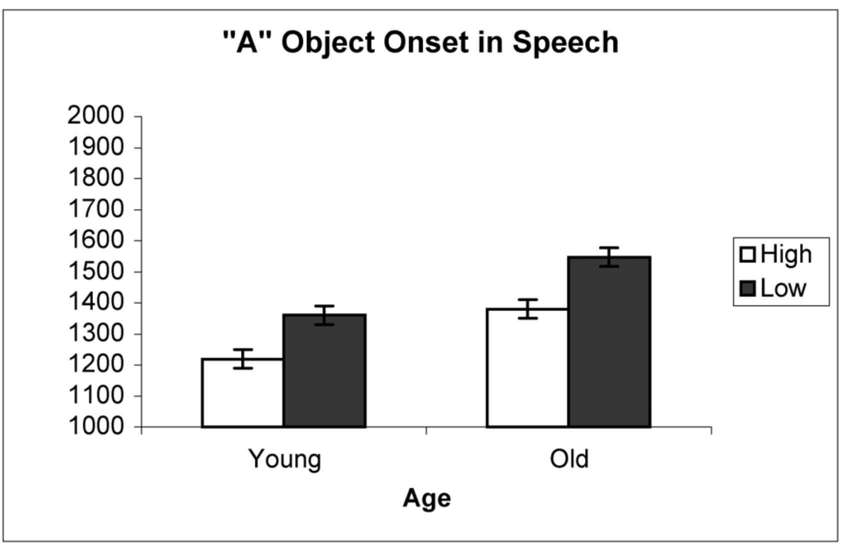

Figure 2. Time until first object name onset as a function of age and first object frequency.

TABLE 2
Inferential statistics before ''A'' onset. Only effects that reach significance by subject or item are included

Factors	F_1	*df*	*MSE*	*p*	F_2	*df*	*MSE*	*p*
''A'' onset								
Age	5.09	1,28	713548	< .04	34	1,88	108670	< .0001
A Freq	25	1,28	112013	< .0001	12.02	1,88	151434	< .009
A Freq × Crit Freq	4.26	1,28	56100	< .05	0.49	1,88	151434	*ns*
Age × A Freq × Crit Freq	6.45	1,28	56100	< .02	1.33	1,88	151434	*ns*
Gaze time on A Object before ''A'' onset								
Age	4.73	1,28	565588	< .04	19.41	1,88	142823	< .0001
A Freq	50	1,28	97179	< .0001	21	1,88	160218	< .0001
Disfluent naming of ''A''								
Age	0.43	1,28	7.24	ns	7.30	1,88	0.48	< .009
A Freq	6.33	1,28	0.69	< .02	1.73	1,88	0.72	ns
Age × A Freq × Crit Freq	6.04	1,28	0.22	< .03	< 1	1,88	0.72	ns
A Freq × Position	4.10	1,28	0.52	< .06	4.06	1,88	0.28	< .05
Age × A Freq × Crit Freq	3.65	1,28	0.52	< .07	10.27	1,88	0.28	< .002
Gaze time on Critical Object B before ''A'' onset								
First Freq	6.94	1,28	16632	< .02	7.81	1,88	12976	< .007
Crit Code	9.82	1,28	8939	< .005	1.48	1,88	14503	*ns*

Note: Position refers to position of the critical object (B or C position). ''Crit'' refers to the critical object in B or C position for which frequency and codability were manipulated. ''A''

Fluency of A. A name was considered disfluent if one of the following occurred before it: a noticeable pause (over 200 ms), a filled pause (''um'' or ''uh''), a false start, a stressed article (''thee'', see Fox Tree & Clark, 1997), or if the noun was corrected (''gira- zebra''). Older adults produced first object names fluently on only 62% (2) of trials compared to the younger adults'' 81% (2), a significant difference. Inferential statistics appear in Table 2. Speakers produced high frequency object A names more fluently than low frequency ones, primarily when the B object was a repeated object rather than a novel critical one. This pattern resulted in a significant interaction between object A name frequency and critical object position. In addition, the main effect of critical object position and the interaction between position and age approached or reached significance in both analyses. Older speakers tended to produce object A names more fluently when the second object was repeated whereas younger adults showed a smaller effect of critical object position. No other effects or interactions approached significance in both analyses.

Thus, these results indicate that older adults are slightly slower than younger adults, the time that speakers spent looking at the objects was closely related to the time to begin saying their names, and this relationship held for both young and older speakers. The results based on the time to name the A object also showed that speech onset did not depend on characteristics of the critical object, regardless of whether that object was in the B or C position. These results suggest that the only age difference prior to speech onset is the extra time needed to retrieve a single noun, the first one produced. We turn next to the eye movement measures for critical objects that allow us to ask whether there is advance name preparation that is simply not sufficient to delay the speakers' onset of speech. In other words, do speakers allocate any attention to either the second or the third object prior to speech? If the answer is no, then speakers attend to the B and C objects only while articulating the first part of the sentence and this in turn suggests that speakers engage in minimal advance preparation of words.

Gazes to critical objects. Shown in Figure 3 is the time that young and older speakers spend gazing at the critical object when in the B position.

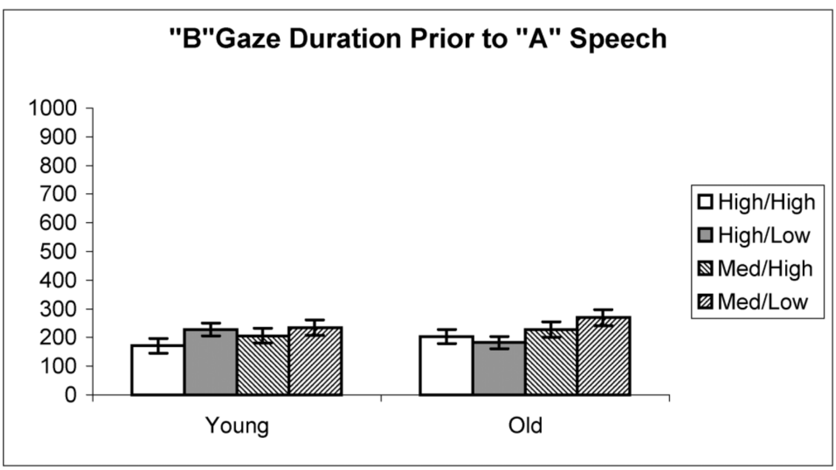

Figure 3. Gaze durations on critical object prior to the production of the first object name in the utterance as a function of age, critical object position, object codability, and object name frequency.

The basic pattern is simple. When the critical object is in the B position, young and old speakers make the temporal equivalent of a single fixation on the B object. Both young and old speakers spend essentially no time on the C object. In the analysis, four trials were dropped because no gaze on these objects was recorded. Older and younger adults showed very similar amounts of time spent gazing at objects in B and C positions prior to the onset of A. Older adults gazed for 138 (13) ms at critical objects in B position and very rarely in the C position, resulting in a minuscule mean of .00001 (.0001) ms. Similarly, younger adults spent 118 (10) ms on critical objects in B position and 7 (4) ms in C. Due to the zero values for objects in C position, ANOVAs were conducted only on gaze time on critical objects in B position prior to ''A''. Speakers spent more time (44 ms) gazing at critical objects when A objects had high rather than low frequency names.[5] Note that this occurs even though there is no effect of the B object characteristics on speech onset. Analyses of gaze durations before speech for repeated objects yielded an identical pattern of results and very similar means.

Summary of processing before ''A''. Taken as a whole, these results suggest minimal advance preparation of object names by young and older adults. For all speakers, speech onset was influenced only by characteristics of the A object. Attention (as reflected in gaze durations) was allocated almost exclusively on object A, with less than 150 ms prior to the onset of its name spent gazing at the B object, and essentially no time on the C object. These results alone allow us to conclude that young and older adults engaged in a last-second approach to preparing words in these utterances. The primary difference being that older adults took slightly more time to retrieve first nouns as reflected in their longer gaze durations for A objects and later A onsets. The consistency across the different measures provides convergent support for this conclusion. Assuming little benefit of extrafoveal preview, the amount of time spent on objects in position B prior to saying A is roughly equivalent to the time needed for object recognition (Potter, 1975). The marginal effects and interactions involving critical object position in the fluency of ''A'' suggests that speakers may have hesitated when they recognised that the B object would be a hard (critical) object to name rather than an easy (repeated) one.

---

[5] Although first object names were matched for length pre-experimentally, speakers tended to replace low frequency names with longer words than high-frequency names, resulting in longer low frequency names. These longer first names allowed more time after the onset of speech for speakers to prepare second nouns (see Griffin, 2003, for a full account of this reversed length effect). The presence of this effect suggests that speakers were making an effort to produce object names fluently and were willing to briefly buffer short first object names.

The remaining analyses focus on the timing during speech and eye positions during speech. These data will help to elucidate how younger and older adults respond to difficulties in word retrieval. This is especially important given that both the young and the older adults are following a fairly risky strategy of retrieving words just moments prior to their articulation in the utterance.

During speech

Lag to critical name onset. For each trial, we calculated the time from the onset of the object A's name to the onset of the critical object's name (see Figure 4). This latency measure includes time spent articulating the first noun[6] and other intervening words and disfluencies. However, the measure excludes time spent in disfluencies that occurred immediately before the first noun, which are more likely to be associated with the first noun than later ones. Of course, speakers took more time to begin to utter the critical noun when it was in position C rather than B, resulting in a significant main effect of critical object position (see Table 3 for inferential statistics). Older speakers tended to take more time from the beginning of A to the beginning of the critical noun. Interestingly, the age difference was about 200 ms regardless of whether the critical object was in position B or C. Specifically, in the B position it took older adults 1376 (59) ms to begin uttering the critical object name, and when in C position, 2315 (69) ms. Younger adults took 1186 ms (56) before B nouns and 2116 (67) ms before C nouns. This 200 ms age difference was significant by items but not subjects. This is a fairly small age difference, under 10% for the time between the A noun and C noun. The fact that this effect did not increase from the B to the C noun suggests that there was not an overall age-related decrease in speech *rate.*

Younger and older speakers were similarly influenced by codability. Both groups took more time to begin uttering medium rather than highly codable object names, a 502 ms difference for older adults and 406 ms for younger. As expected, speakers took more time to begin saying the critical object's name when it was lower in frequency, and this slowing was similar for both older and younger adults (354 ms and 319 ms respectively). The interaction between codability and critical name frequency was significant, due to an overadditive effect of frequency for medium codable objects. That is, speakers were particularly slow in retrieving low frequency names for medium codable objects. Having already produced the first nouns, the

[6] Analyses were also performed excluding trials with non-targets produced for the first object name to ensure the results were not due to differences in the words used and their lengths. The results were the same as those including those trials.

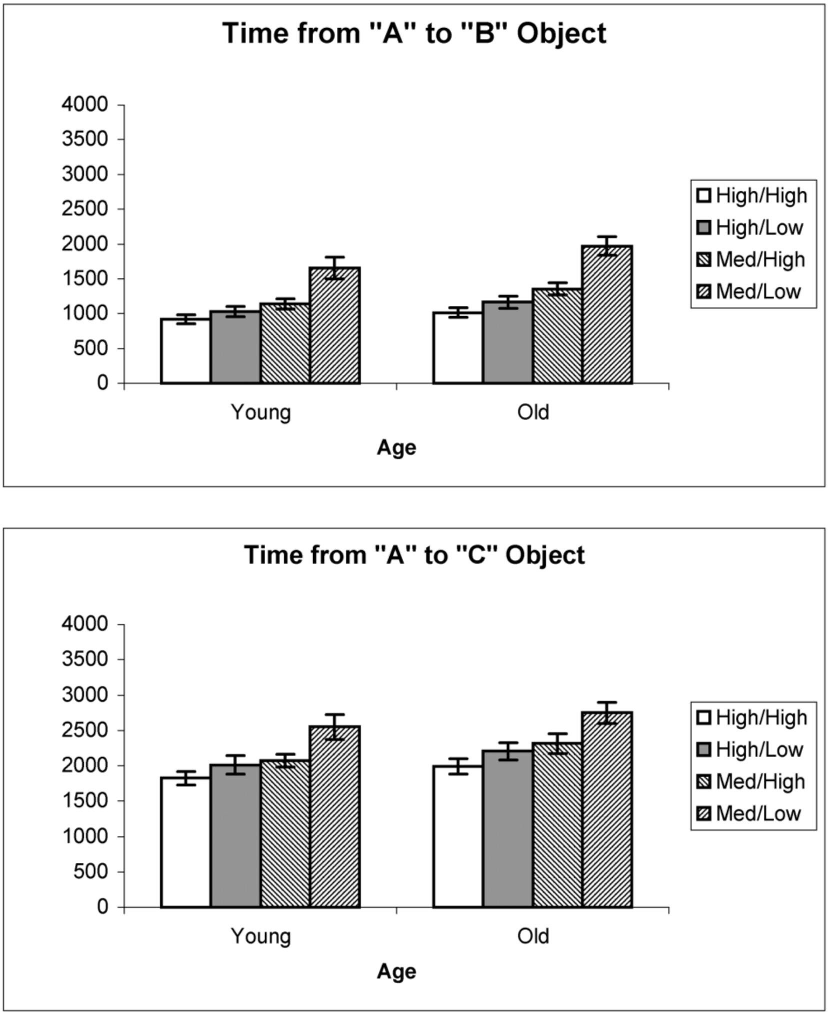

Figure 4. Time from the onset of articulation of the first object name until the onset of articulation of the critical object name as a function of age, critical object position, object codability, and object name frequency.

groups were minimally affected by A frequency. However, the frequency of A entered into a three-way interaction with Position and Critical object codability in which the codability effect for critical objects in position B after high frequency names for A was greater than for the other three cells. These results show that the timing of nouns within speech in both young and older speakers reflected the difficulty of word retrieval.

TABLE 3
Inferential statistics for measures of processing during speech. Only effects that reach significance by subject or item are included

Factors	F_1	*df*	*MSE*	*p*	F_2	*df*	*MSE*	*p*
Time between "A" and Critical Name Onsets								
Age	1.95	1,28	2316769	*ns*	6.66	1,88	615056	< .02
Position	212	1,28	493163	< .0001	678	1,88	125250	< .0001
Crit Codability	118	1,28	208363	< .0001	34.88	1,88	615056	< .0001
Crit Freq	46	1,28	290322	< .0001	15.94	1,88	615056	< .0002
Crit Code × Crit Freq	18.40	1,28	193986	< .0003	4.32	1,88	615056	< .05
Position × First Freq × Crit Code	4.76	1,28	195101	< .04	4.14	1,88	160626	< .05
Fluency of Critical Name								
Age	4.64	1,28	3.32	< .05	11.39	1,88	1.01	< .002
Crit Codability	56	1,28	0.58	< .0001	27	1,88	1.01	< .0001
Crit Freq	39	1,28	0.60	< .0001	16.33	1,88	1.01	< .0002
Crit Code × Crit Freq	41	1,28	0.26	< .0001	5.00	1,88	1.01	< .03
Age × Crit Code × Crit Freq	11.53	1,28	0.26	< .003	< 1	1,88	1.01	*ns*
Age × Position × Crit Freq × A Freq	2.66	1,28	0.59	*ns*	5.47	1,88	0.41	< .03
Gaze time on Critical Object between "A" and "B"								
Age	4.89	1,28	2682416	< .04	6.74	1,88	1350784	< .02
Position	45	1,28	1156355	< .0001	176	1,88	264243	< .0001
Crit Codability	168	1,28	504963	< .0001	49	1,88	1350784	< .0001
Critical Freq	94	1,28	403785	< .0001	20	1,88	1350784	< .0001
Crit Code × Crit Freq	42	1,28	203583	< .0001	3.97	1,88	1350784	< .05
Age × A Freq	5.39	1,28	262974	< .03	3.86	1,88	323940	< .06

Note: Position refers to position of the critical object (B or C position). "Crit" refers to the object in B or C position for which frequency (Freq) and codability (Code) were manipulated. "A" refers to object in A position. Proportions were arc-sine transformed for these analyses (Winer et al., 1990).

Fluency of critical name. The next set of analyses deal with the fluency of the critical object names. It is important to bear in mind that the fluency of object names is also closely related to the timing of the utterance. That is, names that follow disfluencies are likely to take more time to produce than more fluent names. Both overall timing and fluency of the names reflect the ease and success of the production process. Thus, it is more accurate to say that both are reflections of the production process rather than to say either that speakers are slower because they are more disfluent or speakers are more disfluent because they are slower.

A name was considered disfluent if one of the following occurred before it: a noticeable pause (over 200 ms), a filled pause (''um'' or ''uh''), a false start (''the gira- zebra''), a stressed article (''thee''; see Fox Tree & Clark, 1997), or if the noun was corrected (''the giraffe, zebra''). The two 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 65% fluently. Unsurprisingly, medium codable objects elicited disfluencies more often than highly codable ones did, with a difference of 23% for older speakers and 18% for younger. Among highly codable object names, both groups showed small frequency effects in the proportion of fluent trials, in the order of 5% for young and 1% for old. A striking age difference in fluency occurred in retrieving low frequency names for medium codable objects. There was a 14% frequency effect for young adults and 32% for older adults. Thus, in addition to significant main effect of frequency, ANOVAs revealed a significant two-way interaction of codability and frequency. The three way interaction of codability, frequency, and age was significant by subjects but not by items. So, as competition in name selection increased and word frequency decreased, older speakers tended to be less fluent than young adults.

Gazes to critical objects. The time spent gazing at critical objects between the onset of A and the onset of the critical object's name mirrored effects in latencies between names in most respects (see Figure 5). Older adults spent 318 ms longer than younger ones gazing at critical objects, resulting in a significant main effect of age. Speakers of all ages spent about 400 ms longer gazing at critical objects in C position rather than B position. This very large effect is most likely due to the absence of any additional objects in the display to retrieve names for after the C object (see Van der Meulen, 2001, for a similar observation). This effect of position did not interact with other factors. Reinforcing the notion that critical word retrieval occurred during speech, gaze times showed large and robust effects of critical object codability, frequency, and an interaction between them. In other words, the same object properties that influenced the timing and fluency of critical nouns also had robust effects on the timing of gazes on these objects. In addition, there was a significant interaction between age and frequency of ''A''. Older adults gazed at critical objects for 106 ms more between names when they followed A objects with low rather than high frequency names, whereas younger adults showed the opposite pattern. No other effects in gaze times interacted with age group.

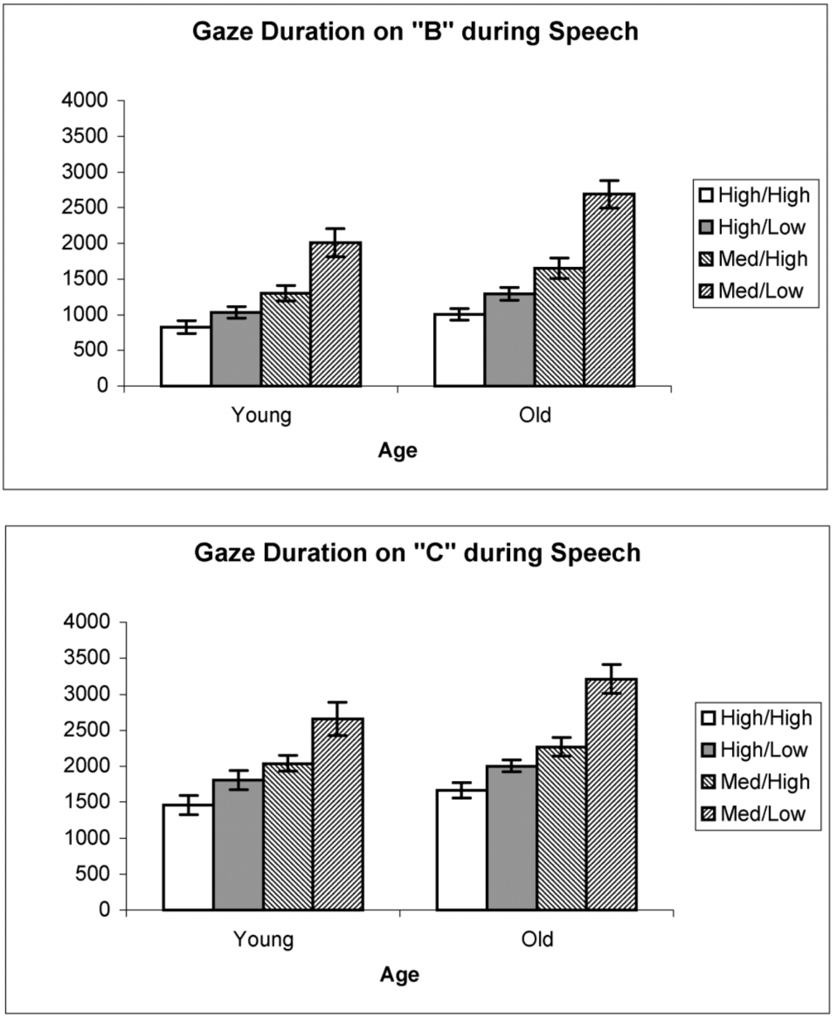

Figure 5. Gaze durations on critical objects as a function of age, critical object position, object codability, and object name frequency.

Summary. The primary finding is the very strong consistency in the timing of vocal responses and eye movements for the two groups. To further reinforce the consistency in the two groups, we have plotted the proportion of trials when speakers were gazing at object B relative to the onset of the object B name in the speakers' utterances (Figure 6). There appears to be no difference between young and old in the time that speakers begin gazing at the object relative to the onset of naming that object in the utterance. With the exception of a 200–300 ms additive

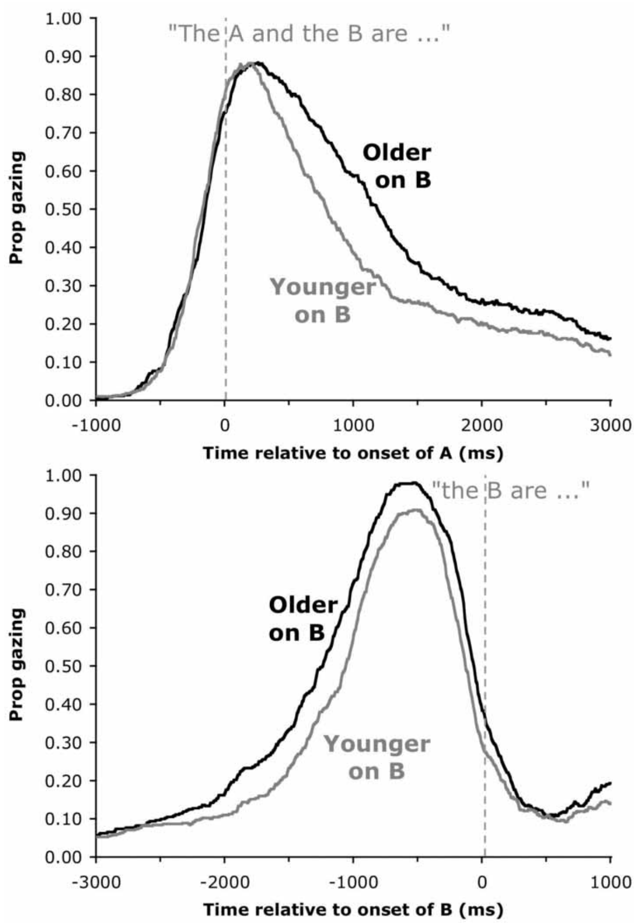

Figure 6. Proportion of speakers gazing at object B relative to the onset of object A's name and object B's name in utterance.

difference between younger and older speakers, the two groups were identical in the time course of their utterances. The two groups spoke at very similar rates, had similar gaze patterns, and experienced similar delays based on the difficulty of retrieving words. The strong consistency across groups in timing suggests that both groups retrieve their words a moment before uttering them.

GENERAL DISCUSSION

We can summarise the main point from all of these analyses quite simply: all of these results demonstrate that both young and older speakers followed an incremental word retrieval strategy in which they uttered object names shortly after retrieving them. Indeed, analyses suggested that the time course of word retrieval for the two groups was very similar. That is, neither group tended to buffer object names for long while retrieving words that came later in the utterance. The lack of an increased scope of preparation for the older adults had dramatic consequences for the fluency of their speech. Specifically, when the upcoming words required selecting between multiple names (medium codable) and phonological encoding was slower (low frequency names), older adults were particularly disfluent (62% disfluent trials vs. 37%). Converging evidence is provided by the analyses of eye movements over the pictures prior to and after the onset of speech. Properties of the B and C objects did not affect the onset of object A's name, and eye movement measures paralleled latency effects, suggesting that prior to the onset of speech, attention was directed primarily to the first object to be mentioned in the sentence.

The scope of word preparation in language production has important consequences for how quickly and fluently utterances may be articulated. Generally, studies suggest that greater preparation prior to the onset of speech results in a faster and more fluent delivery (e.g., Griffin & Bock, 2000). We suggested in the introduction that an increase in preparation might be one method that older adults might use in order to alleviate the influence of any age-related increases in production difficulties (Griffin & Spieler, 2000). However, the present results suggest that older adults appear to make no such adjustments. Thus, the slowing and increased disfluencies of the older speakers are a relatively straightforward reflection of the same production difficulties seen in more isolated word production studies. This age-related increase in disfluency was especially apparent when there was competition at the lexical level. Words that are medium codable have multiple acceptable names and older adults encounter problems that result in a failure to have the word selected and phonologically encoded by the time the object must be produced.

Obviously, there are several important differences between production in the present experiment and natural settings for language production. However, not all of these differences put the older adults at a disadvantage in the present experiment with respect to the fluency of their utterances. First, neither young nor older speakers were informed of any need to begin or end their sentences within a minimum time. One might expect that speakers would be most likely to utter their words immediately after retrieval when they are placed under tight temporal constraints (as in Ferreira & Swets, 2002; Griffin & Spieler, 2000). While it is difficult to say with certainty, the lack of any overt timing demands makes it less likely that the older adults selected a mode of production that placed greater emphasis on time than is the case in their normal speech.

Preparing one's words far in advance requires prepared speech to be buffered. This added memory component would seem to place some restrictions on the viability of increased preparation as a general strategy for maximising fluency. In particular, this may be a strategy that is only available for relatively high-functioning older adults who can accommodate the increased working memory demands (see Martin & Freedman, 2001) associated with increased preparation. However, an inspection of the WAIS performance measures of our participants reveals that the older adults were very high functioning, scoring in the 82nd percentile of the population for their age group based both on vocabulary and digit-symbol performance. This high level of functioning would seem, if anything, to underestimate the size of the age-related increase in disfluencies observed in the present experiment. Of course, the high degree of consistency in timing between young and older speakers may also be partly a function of the high level of functioning for all of these speakers. It is an interesting possibility that older speakers functioning at a lower level may exhibit still greater age differences when they prepare their utterances. However, it is extremely unlikely that such older speakers would differ in a way that would make them *more* fluent than the current group of older speakers. Instead, speakers in more open-ended or natural speaking situations may alter the content or syntactic structure of their utterances to cope with variations in the difficulty of word retrieval (Ferreira & Dell, 2000; Ferreira & Firato, 2002).

Caveats and concerns

Age differences in the present experiment were primarily main effects in which the older speakers appeared to spend more time retrieving words than young adults did. This may have been due to age-related decreases in the speed of word retrieval, but there are also a few alternative hypotheses. For instance, speakers direct their attention to objects prior to moving their eyes to the objects' locations. The eyes probably lag behind attention by about 200 ms. During this time, speakers may use extrafoveal visual information about the object to which they are moving their eyes to begin identifying the object. Extrafoveal sensitivity decreases with age, making it likely that younger adults could preprocess objects more than older adults could (see Irwin, 2004). As a result, the age-related differences in the timing of speech and gaze durations may be attributable to differences in eyesight.

Another concern is the difference in name preferences in older and younger adults. Often, departures from the names provided by normative samples are considered naming errors. However, even on occasions when an older adult provides the same name for an object as a younger adult, the name may be more difficult to retrieve because it has fiercer competitors. For critical objects, we attempted to equate naming difficulty across conditions for younger and older speakers by categorising items according to how each group named them. This appeared to work. Name agreement for critical objects did not differ significantly for the two age groups, 76.7% (1.9) and 72.2% (2.3) for younger and older speakers respectively. The age main effect remained significant at least in item analyses for latencies and gaze times for critical objects during speech, even when name agreement for items for each group were partialled out. Both groups showed very high name agreement for A objects, 88.3% (1.2) for old and 94.6% (0.9) for young. Nonetheless, older adults agreed on dominant names for first objects significantly less often than younger adults did, $F_1(1, 28) = 12.19$, $MSE = 0.62$, $p < .002$; $F_2(1, 88) = 13.82$, $MSE = 0.43$, $p < .0001$. The higher level of agreement for the younger speakers is likely to translate into slightly faster word selection than would be the case if the young matched the agreement level of the older speakers. When A's name agreement within each group was partialled out, the differences in latency for A shrank to a non-significant 84 ms and the difference in gaze time on object A was a very small 42 ms. Group differences in the fluency of A disappeared as well, although the frequency effects were still reliable in all analyses.

We call attention to these concerns and additional analyses because they are not unique to our experiment. They suggest that age effects in object naming may frequently be overestimated due to differences in the number of competing responses for older and younger speakers. That said, a small age-related decrease in the speed of word retrieval appears to remain for referring to objects within multiword utterances.

Relation to more general theories in cognitive ageing

Theories of age-related changes in cognitive processing are often based on very simple experimental paradigms. The result of this is that these theories can be difficult to apply to situations such as language production where behaviour unfolds over time. For example, generalised slowing accounts suggest that with age, the rate of information processing decreases and that this in turn results in the wide range of age differences reported in the literature. There has been so much discussion on this issue that we will not attempt any further review (see, e.g., Cerella, 1985; Fisher & Glaser, 1996; Myerson, Hale, Wagstaff, Poon, & Smith, 1990; Spieler, 2001). The relevant issue with respect to generalised slowing accounts is that these theories often do not account for simultaneous changes in both the what and when of behaviour. In the present context, older adults are somewhat slower to initiate speech, speech rate is somewhat slower, and the older adults are considerably less fluent. Generalised slowing might simply predict that older adults would speak more slowly, and they do. To understand why older adults' speech is both slowed and less fluent, particularly in situations when there is lexical competition, one would have to draw on theories specifically concerned with the preparation and production of language, probably along with considerably more empirical evidence than we currently have available (for further discussion, see Griffin & Spieler, in press). There are some theoretical frameworks that might be more specifically useful.

One theory that is conceptually closer to the production theories that we have drawn on and has been applied to age differences in language processing is the Node Structure Theory (NST) of MacKay and colleagues (MacKay, 1982; MacKay & Burke, 1990). The model consists of multiple layers of processing nodes that code for semantic, phonological, and articulatory information leading to production. The model also incorporates timing and sequencing nodes that allow one to control the ordering and rate of activation of individual nodes. This theory has generally been applied to single word retrieval based on word definitions (Burke et al., 1991; James & Burke, 2000) and picture naming (e.g., Taylor & Burke, 2002). For example, the age difference in the probability of tip of the tongue (TOT) states is accounted for by the fact that the mapping of semantics to lexical nodes is a mapping of many semantic nodes to a small number of lexical nodes. In contrast, the mapping of lexical nodes to phonological nodes is one to many. Generally speaking, a many-to-few mapping is more fault tolerant than a one-to-many mapping because any degradation in activation along one or a few connections can be overcome by the many other connections. Age differences in TOT probabilities are explained by a general degradation in the communication between nodes in the network that, because of the comparative robustness of the semantic to lexical mapping, has a larger effect on the process of phonological retrieval. Note also that this model might predict older adults should be at a disadvantage when there are multiple competing lexical names for an object because the semantics would then map to multiple lexical entries. In the present study, older adults were particularly disfluent when the object was medium codable and the object name was low in frequency, consistent with NST.

The present results, in conjunction with other results (Griffin & Spieler, 2000) suggests a correlation between age differences in planning and age differences in the fluency of speech. That is, in situations where older adults appear to prepare words further in advance than younger adults (e.g., Griffin & Spieler, 2000), age differences in fluency are attenuated, whereas when both young and older adults are highly incremental in preparing upcoming words, older adults are, as seen here, less fluent in their resultant speech. Although finding that both young and older adults elect to do very little advance word preparation and that older adults are more disfluent than younger adults are both rather prosaic results, the methods used in the present study have the potential to help understand the relation between ageing, planning, and fluency in language production.

REFERENCES

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* (Version (Release 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].

Banseel, P., Griffin, Z. M., & Spieler, D. H. (2001). *Matlab routine for measuring speech onset and offset*. Retrieved from http://oak.psych.gatech.edu/~spieler/software.html.

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

Bock, K. (1995). Sentence production: From mind to mouth. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition. Vol 11: Speech, language, and communication* (pp. 181–216). Orlando, FL: Academic Press.

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.

Bowles, N. L. (1993). Semantic processes that serve picture naming. In J. Cerella, J. Rybash, W. Hoyer, & M. L. Commons (Eds.), *Adult information processing: Limits on loss* (pp. 303–323). San Diego, CA: Academic Press.

Bowles, N., Obler, L. K., & Albert, M. L. (1987). Naming errors in healthy aging and dementia of the Alzheimer's type. *Cortex*, *23*, 519–524.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Review, 109*, 204–223.

Brown, R., & McNeill, D. (1966). The ''tip of the tongue'' phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325–337.

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 and Language, 30*, 542–579.

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.

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.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*, 177–208.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. C. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning Memory and Cognition*, *27*, 1430–1450.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, *98*, 67–83.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, *5*, 313–349.

Dell, G. S. (1995). Speaking and misspeaking. In L. R. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science. Vol. 1: Language* (2nd ed., pp. 183–208). Cambridge, MA: MIT Press.

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, *23*, 517–542.

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

Ferreira, V. S., & 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.

Feyereisen, P. (1997). A meta-analytic procedure shows age-related decline in picture naming: Comments on Goulet, Ska, & Kahn (1994). *Journal of Speech and Hearing Research*, *40*, 1328–1333.

Fisher, D. L., & Glaser, R. A. (1996). Molar and latent models of cognitive slowing: Implications for aging, dementia, depression, development, and intelligence. *Psychonomic Bulletin and Review*, *3*, 458–480.

Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing ''the'' as ''thee'' to signal problems in speaking. *Cognition*, *62*, 151–167

Fromkin, V. (1968). Speculations on performance models. *Journal of Linguistics*, *4*, 1–152.

Goodglass, H., Theurkauf, J. C., & Wingfield, A. (1984). Naming latencies as evidence for two modes of lexical retrieval. *Applied Psycholinguistics*, *5*, 293–303.

Goulet, P., Ska, B., & Kahn, H. J. (1994). Is there a decline in picture naming with advancing age? *Journal of Speech and Hearing Research*, *37*, 629–644.

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., & Bock, K. (1998). Constraint, word frequency, and levels of processing in spoken word production. *Journal of Memory and Language*, *38*, 313–338.

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

Griffin, Z. M., & Huitema, J. S. (1999). *Beckman Spoken Picture Naming Norms*. [On line]. Retrieved from http://www.langprod.cogsci.uiuc.edu/~norms.

Griffin, Z. M., & Spieler, D. H. (2000, November). *Speech planning in younger and older adults.* Poster presented at the annual meeting of the Psychonomic Society.

Griffin, Z. M., & Spieler, D. H. (in press). Observing the what and when of language production for different age groups by monitoring speakers' eye movements. *Brain and Language*.

Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighborhood effects in speech production. *British Journal of Psychology*, *89*, 151–174.

Horton, W. S., Metzing, C., & Gerrig, R. J. (2002). *Tracking speakers' use of internal and external information during the construction of referring expressions.* [Unpublished Manuscript]

Huitema, J. S. (1996). [The Huitema picture collection]. Unpublished.

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology 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, and Cognition*, *26*, 1378–1391.

Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa, et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 432–438.

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.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.

Kelly, M. H. (1986). *On the selection of linguistic options.* Unpublished doctoral dissertation, Cornell University, Ithaca, NY, USA.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, *14*, 185–209.

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology, 99*, 199–208.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*, 223–232.

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.

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 ultilization* (pp. 213–263). Amsterdam: 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.

Meyer, A. S., & Van der Meulen, F. F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin and Review*, *7*, 314–319.

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

Mookerjee, A., Spieler, D. H., & Griffin, Z. M. (2000). *Eye2: DOS software for speech and eye movement recording*. [On-line] Retrieved from http://psychology.gatech.edu/spieler/software.html

Myerson, J., Hale, S., Wagstaff, D., Poon, L. W., & Smith, G. A. (1990). The information loss model: A mathematical theory of age-related cognitive slowing. *Psychological Review*, *97*, 475–487.

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.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*, 273–281.

Paivio, A., Clark, J. M., Digdon, N., & Bons, T. (1989). Referential processing: Reciprocity and correlates of naming and imaging. *Memory and Cognition*, *17*, 163–174.

Potter, M. C. (1975). Meaning in visual search. *Science*, *187*, 965–966.

Ramsay, C. B., Nicholas, M., Au, R., Obler, L. K., & Albert, M. L. (1999). Verb naming in normal aging. *Applied Neuropsychology*, *6*, 57–67.

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 and Language*, *35*, 585–605.

Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S. & Brecher, A. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, *24*, 121–133.

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, *73*, 205–246.

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.

Spieler, D. H. (2001). Modelling age-related changes in information processing. *European Journal of Cognitive Psychology*, *13*, 217–234.

Taylor, J. K., & Burke, D. M. (2002). Asymmetric aging effects on semantic and phonological processes: Naming in the picture-word interference task. *Psychology and Aging*, *17*, 662–676.

Van der Meulen, F. (2001). *Moving eyes and naming objects.* Unpublished Dissertation, Katholieke Universiteit Nijmegen.

Van Gorp, W. G., Satz, P., & Kiersch, M. E. (1986). Normative data on the Boston Naming Test for a group of normal older adults. *Journal of Clinical and Experimental Neuropsychology*, *8*, 702–705.

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, *37*, 356–381.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd edn). New York: McGraw-Hill.

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.

Appendix

Critical objects of high or medium codability and names of high or low frequency based on younger adults. Alternative names appear after a slash and optional elements in parentheses

Quadruple	*High-High*	*High-Low*	*Medium-High*	*Medium-Low*
1	apple	axe	hat/cap	oven/stove
2	baby	button	board/wood	(rasp)berries
3	cake	cane	coat/jacket	crane/bulldozer
4	moon	maze	matches/matchbook	mop/broom
5	star	screw	stairs/staircase	strainer/sieve
6	window	wreath	(wine)glass	waffle iron/maker
7	table	plug	(frying)pan	palette/paints
8	bowl	butterfly	tray/platter	donkey/mule
9	bomb	well	weights/barbells	(killer)whale
10	bottle	vase	chest/trunk	vise/clamp
11	shoe	slide	(swimming) pool	sled/toboggan
12	tooth	doorknob	TV/television	limousine/limo