Lipreading, Processing Speed, and Working Memory in Younger and Older Adults

Julia E. Feld
Mitchell S. Sommers
Washington University in
St. Louis, Missouri

Purpose: To examine several cognitive and perceptual abilities—including working memory (WM), information processing speed (PS), perceptual closure, and perceptual disembedding skill—as factors contributing to individual differences in lipreading performance and to examine how patterns in predictor variables change across age groups.
Method: Forty-three younger adults (mean age = 20.8 years, *SD* = 2.4) and 38 older adults (mean age = 76.8 years, *SD* = 5.6) completed tasks measuring lipreading ability, verbal WM, spatial WM (SWM), PS, and perceptual abilities.
Results: Younger adults demonstrated superior lipreading ability and perceptual skills compared with older adults. In addition, younger participants exhibited longer WM spans and faster PS than did the older participants. SWM and PS accounted for a significant proportion of the variance in lipreading ability in both younger and older adults, and the pattern of predictor variables remained consistent over age groups.
Conclusions: These findings suggest that the large individual variability in lipreading ability can be explained, in part, by individual differences in SWM and PS. Furthermore, as both of these abilities are known to decline with age, the findings suggest that age-related impairments in either or both of these abilities may account for the poorer lipreading ability of older compared with younger adults.

KEY WORDS: lipreading, aging, cognition

Lipreading[1]—the perception of speech by interpreting visually available movements of the face, mouth, and tongue—is useful to both hearing-impaired and normal-hearing people. When acoustic information about speech is degraded, either by noise or by hearing impairment, being able to see a speaker's articulatory movements as well as hear them significantly increases speech intelligibility compared with looking or listening alone (Erber, 1969; Grant, Walden, & Seitz, 1998; Sumby & Pollack, 1954). This enhancement in performance is partially due to the fact that visual information about speech can complement auditory information, especially in noisy or reverberant settings (MacLeod & Summerfield, 1987). For instance, acoustic cues to place of articulation for consonants are often difficult to perceive in noisy environments, but the shape of the mouth and articulatory movements that correspond to those consonants are often visually clear. Thus, the availability of visual speech cues provides individuals with an opportunity to compensate

---

[1] The term *lipreading* is used to refer to tasks in which no auditory information is available. The term *speechreading* is used to refer to processing visual speech information in the presence of congruent auditory signals.

for the reduction in available auditory information by providing an alternative modality to obtain phonetic information, resulting in auditory–visual (AV) scores that are often higher than would be expected simply by adding the auditory-only (A) and visual-only (V) performances (Grant et al., 1998; also see Zekveld, Kramer, Vlaming, & Houtgast, 2008, for evidence for a similar superadditive pattern in auditory speech perception augmented by visually presented masked text). This benefit of AV compared with A presentations can be substantial. For example, Middelweerd and Plomp (1987) demonstrated that adding visual speech information was, on average, equivalent to a 4.3-dB improvement in signal-to-noise (S/N) ratio.

The addition of visual input does not benefit all participants uniformly. MacLeod and Summerfield (1990) found improvements in S/N ratios that ranged from 2.7 to 9.5 dB when participants could both see and hear the talker, compared with listening alone. To put this in perspective, a 3-dB increase in S/N ratio translates into approximately a 22% improvement in speech perception, whereas a 10-dB increase in S/N produces about a 74% improvement, on the basis of a 7.4%-per-dB increase (MacLeod & Summerfield, 1990). Of particular importance, the amount of benefit obtained from the addition of visual speech information as a supplement to auditory information (henceforth referred to as *visual enhancement*) correlated strongly with silent lipreading ability ($r = .89$, $n = 20$, $p < .01$). That is, those individuals who were better lipreaders also exhibited the greatest visual enhancement. Recently, Sommers, Tye-Murray, and Spehar (2005) demonstrated the importance of age-related changes in lipreading by comparing visual enhancement in younger and older adults. Sommers et al. reported that older adults obtained less visual enhancement than younger adults before controlling for lipreading performance but had nonsignificant differences in visual enhancement after controlling for V performance. Taken together, these findings suggest that lipreading ability is one of the principal factors that determine how much benefit individuals get from the addition of visual speech information and that this benefit varies dramatically even within relatively homogenous populations.

The importance of lipreading for AV speech perception has led to considerable research on how this ability varies across individuals and populations. Although differences in methodology make direct comparisons across studies difficult, there is overwhelming evidence for high levels of variability in lipreading performance. For example, accuracy has been shown to range from 0% to 94% correct for hearing-impaired children (Lyxell & Holmberg, 2000), from 0% to 41% in normal-hearing children (Lyxell & Holmberg, 2000), from 0% to 65% correct in normal-hearing adults (Auer & Bernstein, 2007), and from 0% to 85% correct in adults with early-onset hearing impairment (Auer & Bernstein, 2007).

One consistent finding with respect to differences in lipreading across populations is that older adults generally exhibit reduced lipreading abilities compared with younger adults (Cienkowski & Carney, 2002; Dancer, Krain, Thompson, Davis, & Glen, 1994; Honnell, Dancer, & Gentry, 1991; Shoop & Binnie, 1979; Sommers et al., 2005; Spehar, Tye-Murray, & Sommers, 2004). Note that this age-related loss in lipreading abilities is somewhat unexpected for at least two reasons. First, older adults may have an increased need to rely on the visual perception of speech because of age-related hearing loss. Second, because age-related hearing loss develops gradually, older adults have an opportunity to learn to encode visual speech information over the course of several years. Despite these considerations, however, numerous studies have found that for postlingual hearing loss (for evidence of improved lipreading in congenitally deafened individuals, see Bernstein, Demorest, & Tucker, 2000), there is little relationship between hearing status and lipreading (Clouser, 1977; Farrimond, 1959; Lyxell & Rönnberg, 1989, 1991; Owens & Blazek, 1985; Rönnberg, 1990; but see Auer & Bernstein, 2007; Bernstein, Auer, & Tucker, 2001); furthermore, as noted, younger adults consistently outperform older adults on lipreading tasks.

The absence of correlations between hearing loss and lipreading has led researchers to investigate a range of perceptual and cognitive abilities as factors that might contribute to individual differences in V speech perception. In general, these studies have failed to find consistent correlates of lipreading either within or across different populations. For example, several studies have reported that overall intelligence is a relatively poor predictor of lipreading ability (Elphick, 1996; for a review, see Jeffers & Barley, 1971), as are verbal reasoning abilities (Jeffers & Barley, 1971; Summerfield, 1991), vocabulary (Lyxell & Rönnberg, 1992; Simmons, 1959), and education level (Dancer et al., 1994). In some studies (Dancer et al., 1994; Johnson, Hicks, Goldberg, & Myslobodsky, 1988), female participants have been found to outperform male participants on measures of lipreading, but these effects are usually small and often fail to reach significance (Aloufy, Lapidot, & Myslobodsky, 1996; Irwin, Whalen, & Fowler, 2006; Tye-Murray, Sommers, & Spehar, 2007).

Another set of abilities that has been examined as possible predictors of lipreading is complex perceptual tasks in which participants are asked either to extract meaningful information from a complex array or to make inferences about missing information. For instance, Sharp (1972) found that good lipreaders were significantly better than poor lipreaders at some tasks of *visual closure*—the ability to integrate dissociated parts to form a unified whole and to locate specific figures within larger, more complex figures. Lyxell and Rönnberg (1989) found significant correlations between lipreading performance and verbal inference-making ability—the ability to fill in missing letters to form a word and to fill in words to form sentences. Similarly, Sanders and Coscarelli (1970) found significant differences between good and poor lipreaders using measures of word, sentence, and picture completion, all of which assessed visual synthesis—the ability to generate complete representations from partial information.

Despite some progress in establishing predictors of lipreading ability, several difficulties with previous studies in the area make it difficult to draw any firm conclusions about predictors of V speech perception. First, the perceptual constructs that have been investigated tend to be poorly defined and operationalized, often rendering contradictory results. For instance, Sharp (1972) used five perceptual tasks thought to measure the same underlying construct. Three of these measures showed significant differences between good and poor lipreaders, and two did not. Similarly, Simmons (1959) found significant correlations between lipreading and one measure of synthetic ability but nonsignificant correlations between two other measures of the same construct. One of the significant relationships with lipreading ability found by Simmons was the ability to decipher fragmented sentences, an ability that, in another case, has also failed to significantly predict lipreading ability (Watson, Qui, Chamberlain, & Li, 1996). The situation is further complicated because there has been little standardization of tasks used to predict V speech perception. Thus, even when significant correlations are found, it is unclear what trait has been isolated and how it relates to other cognitive constructs.

The failure of past studies to establish reliable predictors of lipreading ability has led more recent investigations to examine the relationship between V speech perception and cognitive abilities. Unfortunately, many of these studies (often from the same laboratories) have produced somewhat inconsistent results. For example, measures of working memory (WM) have been shown to correlate significantly with lipreading in some studies (Lidestam, Lyxell, & Andersson 1999; Lyxell & Holmberg, 2000) but not in others (Lyxell & Rönnberg, 1989). Lyxell and Rönnberg (1993) found correlations between verbal WM (VWM) and lipreading in a condition with meaningful background noise but not between VWM and silent lipreading. Similar conflicting results have been found between studies examining processing speed (PS) and lipreading ability, with some studies finding significant relationships (Lyxell & Holmberg, 2000) and others failing to find systematic relationships (Lidestam et al., 1999).

The confusion in the literature about what factors predict lipreading may be due in part to the inclusion of a limited number of predictor variables in each study. For example, to our knowledge, there are no studies of factors predicting lipreading that have obtained measures of perceptual abilities, WM, and PS from the same participants. Of particular relevance to the current experiment is that unless a range of cognitive and perceptual abilities are measured within the same individuals, shared variance among the different predictor variables of lipreading may lead to dramatically different conclusions depending on the set of predictor variables included in a given study. For example, PS and WM are highly correlated (Fry & Hale, 2000), making it essential to include both measures as a means of identifying their independent contributions to lipreading ability. In addition, many studies have used limited measures of a given construct or have not distinguished between different components of the constructs. For example, despite considerable evidence that age-related declines in WM capacity are greater for spatial than for verbal tasks (Jenkins, Myerson, Joerding, & Hale, 2000; Myerson, Hale, Rhee, & Jenkins, 1999), to our knowledge, studies examining spatial WM (SWM) as a potential predictor of lipreading ability are absent from the literature.

The current study had two main goals. First, it was designed to overcome some of the limitations of previous studies examining predictors of lipreading ability by using a wide range of predictor variables and including multiple measures of those variables to assess both verbal and spatial processing. Second, the study was designed to determine whether the predictors of lipreading ability change as a function of age. One explanation for the age-related impairments in lipreading observed in previous investigations (cf. Sommers et al., 2005) is that older adults use different and less efficient processes for visual speech perception than do younger adults. A second possibility is that younger and older adults rely on similar mechanisms for visual speech perception and that age differences simply reflect declines in one or more of these abilities. Determining the extent to which lipreading is mediated by distinct mechanisms in older and younger adults would serve as a first step toward understanding why, despite a greater need to rely on visual speech information, older adults are less able than younger adults to lipread.

Method

Participants

Forty-three younger adults (mean age = 20.6 years, *SD* = 2.4; 10 men and 33 women) were recruited from the Washington University student population through the university's participant recruitment Web site. Thirty-eight community-dwelling older adults (mean age = 76.8 years, *SD* = 5.6; 5 men and 33 women) were recruited from the participant pool maintained by the Aging and Development Program at Washington University in St. Louis. The ratio of men to women is typical of studies involving older adults, and the younger group was selected to approximately mirror this ratio. Using the Telephone Interview for Cognitive Status (Brandt & Folstein, 2003), older participants were screened for normal cognitive function (maximum possible value = 41, nonimpaired range = 33–41, group average = 35.4). To minimize the influence of visual pathologies on lipreading ability, all participants were screened for normal vision before testing. Participants whose vision or corrected vision exceeded 20/40, as determined by a Snellen eye chart, were excluded. All participants reported English as their native language and received $20 or two course credits for 2 hr of participation.

Procedure

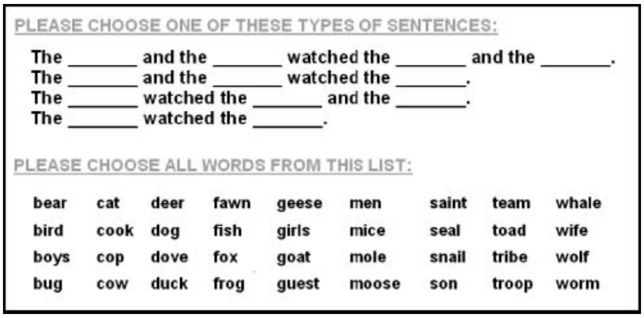
Participants were tested individually and completed a lipreading task and nine predictor tasks: (a) four WM span tests, (b) three PS tasks, and (c) two measures of perceptual abilities. Participants sat in front of a 17-in. Touchsystems monitor (ELO-170C) and responded to test stimuli by touching the screen, pressing buttons, or speaking aloud. Tests were administered using E-Prime and SuperLab presentation software. Instructions were given orally and in writing, and participants completed practice trials for each task to demonstrate that they understood the instructions.

Stimuli

Lipreading Task

The Build-a-Sentence (BAS) test is a closed-set lipreading task that was developed to avoid floor level performance that can sometimes occur in tests of lipreading (Tye-Murray, Sommers, et al., 2008). In the BAS test, participants view a female speaker from the shoulders up producing one of four possible sentence constructions. Every sentence contained the verb "watched," with either one or two blanks preceding and following the verb. After seeing each sentence produced by a speaker, participants viewed a response screen displaying all four potential sentence types with blanks in the place of words. In addition, they saw 36 word options for completing the blanks. Figure 1 shows the response screen, with all potential sentence forms and word choices. All of the word choices had one of nine initial consonants (b, c, d, f, g, m, s, t, w), with four choices corresponding to each of the nine word-initial phonemes. Participants were asked to respond verbally by producing the entire sentence frame with their choice of words for each blank. Participants were encouraged to guess whether they were uncertain about a response. Materials were high-quality digital recordings of an actor speaking sentences in a General American English dialect. All materials were recorded with sound but administered without (for further detail on test construction, see Tye-Murray, Sommers, et al., 2008).

Figure 1. Response screen for the lipreading Build-a-Sentence test.



PLEASE CHOOSE ONE OF THESE TYPES OF SENTENCES:

The ______ and the ______ watched the ______ and the ______.
The ______ and the ______ watched the ______.
The ______ watched the ______ and the ______.
The ______ watched the ______.

PLEASE CHOOSE ALL WORDS FROM THIS LIST:

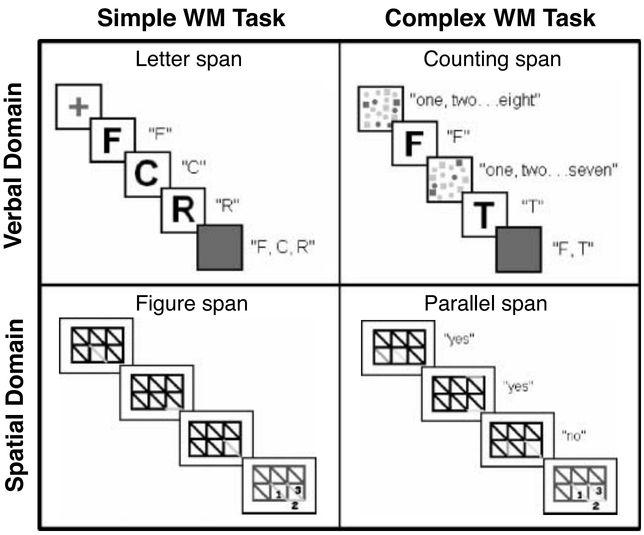
bear	cat	deer	fawn	geese	men	saint	team	whale
bird	cook	dog	fish	girls	mice	seal	toad	wife
boys	cop	dove	fox	goat	mole	snail	tribe	wolf
bug	cow	duck	frog	guest	moose	son	troop	worm

Preliminary validation studies of the BAS test. The BAS test has been demonstrated to be a reliable measure of lipreading, with a test–retest correlation of $r = .90$. Scores on the BAS test are significantly correlated with other sentence-length measures of lipreading ability, including Boothroyd, Hanin, and Hnath-Chisolm's (1985) City University of New York Sentence Test of Speech Perception ($r = .64, p < .001$; Tye-Murray, Spehar, et al., 2008). Pilot studies in our own laboratory have also found correlations between the BAS test and measures of V consonant recognition ($r = .73, p < .001$), V vowel recognition ($r = .75, p < .001$), and lipreading single words without constraining context ($r = .84, p < .001$) using the Children's Audiovisual Enhancement Test (Tye-Murray & Greers, 2001; for details on test construction, see Sommers et al., 2005).

WM

Participants completed both the SWM and VWM tasks. Figure 2 presents schematic illustrations of each task. One task in each domain was a simple span task that involved only recall of the presented items. The second task in each domain was a complex span task, in which participants had to maintain and recall information while completing a secondary task (Engle, Tuholski, Laughlin, & Conway, 1999; Jenkins, Myerson, Hale, & Fry, 1999). In each of the WM tasks, the to-be-remembered items were presented on the screen individually every 2 s. At the end of each trial in the verbal tasks, participants were prompted by the appearance of a green box to recall the series aloud, and their responses were digitally recorded. At the end of each trial in the spatial tasks, participants were prompted by a change in the color of the display (from black to green) to recall the series. Responses were recorded using a touch-sensitive

Figure 2. Working memory (WM) task examples.



computer screen. To ensure that participants attended to the secondary tasks on the complex measures, all participants were required to maintain 85% accuracy on secondary tasks for both the SWM and VWM complex tasks. Order of presentation of trial lengths was randomized once, and the same order was used for all participants so that each participant saw every series length, regardless of performance. No feedback was given during practice or testing.

Simple VWM (letter span). Participants viewed a series of capitalized, black consonants on a beige background, presented in the center of the computer screen to appear approximately 1.5 in. high. The letters were presented once every 2 s, and participants were instructed to speak the letters aloud as they appeared. At the end of a series, a green box replaced the last letter, and participants recalled the consonants aloud in the order they were presented. Two trials of each list length (from 2 letters to 11 letters) were presented to each participant, for a total of 20 trials. Prior to the main test, participants completed six practice trials: two trials each at series lengths of two, three, and four items.

Complex VWM (counting span). The procedure for the counting span task was the same as for the letter span task, but in addition, participants performed a secondary task between each letter presentation. In this secondary task, participants viewed an array of geometric shapes that consisted of green and blue circles and squares, and they counted aloud the number of blue circles (shown in light gray in Figure 2) present. Two trials of each list length (from two letters to seven letters) were presented to each participant, for a total of 12 trials. Prior to the main test, participants received four practice trials of two-item series.

Simple SWM (figure span). Participants viewed a black grid on a beige background in the center of the computer screen. Individual segments of the grid were highlighted in red (shown in light gray in Figure 2) for 2 s. After a series of segments were presented, the grid turned green, and participants recalled the locations of the red line segments by touching the segments on the computer screen. Two trials of each list length (from three segments to 11 segments) were presented to each participant, for a total of 18 trials. Prior to the main test, participants completed six practice trials: two trials each of series lengths of two, three, and four items.

Complex SWM (parallel span). The procedure was the same as for the figure span task, but two segments were highlighted in each presentation: one in red (shown in light gray in Figure 2) and one in blue (shown in dark gray). Participants judged whether the red and blue lines were parallel, and they spoke their responses aloud. After a series of segments were presented, participants recalled the locations of the red lines by touching the segments on the computer screen. Two trials of each list length (from two segments to eight segments) were presented to each participant, for a total of 14 trials. Prior to the main test, participants completed six practice trials: two trials each of series lengths of two, three, and four items.

Trials were only counted as accurate if the participant recalled all the items it contained in the order they appeared. *WM span* was defined as one-half point less than the longest span length at which the participant accurately recalled one of two trials. If the participant accurately recalled both of the trials at that length, an additional half point was awarded. For example, if the participant accurately recalled both of the trials at lengths of 2, 3, 4, and 5, and one of two trials at 6 correct, they would have a span of 5.5. If, however, they recalled both trials at a length of 6, but neither at a length of 7, they would be awarded a span of 6.0.

PS

Three PS tasks were used: a lexical decision task, a category judgment task, and a rhyme judgment task. In each task, words were presented in the center of the computer screen to appear approximately 1.5-in. tall, and participants made judgments about them by pressing keys on the keyboard. They were instructed to respond as quickly and accurately as possible. Prior to each task, participants completed 10 practice trials. Each task had 40 items, 20 of which were "yes" responses and 20 of which were "no" responses. Accuracy and reaction time in milliseconds were recorded by E-Prime stimulus presentation software.

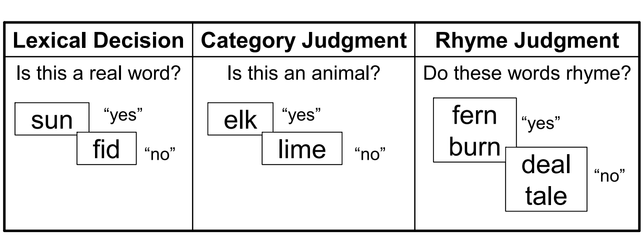
In the lexical decision task, participants saw strings of three letters on the screen. They were instructed to determine whether the letter string formed a real English word and to indicate their response by pressing either a "yes" or "no" key as quickly and accurately as possible. Nonwords were formed by replacing single letters of words that were all phonotactically legal words in English. In the category judgment task, participants saw words appear on the computer screen and made judgments as to whether the word was an animal by pressing either a "yes" or "no" key. All "yes" targets were typical animal names (monkey, cat, rabbit), and all "no" targets were typical fruit and vegetable names (lime, apple, carrot). In the rhyme judgment task, participants saw pairs of words appear on the computer screen and made judgments as to whether the two words rhymed. The rhyme words were orthographically dissimilar both for "yes" targets and for "no" targets. Figure 3 shows examples for each task type.

A mean reaction time for each test was calculated by averaging the reaction times in milliseconds on correct trials, excluding values that fell outside 2.5 *SD*s of the mean. For both younger and older adults, these outliers composed approximately 2% of the trials. Accuracy for all PS tasks was extremely high (95% for the younger adults and 97% for the older adults).

Perceptual Ability

Perceptual closure. The Fragmented Sentence Task (FST; Watson et al., 1996) contains 35 sentences of 5–12 words taken from the City University of New York Lipreading Test (Boothroyd et al., 1985). These sentences were of low-to-moderate semantic predictability—for example, "It is going to be very windy today," and "The sleeves are too long." Sentences were presented in Arial font on the computer screen. Portions of the overall sentence were erased using Adobe Photoshop by removing randomly shaped clusters of pixels, such that part of every letter was removed, but no letter was completely erased. To determine the percentage of the stimulus to erase, while avoiding both floor and ceiling effects, we conducted a pilot study in which progressively greater amounts of information were removed until word identification averaged 75% across individuals (for additional details, see Watson et al., 1996). On the basis of the pilot data, we elected to remove clusters of 60 pixels from each letter. Participants saw the sentences displayed on the computer screen for 3 s, and they read the sentence aloud to the best of their ability. Figure 4 shows examples of the degraded stimuli. The next trial was initiated by the participant by a key press. The verbal responses were recorded by an experimenter. A participant's score was the number of words in position correctly identified.

Figure 3. Processing speed task examples. For all tasks, only one trial appeared on the screen at a time; however, two trials are shown here to illustrate both positive and negative response types.



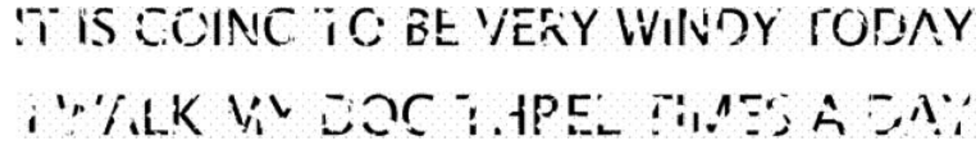
Perceptual disembedding. In the Embedded Figures Test (Witkin, Oltman, Raskin, & Karp, 2002), participants received a paper booklet that had eight simple geometric shapes printed on the back page. Within the booklet were printed complex geometric shapes, which included one of the simple shapes within it and had additional lines to make the simple shape difficult to identify. Participants were instructed to locate a specific simple form in each complex shape and to trace it in pencil. There were two blocks—each contained nine trials—and participants were instructed to complete as many trials as possible in 5 min. Prior to testing, participants completed a 2-min practice block. A participant's score was the number of forms correctly identified in the two 5-min trials.

Results

Means and standard deviations for the predictor (cognitive and perceptual tasks) and criterion (lipreading) measures for both younger and older adults are displayed in Table 1. In line with previous research, lipreading ability was highly variable between individuals, with the correct percentage ranging from 1% to 72% correct (11%–72% for younger adults and 1%–56% for older adults). Overall, lipreading scores for younger adults were significantly higher than for older adults, $t(80) = 6.17$, $p < .001$, although there was considerable overlap in ability between older and younger adults. In addition to higher lipreading scores, younger adults had longer WM spans, had faster PS, and accurately identified more items in the perceptual tasks than did older adults.

To examine the relationships between individual predictor variables and the lipreading task, Pearson product–moment correlations were calculated between lipreading scores and the predictor variables. The results of these analyses are shown in Table 2. Lipreading correlated significantly with each of the cognitive and perceptual constructs measured. For each predictor construct (VWM, SWM, PS, and perceptual ability [PA]), there were robust correlations between the multiple measures of that construct. Accordingly, composite scores were calculated for each predictor construct. Composite

Figure 4. Sample stimuli from the Fragmented Sentence Task.



VWM scores were obtained by calculating the mean *z* score span for letter and counting span, and composite SWMs were calculated in the same manner but using figure span and parallel span. A PS composite was calculated by obtaining the mean *z* score reaction time in milliseconds for the three PS tasks. A PA composite was obtained by averaging the correct percentage for the Embedded Figures Test and the FST. In addition, the results displayed in Table 2 highlight one of the principal difficulties with interpreting findings from previous studies investigating predictors of lipreading—namely, the presence of strong interrelations between all of the predictor variables.

To determine the unique variance that each predictor variable contributed to lipreading performance, we conducted a hierarchical multiple regression analysis with lipreading ability as the criterion variable. Priority of entry for the variables was specified, with SWM entered first, PS entered second, age entered third, PA entered fourth, and VWM entered last. This order was determined by a preliminary stepwise regression analysis that showed that SWM and PS accounted for the greatest amounts of variance in the criterion variable. This preliminary regression also revealed that despite significant correlations between lipreading and age, VWM, and PA, the variance in lipreading accounted for by these predictors was completely redundant to that accounted for by SWM and PS. The results of the hierarchical regression are shown in Table 3. To test whether the predictive power of SWM and PS was due solely to order of entry, another regression was conducted, including SWM and PS as the last, rather than the first, variables entered. After controlling for age, VWM, PA, and PS, SWM still accounts for a small but significant portion of the variance in lipreading ability ($\Delta R^2 = .04, p < .05$). This is also true for PS after controlling for age, VWM, PA, and SWM ($\Delta R^2 = .05, p < .01$). This demonstrates that PS and SWM account for additional unique variance in lipreading ability beyond that explained by other measures, and their observed influence in the regression analysis is not simply due to order of entry.

Table 1. Means and standard deviations for tasks by age group.

	Younger		Older		
Ability	***M***	***SD***	***M***	***SD***	***t***
BAS test lipreading (% correct)	0.44	0.13	0.26	0.14	6.00***
Verbal working memory (VWM)					
Simple VWM (letter span)	5.2	0.9	4.5	0.8	3.77***
Complex VWM (counting span)	4.5	1.3	3.3	1.0	4.58***
Spatial working memory (SWM)					
Simple SWM (figure span)	6.9	1.8	4.4	1.5	6.78***
Complex SWM (parallel span)	5.4	1.7	3.0	1.3	7.13***
Processing speed (reaction time in milliseconds)					
Category judgment	575	80	742	120	−7.51***
Rhyme decision	1,332	295	1,803	362	−6.36***
Lexical decision	627	116	903	187	−7.86***
Perceptual ability (% correct)					
Embedded Figures Test	0.70	0.24	0.23	0.21	9.30***
Fragmented Sentence Task	0.71	0.10	0.63	0.14	3.00***

Note. BAS test = Build-a-Sentence test.

***$p < .001$.

Table 2. Pearson correlations between lipreading ability and predictor variables.

		VWM		SWM		PS			PA	
Variable	**Age**	**L span**	**C span**	**F span**	**P span**	**Animal**	**Rhyme**	**Word**	**FST**	**EFT**
BAS	−.56**	.27*	.43**	.59**	.54**	−.60**	−.52**	−.53**	.25*	.54*
Age	—	−.39**	−.45**	−.60**	−.62**	.65**	.59**	.67**	−.27*	−.72**
L span		—	.41**	.33**	.35**	−.24*	−.25*	−.17	.11	.44**
C span			—	.55**	.50**	−.39**	−.37**	−.29**	.39**	.51**
F span				—	.80**	−.51**	−.36**	−.47**	.26*	.59**
P span					—	−.60**	−.34**	−.49**	.34**	.68**
Animal						—	.71**	.85**	−.33**	−.59**
Rhyme							—	.79**	−.24*	−.45**
Word								—	−.21	−.52**
FST									—	.31**

Note. PS = processing speed; PA = perceptual ability; L span = letter span; C span = counting span; F span = figure span; P span = parallel span; Animal = category judgment task; Rhyme = rhyme judgment task; Word = lexical decision task; FST = Fragmented Sentence Task; EFT = Embedded Figures Test.

*$p < .05$. **$p < .01$.

As shown by the standardized regression coefficient (β) in the first column, the largest correlations with lipreading were observed for SWM and PS scores, and this was true for both the younger and older groups. The second column (R^2) is the total variance accounted for as each predictor was added to the model. The third column (ΔR^2) shows the amount of variance explained independently by each predictor given order of entry. In this case, SWM and PS both contribute significant unique variance to lipreading ability and, in total, account for 46% of the variance in lipreading for all participants combined. Of particular importance to the current study is that after controlling for variance in both SWM and PS, age did not contribute significant unique variance to lipreading ability.

Table 3. Summary of multiple regression for variables predicting lipreading ability ($N = 81$).

	All			Younger			Older		
Variable	**β**	**R^2**	**ΔR^2**	**β**	**R^2**	**ΔR^2**	**β**	**R^2**	**ΔR^2**
Step 1									
SWM	.60	.36	.36**	.39	.15	.15*	.35	.13	.13*
Step 2									
PS	−.39	.46	.10**	−.30	.24	.09*	−.29	.21	.08†
Step 3									
Age	−.10	.47	.01	.17	.27	.03	−.05	.21	.00
Step 4									
PA	.16	.48	.01	.12	.29	.02	.19	.24	.03
Step 5									
VWM	.05	.48	.00	.05	.29	.00	.02	.24	.00

†$p = .06$. *$p < .05$. **$p < .01$.

Discussion

The present study was conducted to investigate factors that contribute to lipreading ability and to determine whether similar abilities mediate lipreading in younger and older adults. Consistent with previous findings (Cienkowski & Carney, 2002; Dancer et al., 1994; Honnell et al., 1991; Shoop & Binnie, 1979; Sommers et al., 2005; Spehar et al., 2004), older adults exhibited poorer lipreading performance than younger adults. Lipreading ability correlated significantly with all of the predictor measures, including VWM and SWM, PS, and performance on verbal and spatial perceptual tasks. Multiple regression analyses, however, indicated that only SWM and PS accounted for significant unique variance.

One potential limitation of the current findings is the use of the BAS test as a measure of lipreading ability. Although the BAS test has high reliability (Tye-Murray, Spehar, et al., 2008), it has relatively low ecological validity; it uses semantically ambiguous sentences in a closed-set format, neither of which replicate typical listening situations. We elected to use the BAS test because it overcomes floor effects that are often seen in studies assessing lipreading with sentence-length materials (Dancer et al., 1994; Tye-Murray et al., 2007). In addition, it allowed us to provide syntactic information that has been shown to improve speech perception compared with single words (Sommers & Danielson, 1999). The pilot results examining correlations between the BAS test and other measures of lipreading seem to indicate that similar results would be obtained with other measures as well. Nevertheless, it will be important to replicate these results using speech materials that more closely approximate real-world listening situations.

These findings replicate and extend earlier findings by Lidestam et al. (1999) and Lyxell and Holmberg (2000) that indicated significant correlations between VWM and lipreading. However, neither of those earlier studies included separate measures of SWM, and it therefore remained unclear which type of WM was related to lipreading ability. The importance of this issue is illustrated in Table 2, which indicates significant correlations between VWM and SWM as well as between other variables that we used to predict lipreading. This collinearity among the predictor variables suggests that additional analyses are required to determine the factors that uniquely account for performance in lipreading tasks. In the present study, we used multiple regression analyses and assessed SWM to demonstrate that traditional measures of visuo-spatial processing, and not measures of VWM, account for unique variance in individual differences in lipreading.

In this regard, it is interesting to note that the extant cognitive aging literature has consistently demonstrated significantly greater declines in SWM than in VWM (Jenkins et al., 2000; Myerson et al., 1999). Considered with the present findings, the greater age-related declines in SWM than in VWM suggest that older adults' poorer lipreading abilities may be in part a consequence of reduced SWM. Moreover, to our knowledge, the current findings also provide the first evidence that older and younger adults rely on a similar set of mechanisms to understand visual only speech information. Therefore, age-related declines in lipreading ability would seem to result from impairments to fundamental cognitive abilities rather than to any age-related change in the mechanisms mediating V performance in younger and older adults.

Because language processing requires temporarily storing and manipulating rapidly presented information, it is not surprising that WM is one component skill necessary for successful interpretation of the information (for a meta-analysis of WM and language comprehension, see Daneman & Merikle, 1996). Although previous studies have not included measures of SWM as a predictor of lipreading performance, the significant relationship between these two measures observed in the present study is perhaps not surprising. SWM has been found to be important for understanding visually mediated language tasks, such as reading (Baddeley, 2003). Thus, although systematic models of lipreading are not yet available, the current findings provide an important constraint for the development of such models. Specifically, the results suggest that any comprehensive model of lipreading will need to include a component that functions to store a sequence of visually observed movements and then combine those movements into a unified percept.

The relationship between cognitive mechanisms and lipreading fits the framework proposed by Pichora-Fuller, Schneider, and Daneman (1996), who found that in an auditory context, as perceptual processing becomes more difficult (either by a decreased S/N ratio or, in this case, the impoverished visual signal), the demands placed on cognitive mechanisms (such as WM) are increased. Even in ideal circumstances, lipreading is a perceptually difficult task because there is not a direct correspondence between lip movement and sound, resulting in an incomplete and often ambiguous signal. The ambiguity of the signal is likely one reason that word identification in V is more difficult than in A, and the task difficulty caused by this ambiguity may also help to explain why successful lipreading performance depends on cognitive abilities, such as WM.

The present results also serve to explain the somewhat surprising finding that lipreading is largely unrelated to postlingual hearing loss (Clouser, 1977; Farrimond, 1959; Lyxell & Rönnberg, 1989, 1991; Owens & Blazek, 1985; Rönnberg, 1990) in that cognitive (SWM and PS) abilities, rather than absolute sensitivity, seem to be the primary determinants of lipreading performance. Furthermore, if, as the data suggest, lipreading ability depends on fundamental cognitive traits that are stable in adulthood, it is not surprising that lipreading training often leads to only modest improvements in V performance that may or may not be clinically significant (Tye-Murray, 2008).

As discussed in the introduction, lipreading ability is a critical factor in determining how much individuals benefit from the addition of visual speech information (Macleod & Summerfield, 1990; Sommers et al., 2005). Considered with the current findings, this result suggests that older adults are at a significant disadvantage during face-to-face communication; presbycusic hearing loss reduces overall audibility, and impairments in lipreading make it difficult to compensate for the age-related sensory decline using visual speech information. The situation is further complicated because, as noted, training on lipreading has met with only limited success. However, it may be that a combination of practice with lipreading and training on basic cognitive abilities, such as SWM, will be able to produce clinically significant benefits for hearing-impaired individuals.

Acknowledgments

Portions of these data were presented at the 48th annual meeting of the Psychonomics Society on November 15, 2007, in Long Beach, California, and at the Cognitive Aging Conference on April 12, 2008, in Atlanta, Georgia. We gratefully acknowledge the support of the Gustavus and Louise Pfeiffer Foundation and National Institutes of Health Grant AG 022448. We thank Anna Dinndorf and Rachel Wilkinson for their assistance in data collection.

References

Aloufy, S., Lapidot, M., & Myslobodsky, M. (1996). Differences in susceptibility to the "blending illusion" among native Hebrew and English speakers. *Brain and Language, 53,* 51–57.

Auer, E. T., Jr., & Bernstein, L. E. (2007). Enhanced visual speech perception in individuals with early-onset hearing impairment. *Journal of Speech, Language, and Hearing Research, 50,* 1157–1165.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36,* 189–208.

Bernstein, L. E., Auer, E. T., Jr., & Tucker, P. E. (2001). Enhanced speechreading in deaf adults: Can short-term training/practice close the gap for hearing adults? *Journal of Speech, Language, and Hearing Research, 44,* 5–18.

Bernstein, L. E., Demorest, M. E., & Tucker, P. E. (2000). Speech perception without hearing. *Perception & Psychophysics, 62,* 233–252.

Boothroyd, A., Hanin, L., & Hnath-Chisolm, T. (1985). *A sentence test of speech perception: Reliability, set-equivalence, and short-term learning* (Internal Report RCI 10). New York, NY: City University of New York.

Brandt, J., & Folstein, M. (2003). *The Telephone Interview for Cognitive Status*. Lutz, FL: Psychological Assessment Resources.

Cienkowski, K. M., & Carney, A. E. (2002). Auditory–visual speech perception and aging. *Ear and Hearing, 23,* 439–499.

Clouser, R. A. (1977). Relative phoneme visibility and lipreading performance. *Volta Review, 79,* 27–34.

Dancer, J., Krain, M., Thompson, C., Davis, P., & Glen, J. (1994). A cross-sectional investigation of speechreading in adults: Effects of age, gender, practice, and education. *Volta Review, 96,* 31–40.

Daneman, M., & Merikle, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3,* 422–433.

Elphick, R. (1996). Issues in comparing the speechreading abilities of hearing-impaired and hearing 15 to 16 year-old pupils. *British Journal of Educational Psychology, 66,* 357–365.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128,* 309–331.

Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research, 12,* 423–425.

Farrimond, T. (1959). Age differences in the ability to use visual cues in auditory communication. *Language and Speech, 2,* 179–192.

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory and fluid intelligence in children. *Biological Psychology, 54,* 1–34.

Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory–visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory–visual integration. *The Journal of the Acoustical Society of America, 103,* 2677–2690.

Honnell, S., Dancer, J., & Gentry, B. (1991). Age and speechreading performance in relation to percent correct, eyeblinks, and written responses. *Volta Review, 93,* 207–212.

Irwin, J. R., Whalen, D. H., & Fowler, C. A. (2006). A sex difference in visual influence on heard speech. *Perception & Psychophysics, 68,* 582–592.

Jeffers, J., & Barley, M. (1971). *Speechreading (lipreading)*. Springfield, IL: Thomas.

Jenkins, L., Myerson, J., Hale, S., & Fry, A. (1999). Individual and developmental differences in working memory across the life span. *Psychonomic Bulletin & Review, 6,* 28–40.

Jenkins, L., Myerson, J., Joerding, J. A., & Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. *Psychology and Aging, 15,* 157–175.

Johnson, F. M., Hicks, L. H., Goldberg, T., & Myslobodsky, M. S. (1988). Sex differences in lipreading. *Psychonomic Bulletin & Review, 26,* 106–108.

Lidestam, B., Lyxell, B., & Andersson, G. (1999). Speechreading: Cognitive predictors and displayed emotion. *Scandinavian Audiology, 28,* 211–217.

Lyxell, B., & Holmberg, I. (2000). Visual speechreading and cognitive performance in hearing-impaired and normal hearing children (11–14 years). *British Journal of Educational Psychology, 70,* 505–518.

Lyxell, B., & Rönnberg, J. (1989). Information-processing skill and speech-reading. *British Journal of Audiology, 23,* 339–347.

Lyxell, B., & Rönnberg, J. (1991). Word discrimination and chronological age related to sentence-based speech-reading skill. *British Journal of Audiology, 25,* 3–10.

Lyxell, B., & Rönnberg, J. (1992). The relationship between verbal ability and sentence-based speechreading. *Scandinavian Audiology, 21,* 67–72.

Lyxell, B., & Rönnberg, J. (1993). The effects of background noise and working memory capacity on speechreading performance. *Scandinavian Audiology, 22,* 67–70.

MacLeod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *British Journal of Audiology, 21,* 131–141.

MacLeod, A., & Summerfield, Q. (1990). A procedure for measuring auditory and audio–visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *British Journal of Audiology, 24,* 29–43.

Middelweerd, M. J., & Plomp, R. (1987). The effect of speechreading on the speech-reception threshold of sentences in noise. *The Journal of the Acoustical Society of America, 82,* 2145–2147.

Myerson, J., Hale, S., Rhee, S. H., & Jenkins, L. (1999). Selective interference with verbal and spatial working memory in young and older adults. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 54,* 161–164.

Owens, E., & Blazek, B. (1985). Visemes observed by hearing-impaired and normal-hearing adult viewers. *Journal of Speech and Hearing Research, 28,* 381–393.

Pichora-Fuller, M., Scheider, B., & Daneman, M. (1996). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America, 97,* 593–608.

Rönnberg, J. (1990). Cognitive and communicative function: The effects of chronological age and "handicap age." *European Journal of Cognitive Psychology, 2,* 253–273.

Sanders, J. W., & Coscarelli, J. E. (1970). The relationship of visual synthesis skill to lipreading. *American Annals of the Deaf, 115,* 23–26.

Sharp, E. Y. (1972). The relationship of visual closure to speechreading. *Exceptional Children, 38,* 729–734.

Shoop, C., & Binnie, C. A. (1979). The effects of age upon the visual perception of speech. *Scandinavian Audiology, 8,* 3–8.

Simmons, A. A. (1959). Factors related to lipreading. *Journal of Speech and Hearing Research, 2,* 340–352.

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging, 14,* 458–472.

Sommers, M., Tye-Murray, N., & Spehar, B. (2005). Auditory–visual speech perception and auditory–visual enhancement in normal-hearing younger and older adults. *Ear and Hearing, 26,* 263–275.

Spehar, B., Tye-Murray, N., & Sommers, M. (2004). Time-compressed visual speech and age: A first report. *Ear and Hearing, 25,* 565–572.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America, 26,* 212–215.

Summerfield, Q. (1991). Visual perception of phonetic gestures. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 117–137). Hillsdale, NJ: Erlbaum.

Tye-Murray, N. (2008). *Foundations of aural rehabilitation: Children, adults, and their family members*. Clifton Park, NY: Delmar Cengage Learning.

Tye-Murray, N., & Greers, A. (2001). *Children's Audiovisual Enhancement Test*. St, Louis, MO: Central Institute of the Deaf.

Tye-Murray, N., Sommers, M., & Spehar, B. (2007). The effects of age and gender on lipreading abilities. *Journal of the American Academy of Audiology, 18,* 883–892.

Tye-Murray, N., Sommers, M., Spehar, B., Myerson, J., Hale, S., & Rose, N. (2008). Auditory–visual discourse comprehension by older and young adults in favorable and unfavorable conditions. *International Journal of Audiology, 47,* 31–37.

Tye-Murray, N., Spehar, B., Rose, N., Sommers, M., Hale, S., & Myerson, J. (2008, March 27–29). *Assessing speechreading performance: Sentences and discourse comprehension*. Paper presented at the 3rd Med-El Aural Rehabilitation Conference: Assessment, Pediatric Habilitation, Adult Rehabilitation, Den Bosch, the Netherlands.

Watson, C. S., Qiu, W. W., Chamberlain, M. M., & Li, X. (1996). Auditory and visual speech perception: Confirmation of a modality-independent source of individual differences in speech recognition. *The Journal of the Acoustical Society of America, 100,* 1153–1162.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (2002). *A manual for the Group Embedded Figures Tests*. Menlo Park, CA: Mind Garden.

Zekveld, A., Kramer, S., Vlaming, S., & Houtgast, T. (2008). Audiovisual perception of speech in noise and masked written text. *Ear and Hearing, 29,* 99–111.

Received July 8, 2008

Accepted February 6, 2009

DOI: 10.1044/1092-4388(2009/08-0137)

Contact author: Julia E. Feld, Washington University in St. Louis, Psychology Department, Campus Box 1125, St. Louis, MO 63130. E-mail: juliafeld5@gmail.com.