1999

# Visual speech perception: a comparison between younger and older adults

Valerie Ann Pavlovich

# Visual Speech Perception: A Comparison Between Younger and Older Adults

Independent Study
Valerie Ann Pavlovich

Supervised by Nancy Tye-Murray, Ph.D. and Rosalie Uchanski, Ph.D.
Central Institute for the Deaf

In cooperation with Mitch Sommers, Ph.D.
Washington University, Department of Psychology

April 26, 1999

# INTRODUCTION

Speech perception, though most often considered an auditory event, can more accurately be described as a multimodal phenomenon. People with normal hearing and vision use both auditory and visual cues in understanding speech although they may be less aware of the visual speech component (Massaro & Stork, 1998). In the case of people with hearing impairment, visual speech cues help to compensate for the loss of acoustic information. For those with profound hearing impairment, vision might even be the primary modality for speech perception rather than audition. When faced with a difficult listening situation, such as carrying on a conversation in a crowded room, even people with normal hearing do much better when watching the speaker's face more closely. Studies have shown that visual cues can actually influence the auditory perception of speech. The McGurk effect may best illustrate this bimodal nature of speech understanding. For example, if one hears the audio syllable "ba" and sees the visual syllable "ga", he/she will perceive the syllable, "da" (Grant, Walden, & Seitz, 1998). Even infants as young as 4 months of age can make use of the visual cues that are available in speech (Dodd, 1987, chap. 7).

As the number of older adults with hearing impairment continues to grow, so does an interest in the speech perception abilities of this population. Many studies that have examined the speech perception abilities of older adults have concentrated on auditory speech perception or auditory-visual integration (Humes, 1996; Goron-Salant, 1986; Watson, Qiu, Chamberlain, & Li, 1996; Gordon-Salant, 1987a; Gordon-Salant, 1987b; Grant et al., 1998; Helfer. 1998). This study is concerned with the visual speech perception abilities of older adults. In reviewing recent literature available on visual speech perception, one will encounter two terms that are often used interchangeably, lipreading and speechreading. The main distinction between these two terms is that lipreading is understood to make use of visual cues alone and speechreading incorporates an auditory signal as well as gestures and contextual cues (Tye-Murray, 1998; Gagné, 1994). This distinction was not made in earlier studies that examined visual speech perception abilities (Farrimond, 1959; Simmons, 1959; Ewertsen & Nielsen, 1971; Pelson & Prather, 1974; Shoop & Binnie, 1979; Honnell, Dancer, & Gentry, 1991). Therefore, any reference to visual speech perception in this discussion will specifically refer to speechreading (or lipreading), defined by Farwell (1976) as the "method of reading speech by watching the lips and face of the speaker" (p. 19). Speechreading or lipreading for the purpose of this discussion will refer to a visual-only condition (i.e. no auditory signal).

An early study conducted by Thomas Farrimond (1959) evaluated the speechreading abilities of a "normal" population of men to determine whether compensatory visual skills developed with increasing age and hearing loss. Subjects ranged in age from 20 to 60+ years. Two sets of stimuli were used, numbers and sentences. Sentence stimuli were used because they produced a more meaningful representation of speech understanding than nonsense syllables or lists of words. Speech stimuli were presented through a silent film. Prior to each sentence presentation, subjects viewed a short scene of an activity that would provide context for the upcoming sentence. Some sentences were more difficult than others. Farrimond's results showed that overall speechreading ability was greatest for subjects in the age range of 30 to 39 years and progressively declined with increasing age despite the increased prevalence of hearing loss in the older men. Data for both types of stimuli showed this decline in speechreading ability with increasing age, although the scores obtained in the numbers test were approximately half those obtained in the sentence test. It was concluded that, since hearing loss progresses slowly, old listening habits are relied on and though visual cues may be used, speechreading skills do not improve without specific training.

In a related study, Ewertsen and Nielson (1971) examined the effects of aging on the audiovisual, auditory, and visual perception of speech. Their subjects were 30 normal hearing subjects, 10 in each of three age groups, 20, 50, and 70 years. Audiovisual and auditory speech perception were tested at signal-to-noise ratios of -20 dB, -10 dB, 0 dB, and +10 dB. White noise was presented at a constant level of 85 dB while the speech signal was varied from 65 dB to 95 dB in 10 dB increments. Twenty-five words were presented at each level. Results showed that at a signal-to-noise ratio of –20 dB, the auditory-only tests showed complete masking of speech while, in the audiovisual tests, 50% of the words were perceived. The authors attributed this improved perception to the role of vision in the audiovisual condition. Speech perception under a visual-only condition, however, resulted in much poorer performance. Overall results revealed that whether auditory or audiovisual, performance decreased with increasing age.

In 1974, Pelson and Prather looked at the effects of visual message-related cues, age, and hearing impairment on speechreading ability. Subjects were divided into three groups of 12. One group was comprised of young, normal-hearing adults between the ages of 19 and 26 years. Another group was made up of older, normal-hearing adults between the ages of 52 and 61 years. The third group was made up of older, hearing-impaired adults between the ages of 51 and 59 years. The average bilateral sensorineural hearing loss for group three was 50 dB HL (ANSI, 1969) with a mean duration of 10.7 years. Two different speechreading tasks were performed. Each task consisted of 12 sentences. One task was performed with visual-only presentation of the sentences, the other task was visual-only but included a 35-mm color slide of a scene that provided context for the sentences. Results revealed a significant difference between the young, normal-hearing group and both older groups. Scores for the young subjects were an average of 35% and 24% above the older subjects. The older adults with hearing loss performed better than the older adults with normal hearing. All subjects performed better on the task that included the 35-mm slide. Pelson and Prather concluded that aging seems to have a detrimental effect on speechreading performance in older adults, even when visual acuity remains unimpaired.

Shoop and Binnie (1979) also studied the effects of age on the visual perception of speech. Subjects for this study were composed of 110 adult men and women with normal hearing and vision. All subjects were divided into four different age groups ranging from 40 to 71+ years. Two sets of stimuli were used to assess speechreading ability, 20 consonants presented in CV syllables and 10 sentences from List 2 of the CID Sentences. Confusion matrices for individuals and groups were generated for later comparison of consonant errors. Results of the consonant test showed a significant decrease in speechreading ability for subjects in the age group 71+ when compared to subjects between the ages of 40 and 70. A very wide range of performance existed for normally hearing adults over 40 years of age. When comparing the mean performance for sentence stimuli to CV syllable performance, sentence performance showed a more consistent decline with increasing age. The authors concluded that "sentences seemed to be a more sensitive measure of age-related effects on the visual speech perception process than viseme categories within CV syllables, although consistent age effects were shown using both of these materials"(p. 8).

An additional study conducted by Honnell et al. in 1991 looked at age and speechreading performance in relation to percent correct, eyeblinks, and written responses. Subjects in this study were divided into 2 age groups, 20-29 and 60-69. There were 14 subjects in each group, all having normal hearing and vision. The CID Everyday Sentences were used as stimuli and a percent correct score was obtained by counting the number of key words correct. The sentences were recorded audiovisually by an adult male speaker. Subjects were instructed to write verbatim what they thought the speaker said and were highly encouraged to guess. The sentences were presented with no sound. Results showed that the older adults scored significantly poorer in terms of percent correct scores than did the younger subjects. Older subjects also wrote fewer words than the younger subjects did, although there was a high correlation between percent correct and number of words written for the older adults. The author attributed this to perhaps a greater reluctance on the part of the older adults to guess when they weren't sure what the speaker was saying. The two groups did not show significant differences in the mean number of eyeblinks that were recorded.

In contrast to the aforementioned studies, Simmons (1959) failed to find a correlation between age and lipreading ability. Her subjects were 12 men and 12 women whose mean age was 47.13 years with a standard deviation of 9.64. A complete range of ages was not given. All of the subjects had been seen in the hearing clinics of Central Institute for the Deaf and had a mean pure-tone hearing loss of 33.85 dB with a standard deviation of 21.10. Lipreading instruction had been recommended but had not been completed by any of the subjects prior to the study. Lipreading abilities were assessed using three tests: the Utley lipreading test, the Mason lipreading test, and an interview rating. Simmons found no significant differences in speechreading ability across subjects based on age.

There are several possible explanations for this decrease in visual speech perception ability in older adults. One theory suggests that this decrease "may be related to the fact that many of the aging changes in the visual system are temporally related, and speech comprehension requires rapid on-line processing of information"(Helfer, 1998, p.234). Age-related changes in visual information processing such as increased visual persistence and increased susceptibility to backward masking in the visual domain, might limit the use of speechreading by the elderly (DiLollo, Arnett, & Kruk, 1982). Studies also show evidence that older adults are more sensitive than younger adults to auditory distortions that are temporal in nature such as time-compressed speech (Helfer, 1998; Wingfield, 1996). In addition, Helfer indicates that slowing in sensorimotor and perceptual processing with increasing age is almost a universal finding in research on aging.

Poorer visual speech perception performance in older adults might be related to cultural or educational differences between younger and older adults. Honnell et al. (1991) suggested that older adults may have used caution in writing what they thought the speaker said because, years ago, emphasis was more on accuracy, neatness, and certainty. This reluctance to guess may have influenced the performance of the older adults in this study. In fact, Farrimond (1959) found that older subjects showed improvement in speechreading performance if they were encouraged to guess. Honnell et al. (1991) also indicated that speechreading performance might be reduced in older adults due to stress, motivation, time constraints, fatigue, visual processing limitations, memory, and slower visual-neural conduction times. Finally, some studies have indicated that central processes may be involved in the decreased visual speech perception ability of older adults (Farrimond, 1959; Pelson & Prather, 1974; Watson et al., 1996).

A better understanding of visual speech processing abilities in the elderly is important to future clinical procedures and aural rehabilitation programs. The present study, therefore, looked at lipreading abilities to determine how well normal-hearing adults of different ages can process visual speech information with the purpose of answering three questions:

1.) Does lipreading ability vary as a function of age? 2.) Do patterns of lipreading performance change with different speech materials? 3.) Does lipreading performance, under visually challenging conditions, vary with age?

## METHODS

### Speech Materials

Twenty-two English consonants were used as stimuli. Each was video-recorded in a VCV sequence with the vowel /a/ (Table I). There were 12 sets of 22 presentations. Each VCV sequence was only used once during each set. Syllables were produced in a natural manner by a female talker using normal voice and articulation.

Twenty-five sentences were also video-recorded (Table II). Five different talkers were used, 3 females and 2 males. Each talker produced 5 sentences. All speech materials were taken from the Iowa Audiovisual Speech Perception video recordings.

### Subjects

Subjects were recruited from two age groups, 18-25 years or 65+ years. There were 16 subjects in the young group and 7 subjects in the older group. Young subjects had normal hearing. Hearing thresholds were obtained on all older subjects using a two alternative forced choice method. All of the older adults had thresholds within normal limits (defined as less than 20 dB hearing loss, re: ANSI, 1989 standards) for octave frequencies between 250 and 4000 Hz (Yantis, 1994). Frequencies above 4000 Hz were not tested. All subjects had normal or corrected-to-normal vision. All subjects spoke English as their primary language and completed a vocabulary test prior to participation. Efforts were made to obtain an equal number of male and female subjects. None of these subjects had any prior lipreading instruction. Subjects were volunteers sought through the Washington University Department of Psychology.

### Procedures

All subjects were seated in front of a 20-inch monitor, approximately 4 feet away. The image of the talker was a front view including the head and shoulders. To alert subjects, the word "READY" appeared on the screen prior to stimulus presentation. The number of the test item also appeared on the screen. Following stimulus presentation, the words "Respond to item #" appeared on the screen for approximately 6 seconds. The video was presented visually with the auditory signal muted. Subjects were given brief rest periods following each set of 22 presentations.

Two different rates of presentation were used to create a visually challenging condition for the VCV materials only. A normal rate of 30 frames per second was presented for 6 complete sets of VCV items. A faster rate of 40 frames per second was used for another 6 sets of VCV items. Twenty-five sentences were presented to each subject at a normal rate of 30 frames per second. The single test session took approximately 2 hours per subject. All subjects were tested individually and received the test stimuli in the same order as listed in Table III.

Sentences were scored by counting every word, not just key words. Words had to be exactly correct to be counted correct. For example, incorrect plurals were considered an incorrect response.

## RESULTS

Table IV shows percent correct scores for all subjects on all lipreading tasks. Table V shows means and standard deviations of percent correct scores for VCV stimuli.

Table VI shows means and standard deviations of percent word correct scores for sentence stimuli. Analysis of the data shows no significant differences between the younger and older adults on any measure. There was a slight decrease in performance for both the younger and older adults when the rate of presentation for the VCV syllables was increased form 30 to 40 frames per second though this difference was not significant. Performance on the sentence test was significantly lower than performance on the VCV syllables for both groups.

**Table I.** Display of the twenty-two VCV syllables used as stimului.

| | |
|---|---|
| aza | afa |
| aga | ava |
| ana | acha |
| ata | apa |
| ama | aka |
| ara | anga |
| ala | aya |
| aba | awa |
| asha | aha |
| ada | asa |
| aja | azha |

**Table II.** List of the twenty-five sentences used as stimuli.

The father watched his son.
The man tied his scarf.
The clock hangs on the wall.
The corner store was open Saturday.
The two children are laughing.
The gardener planted some roses.
The mouse ate the cheese.
The radio was playing too loudly.
The girl dropped her schoolbooks.
The lion roared at the crowd.
The boy's helping his friend.
The students went home for the holidays.
Some brown leaves fell off the tree.
Father left the kitchen in a mess.
The lady waved goodbye to her friend.
It rained all day yesterday.
The dog jumped on the chair.
The ball broke the window.
The girl planted a vegetable garden.
The mother shut the window.
The shoes were very dirty.
The scissors are quite sharp.
Mother picked some flowers.
The girl has a picture book.
Somebody took the money.

**Table III.** Order of stimuli presentation.

| | |
|---|---|
| 1. 3 sets of VCV syllables | Rate: 30 Frames/sec |
| 2. 3 sets of VCV syllables | Rate: 40 Frames/sec |
| 3. 3 sets of VCV syllables | Rate: 40 Frames/sec |
| 4. 3 sets of VCV syllables | Rate: 30 Frames/sec |
| 5. 25 sentences | Rate: 30 Frames/sec |

**Table IV.** Percent correct scores for VCV and sentence stimuli.

| Subject | 30 Frames per second | 40 Frames per second | Difference 30 FPS-40 FPS | Sentences |
|---|---|---|---|---|
| Young 01 | 34 | 37.5 | -3.5 | 29.4 |
| Young 02 | 34.7 | 33.3 | 1.4 | 20.6 |
| Young 03 | 20.1 | 19 | 1.1 | *N/A |
| Young 04 | 41 | 38.2 | 2.8 | *N/A |
| Young 05 | 39.6 | 43.1 | -3.5 | 24.3 |
| Young 06 | 41 | 42 | -1 | 25.7 |
| Young 07 | 37.5 | 34.7 | 2.8 | 14.7 |
| Young 08 | 39.6 | 36.8 | 2.8 | 26.5 |
| Young 09 | 41.7 | 38.2 | 3.5 | 25.7 |
| Young 10 | 52.1 | 46.5 | 5.6 | *N/A |
| Young 11 | 20.8 | 19.4 | 1.4 | 16.2 |
| Young 12 | 48.6 | 44.4 | 4.2 | 25.7 |
| Young 13 | 38.2 | 43.1 | -4.9 | 29.4 |
| Young 14 | 31.9 | 33.3 | -1.4 | 23.5 |
| Young 15 | 41 | 38.2 | 2.8 | 27.9 |
| Young 16 | 25.7 | 25.7 | 0 | 4.4 |
| Old 01 | 43.1 | 41.1 | 2 | 10.3 |
| Old 02 | 43.1 | 40.6 | 2.5 | 31.6 |
| Old 03 | 29.2 | 36.1 | -6.9 | 17.6 |
| Old 04 | 38.2 | 32.9 | 5.3 | 22.8 |
| Old 05 | 37.4 | 24.4 | 13 | 20.1 |
| Old 06 | 39.4 | 31.7 | 7.4 | 24.3 |
| Old 07 | 36.3 | 28.7 | 7.6 | 19.5 |

*N/A means subject did not complete this portion of the test

**TableV.** Means and standard deviations for VCV stimuli at two different rates.

Vowel-Consonant-Vowel

Percent Correct Scores

| | Normal Rate | | Fast Rate | |
|---|---|---|---|---|
| | Mean | Standard Deviation | Mean | Standard Deviation |
| Younger (N=16) | 37 | 8.8 | 36 | 8.3 |
| Older (N=7) | 38 | 4.7 | 34 | 6.1 |

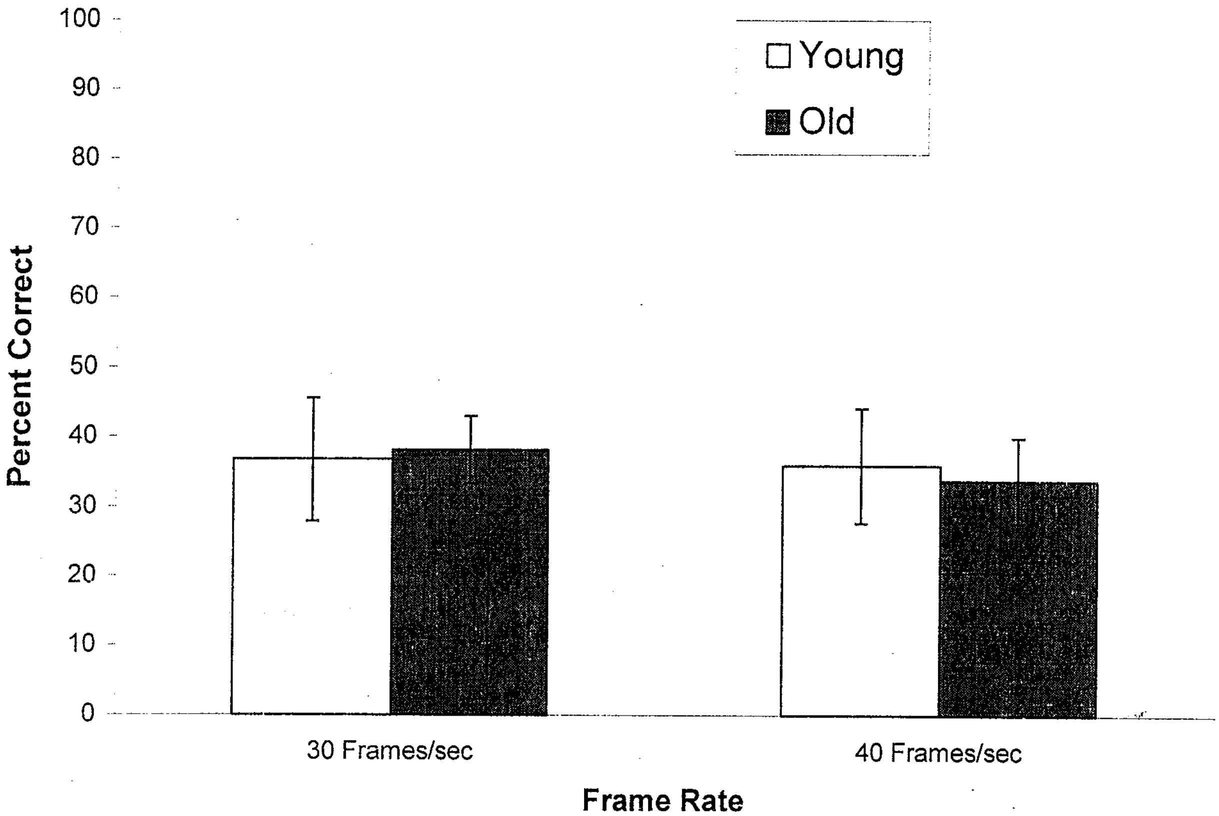
**Table VI.** Percent words correct and standard deviations for sentence stimuli.

Sentences

Percent Correct Scores

| | Mean | Standard Deviation |
|---|---|---|
| Younger (N=13) | 23 | 7.1 |
| Older (N=7) | 21 | 6.5 |

# Figure 1 : VCV Perception



## Figure 2 : Sentence Perception

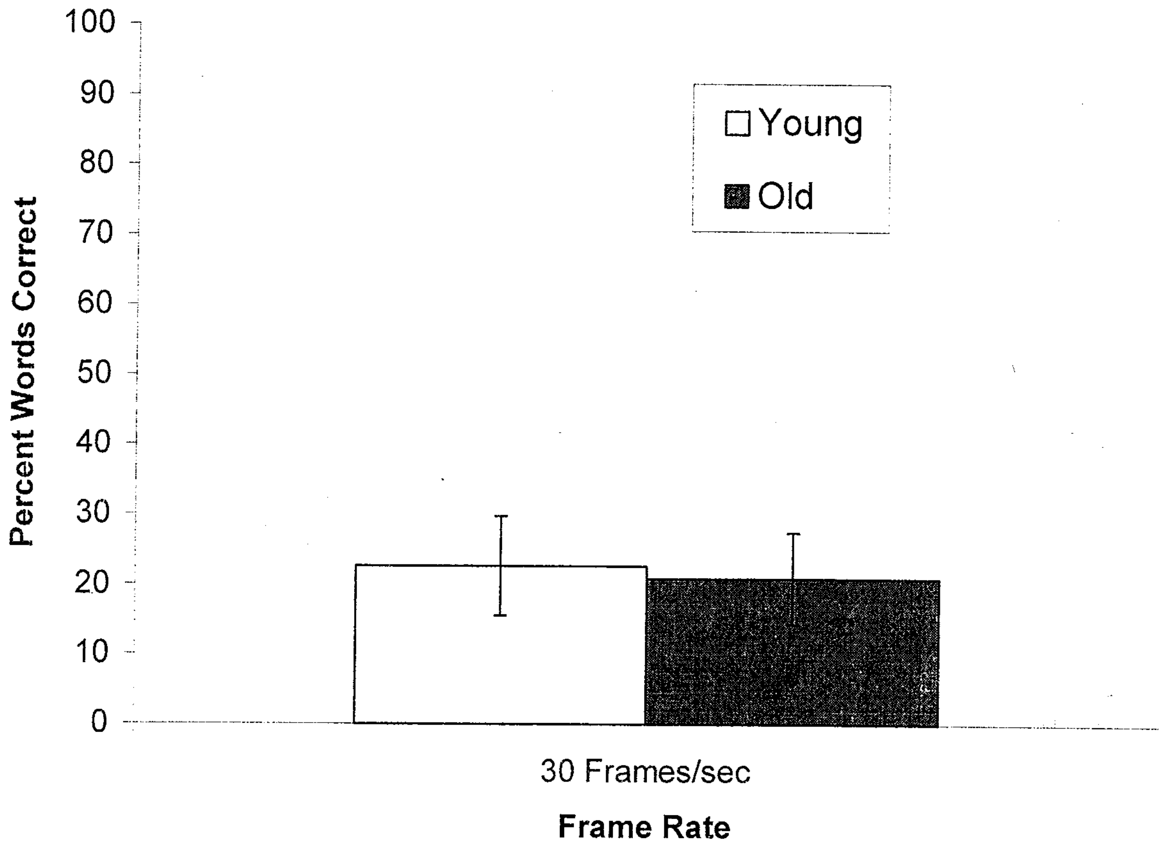
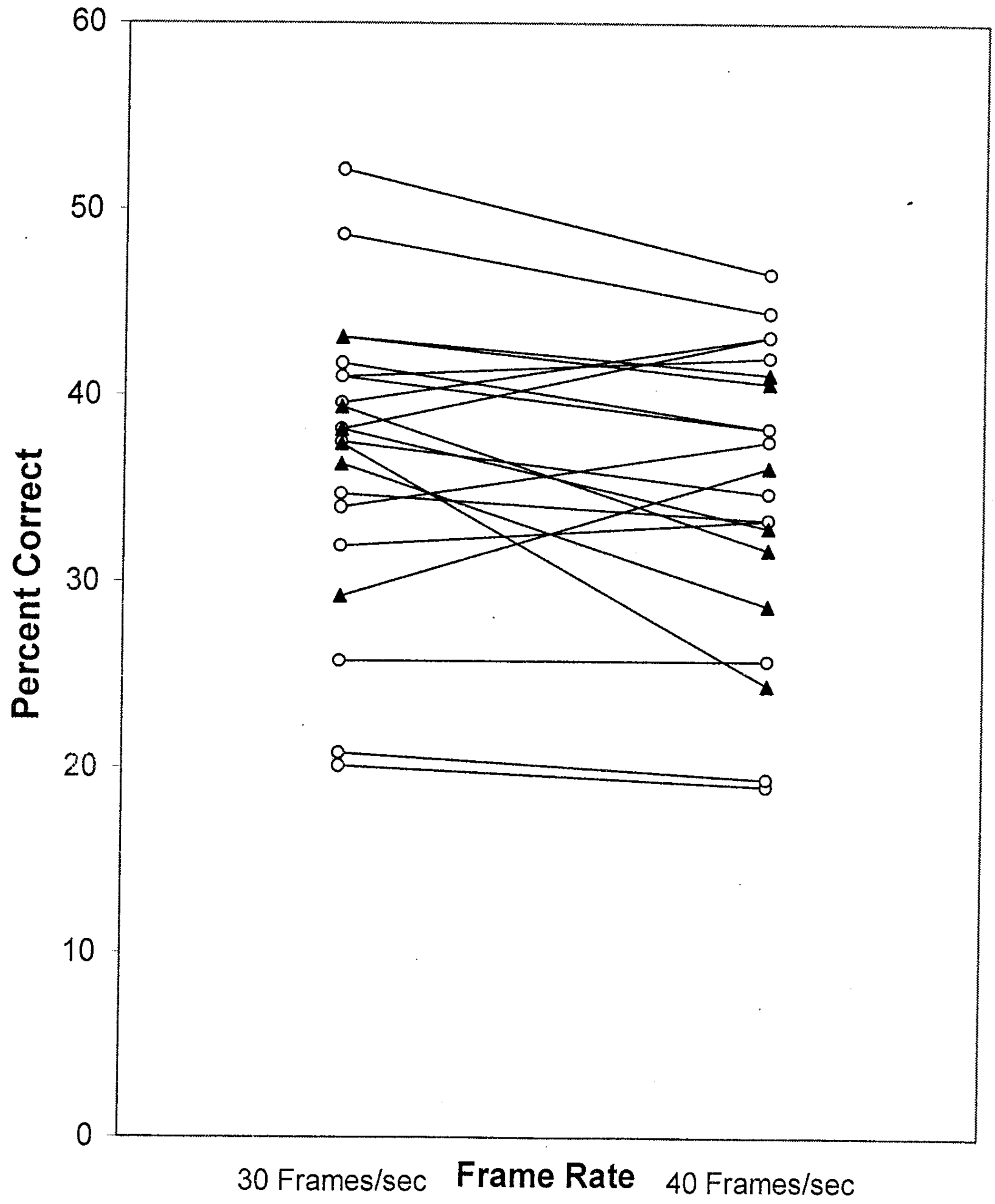


**Figure 3 : Rate differences for Younger and Older Adults**




- Open circles indicate young subjects
- Black triangles indicate older subjects

## DISCUSSION

From these results, one might conclude that under visual-only conditions, older adults process visual speech information equally as well as younger adults. However, "Studies of age differences in rate of visual information processing have shown that young adults can process the contents of visual displays faster than elderly subjects (Hertzog, Williams and Walsh, 1976; Kline and Szafran, 1975; Walsh, 1976, Walsh, Till, and Williams, 1978)" (DiLollo et al., 1982, p. 225). Several earlier studies of visual speech perception in older adults did, in fact, find that speechreading ability decreased with increasing age (Farrimond, 1959; Ewertsen and Nielson, 1971; Shoop and Binnie, 1979; Honnell et al., 1991). In contrast, results of the present study support the research by Simmons (1959) which failed to reveal a significant correlation between age and visual speech perception abilities.

It is unclear why this discrepancy in findings exists between studies as there are many similarities between them. For example, the present study used VCV syllables (with the vowel /a/) containing 22 consonants as well as sentence stimuli for assessing the visual speech perception abilities of younger and older subjects. Shoop and Binnie (1979) used very similar stimuli consisting of CV syllables (with the vowel /a/) and sentence stimuli to assess a similar population of subjects. Shoop and Binnie found a decrease in visual speech perception abilities in their oldest group of subjects, age 71+. Many of the older adults in the present study were also over the age of 70. Perhaps it is the difference in subject size that accounts for the difference. Shoop and Binnie looked at 110 subjects in all, 20 that fell in the age range above 70 years. The present study looked at only 7 older adults in all.

Another possible reason for this contention in results is method of scoring. The present study scored percent correct for sentence stimuli by counting every word that the subject got correct, not just key words. Twenty out of twenty-five sentences began with the word "The", which is highly visible on the lips and may have been anticipated after multiple presentations. This may have helped improve scores on the sentence portion of the test. Honnell et al. (1991) scored sentence stimuli using key words only. Scoring might not completely account for the difference in the findings between these two studies. For instance, consider that the present study counted only words that were exactly correct (i.e. the response "book" for "books" was incorrect) and Honnell et al. (1991) permitted slight variations on words counted as correct (i.e. "cannot" was correct for "can't" and "coat" was correct for "coats"). This difference in scoring might reduce the difference in performance attributable to counting key words as opposed to counting every word correct in a sentence.

The results of this study revealed no significant difference in lipreading performance between VCV syllables and sentences, which is in direct contrast with findings by Binnie and Shoop (1979). These authors noted that speechreading ability decreased significantly more for sentences than for CV syllables, to which they concluded that age-related effects on speechreading performance were stimuli sensitive. Fatigue was also considered as a possible cause of a decrease in speechreading abilities in older adults (Honnell et al., 1991), however, studies that included test times showed that duration of test sessions was not a factor. Honnell et al. reported a testing time of 30 minutes and Pelson and Prather (1974) reported a single test session time of 45 minutes. Both of these studies suggested age-related effects on speechreading abilities. However, the present study reported a single test session time of approximately 2 hours duration yet failed to show a decline in speechreading performance in older subjects.

In conclusion, it is difficult to explain why some studies reveal age-related effects on lipreading abilities while others fail to show these effects. There is a need for ongoing research in this area as many older adults will acquire hearing loss and will become more dependent on visual speech perception abilities for communication.

## ACKNOWLEDGEMENT

Special thanks to Dr. Mitch Sommers, Ph.D., of the Department of Psychology at Washington University for his assistance in subject recruitment, data collection, and for the use of laboratory facilities and equipment.

## REFERENCES

DiLollo, V., Arnett,J.L., & Kruk, R.V. (1982). Age-related changes in rate of visual information processing. Journal of Experimental Psychology: Human Perception and Performance, 8, 225-237.

Dodd, B. (1987). The acquisition of lipreading skills by normally hearing children. In B. Dodd & R. Campbell (Ed.), Hearing by Eye: The Psychology of Lipreading (pp. 163-175) Hillsdale, N.J.: Erlbaum .

Ewertsen, H.W. & Nielsen, H.B. (1971). A comparative analysis of the audiovisual, auditive and visual perception of speech. Acta Otolaryng, 72, 201-205.

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

Farwell, R.M. (1976, February). Speech reading: A research review. A.A.D., 19-28.

Gagné, J.P. (1994). Visual and audiovisual speech perception training: Basic and applied research needs. Journal of the Academy of Rehabilitative Audiology, 27 (Suppl. 1), 133-159.

Gordon-Salant, S. (1986, December). Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing. Journal of the Acoustical Society of America, 80 (6), 1599-1607.

Gordon-Salant, S. (1987). Age-related differences in speech recognition performance as a function of test format and paradigm. Ear and Hearing, 8 (5), 277-281.

Gordon-Salant,S. (1987). Consonant recognition and confusion patterns among elderly hearing-impaired subjects. Ear and Hearing, 8 (5), 270-176.

Grant, K.W., Walden, B.E., & Seitz, P.F. (May, 1998). Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. Journal of the Acoustical Society of America, 103 (5, Pt. 1) 2677-2689.

Helfer, K.S. (1998). Auditory and auditory-visual recognition of clear and conversational speech by older adults. Journal of the American Academy of Audiology, 9, 234-242.

Honnell, S., Dancer, J., & Gentry, B. (1991, May). Age and speechreading performance in relation to percent correct, eyeblinks, and written responses. Volta Review, 93, 207-213.

Humes, L.E. (1996). Speech understanding in the elderly. Journal of the American Academy of Audiology, 7, 161-167.

Massaro, D.W. & Stork, D.G. (1998, May-June). Speech recognition and sensory integration. American Scientist, 86, 236-244.

Pelson, R.O. & Prather, W.F. (1974). Effects of visual message-related cues, age, and hearing impairment on speechreading performance. Journal of Speech and Hearing Research, 17, 518-525.

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, December). Factors related to lipreading. Journal of Speech and Hearing Research, 2 (4), 340-352.

Tye-Murray, N. (1998). Speechreading. In J.L. Danhauer (Ed.), Foundations of Aural Rehabilitation: Children, Adults, and their Families (pp. 193-213). San Diego, CA.: Singular.

Watson, C.S., Qui, W.W., Chamberlain, M.M., & Li, X. (1996, August). Auditory and visual speech perception: Confirmation of a modality-independent source of individual differences in speech recognition. Journal of the Acoustical Society of America, 100 (2, Pt. 1) 1153-1162.

Wingfield, A. (1996). Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory. Journal of the American Academy of Audiology, 7, 175-182.

Yantis, P.A. (1994). Puretone air-conduction threshold testing. In J. Katz (Ed.), Handbook of Clinical Audiology, Fourth Edition (pp. 97-108). Baltimore, MD: Williams & Wilkins.