

Audio-visual speech perception in mild cognitive impairment and healthy elderly controls

Natalie A. Phillips¹, Shari Baum², Vanessa Taler³

¹ Department of Psychology, Concordia University, Montréal, Canada

² School of Communication Sciences and Disorders, McGill University, Montréal, Canada

³ School of Psychology, University of Ottawa, Ottawa, Canada

natalie.phillips@concordia.ca

Abstract

An audio-visual (AV) speech presentation mode can significantly improve spoken word identification. Language comprehension and communication in patients with Alzheimer disease (AD) can be compromised; however, little is known about the extent to which patients might benefit from an AV mode. Patients with mild cognitive impairment (MCI) are at risk for developing AD and can demonstrate parallel but milder difficulties in aspects of language function. Here we report on preliminary findings of a study that investigates the impact of AV speech and sentence context on word identification in patients with MCI and healthy elderly controls. Although both groups performed better in the AV condition compared to an auditory-alone condition and when a constraining sentence context was present, the patients performed worse than controls overall and in the condition that should afford the greatest benefit. This suggests that the cognitive deficits present in MCI may limit their ability to benefit fully from supportive perceptual and linguistic cues.

Index terms: audio-visual speech perception, sentence context, aging, mild cognitive impairment, dementia

1 Introduction

A person with a decreased ability to understand speech can feel isolated from family, friends, and social activities due to a diminished capacity to communicate [1]. There is abundant evidence that caregivers of patients with Alzheimer disease (AD) feel that communication is problematic throughout the disease and that this negatively impacts the caregiver relationship [2, 3, 4, 5, 6, 7]. Therefore, an understanding of the factors that facilitate speech and language comprehension has important implications for enhancing communication and quality of life of patients with AD and in patients at risk for the disease (i.e., those with mild cognitive impairment or MCI).

It is well-known that speech perception is enhanced when one can both hear and see the speech cues produced by one's communication partner [8, 9], analogous to a 5-18 dB increase in the signal-to-noise ratio [9]. Using auditory and visual (i.e., lip-reading) information to understand speech is known as **audio-visual (AV) speech perception**. It has often been demonstrated that even individuals with normal hearing perform significantly

better on AV speech perception tasks than on **auditory-alone (A)** or visual-alone (V) tasks. This is ecologically advantageous because speech perception typically takes place in less-than-ideal acoustic environments (e.g., due to background noise or multiple simultaneous talkers). The AV benefit is shown by comparing speech perception in different perceptual modalities (A, V, and/or AV) and results at least partially from complementary speech information provided by auditory and visual channels [10], [11]. Understanding this phenomenon has great importance since it has the potential to facilitate speech understanding. However, we are aware of no studies that have evaluated this phenomenon in patients with AD, in whom communication problems are known to exist.

The following literature review is organized as follows. First, we introduce the term "mild cognitive impairment," which refers to a group of individuals considered to be at high risk for developing AD. Second, we review the evidence for the contribution of visual cues to enhancing speech perception and the role that cognitive and linguistic factors play. The implications of age-related and dementia-related changes in these factors are then considered. Finally, we report preliminary data from a line of research that addresses these issues in a comprehensive manner.

Mild Cognitive Impairment (MCI). Alzheimer's disease is the most prevalent degenerative neurological disorder in the elderly and is characterised by marked changes in cognition and personality; however, early diagnosis can be difficult. Recent research has focused on individuals with MCI, that is, individuals with subjective memory complaints who have objective evidence of mild memory loss but generally normal cognitive function and no evidence of a meaningful functional impairment. Other cognitive domains such as language may or may not be involved [12]. This population includes individuals considered at high risk for the development of AD over long-term follow-up [13]. Longitudinal studies of samples of similarly mildly impaired patients have shown that approximately one-half to two-thirds progress to a demonstrably demented state within two to five years [14, 15, 16]. These patients are of great interest because they may represent a transitional state between normal cognitive aging and AD, thus providing a unique opportunity to examine cognitive function very early in the disease course.

AV speech perception. Current models of AV speech perception posit the contribution of both peripheral, bottom-up (i.e., sensory) and central, top-down (i.e., cognitive) mechanisms to one's ability to benefit from the combination of auditory and visual signals [17, 18, 19].

Even simplified models incorporate a complex interplay of perceptual and cognitive-linguistic processes, including memory factors. Thus, breakdowns in speech comprehension can occur at multiple levels of processing. Identifying the source of such communicative impairment is critical to addressing it therapeutically.

In AV speech perception tasks, both younger and older adults seem to display a visual enhancement (VE) effect, but there is no consensus about the magnitude of VE in older individuals. Some studies have demonstrated that both young and older adults exhibit reliable VE effects [20, 21, 22] (although see [23]). Nevertheless, there is considerable variability among studies in the stimuli (syllables, isolated words, sentences) and the participants tested (young versus older adults, presence or absence of hearing loss). The importance of increased linguistic content is unclear, with some studies finding a benefit for more linguistically rich materials [20], some finding no benefit [24], and some even demonstrating a negative effect [23]. It is striking that many authors acknowledge that individual differences in cognitive abilities are likely to be important (e.g., [17, 23, 25]); however, very few studies actually measure them [22]. Our group has shown [22] that cognitive factors can predict AV performance and also that an AV presentation mode can impact up-stream cognition, such as long-term memory.

Importantly for the present purposes, there have been no investigations of AV speech perception in cognitively impaired patients with AD or MCI. Nevertheless, there is a small but growing literature indicating that multi-sensory integration may be impaired in patients with AD. Tippet and Sergio [26] demonstrated impaired integration of visuomotor information in AD patients. Patients with questionable impairment made 35% more errors than did controls, with this percentage jumping to over 250% in patients with mild impairment. Festa et al. [27] tested AD patients, patients with Huntington's disease, and normal controls on two visual integration tasks. One task required integration of visual information within the same neocortical processing pathway (i.e., motion and luminance, both processed within the dorsal visual stream) while the other required integration of visual information from neuroanatomically distinct neocortical regions (i.e., motion and colour, processed in the dorsal and ventral visual streams, respectively). The AD patients were selectively impaired in binding features which required greater cross-cortical integration, despite intact performance during baseline conditions. Extended to our current focus, these studies suggest that AV integration may not be intact in dementia.

To summarize, there is a potentially complex picture of AV facilitation in AD. The AV speech mode offers considerable potential to enhance speech perception and comprehension. However, the ability to fully benefit from AV speech is likely influenced by higher-order cognitive factors [17, 18, 22]. Indeed, work by Small and colleagues [28, 29] has demonstrated that working memory capacity plays an important role in the comprehension of spoken sentences in AD. Finally, there is mounting evidence that integration processes may not be normal in AD. Thus, the

issue of whether AD and MCI patients benefit (and to what extent) from the AV speech mode warrants investigation. We are currently investigating this issue in patients with MCI, patients with AD, and in healthy age- and education-matched controls. Below, we report preliminary findings from MCI participants and controls.

2 Materials and Methods

2.1 Participants

MCI subjects (n=9). Patients were recruited from the McGill/Jewish General Hospital Memory Clinic in Montreal, a clinic staffed by neurologists, geriatricians, psychologists, and nurses. Referrals to the clinic come from the McGill Division of Geriatrics, neurologists, and community physicians. These subjects then participate in a series of clinical investigations including full medical, neuropsychological, and neuroradiological evaluations. The clinical diagnosis of MCI was made on the basis of criteria similar to that of [15] and [30]. Subjects were required to have a reported decline (by the individual or family) in memory function which is gradual, of at least six months duration, and documented by impaired performance (i.e., ≥ 1.5 S.D.) on objective neuropsychological tests with appropriate norms for age and/or education. None had significant impairment in activities of daily living and none met the criteria for dementia. MCI patients had MoCA scores > 26 (see below).

Normal Elderly Controls (NEC; n=9): Healthy elderly controls were recruited from the community. These subjects were administered the neuropsychological battery used in the Memory Clinic to verify that they are cognitively intact and all had passing scores on the MoCA. In the present samples, the NECs are matched on sex to the MCI participants, but not on age or education. These important variables will be matched in the complete participant samples (n=20 each). Currently, the MCI participants are older and have fewer years of education than the NECs. Note, however, that both samples have high levels of education for their age-cohorts (i.e., beyond high school education).

Table 1. Mean (and SD) of age (in years), education (in years) and performance on the Montreal Cognitive Assessment (MoCA; maximum score = 30 pts.).

	<i>Sex</i>	<i>Age</i>	<i>Education</i>	<i>MoCA</i>
NEC (n=9)	2 male; 7 female	69.0 (7.8)	15.9 (2.7)	27.4 (2.3)
MCI (n=9)	6 male; 3 female	80.4 (5.6)	13.3 (2.4)	22.9 (3.3)

Inclusion criteria for all participants consisted of:

- English as the maternal or primary spoken language
- good self-reported health (no cerebro-, or cardiovascular disease or neurological disease)
- normal or corrected-to-normal vision as measured by the:
 - MNREAD acuity chart which consist of easy sentences of standard length presented in 19 decreasing print sizes; this measures reading acuity and reading speed [31, 32];
 - Mars Letter contrast sensitivity test which uses letters of varying size and contrast to measure contrast perception [33, 34];
 - Following the Canadian Ophthalmological Society guidelines, only those who obtain a score of 20/60 or better on the MNREAD acuity test and a score of 1.52 or better (older adults) on the Mars Letter test were tested.
- Normal hearing thresholds, as defined by clinically normal or near-normal pure tone air-conduction thresholds in both ears that are ≥ 25 dBHL from 0.25 to 3.0 kHz, and > 10 dB asymmetry at no more than two frequencies;
- Clinically appropriate cognitive function, as determined by the Montreal Cognitive Assessment (MoCA), a test specifically designed to detect MCI in older adults and which does so with high sensitivity and specificity [35].

2.3 Stimuli and Procedure

Stimuli consisted of sentences from the revised Speech Perception in Noise test (revised; SPIN-R test; [36]). This is a well-standardized test of speech perception that has been extensively used in the literature. It consists of 8 lists of 50 inter-mixed high constraint (e.g., “Stir your coffee with a *spoon*.”) and low constraint (e.g., “Bob didn’t think about the *spoon*.”) sentences. The contrast between the two contextual types allows one to measure the facilitation from linguistic/semantic information provided to a person when listening to speech under adverse conditions [37]. There is good evidence [38, 39, 40, 41] that patients with AD benefit considerably from the presence of context constraint during language processing.

A digital video recording was made of a woman speaking two SPIN-R lists. These were presented in each of the A and AV modalities in counterbalanced order across participants. A-only stimuli were presented free-field with the video monitor turned off, along with a simultaneous auditory masking sound that consisted of 12-talker speech babble signal from the SPIN-R test. This masking signal interferes with speech recognition at the central level. The speech-to-babble ratio was set individually for each participant so that he/she was able to identify approximately 40% of a set of calibration words [42]. The purpose of this manipulation was two-fold: first, it avoids ceiling effects by holding auditory word recognition at a sufficiently low level to measure any VE benefit; second, it equates for perceptual load by making

the auditory recognition of words equally difficult for all participants. The AV condition consisted of the simultaneous V and A presentation of sentences on a high resolution video monitor with the masked auditory speech signal present. A and AV conditions were presented in blocked counterbalanced order. Participants were instructed to listen to the entire sentence to identify the target terminal word and were encouraged to guess when necessary. The dependent variable was the correct identification of the terminal word and only exact phonetic matches were scored as correct.

3 Results

The word identification accuracy was analysed in a Modality (2) X Context (2) X Group (2) ANOVA. Mean identification results are presented in Figure 1 below. There was a trend toward a significant group difference ($F(1,16)=3.15, p = .095$), with MCI participants performing more poorly overall than NCs (53% versus 60%). There was a significant effect of context (i.e., more terminal words were identified from high context sentences than from low context sentences, $F(1,16)=107.4, p < .001$) and a significant effect of modality, such that there was improvement in the AV condition relative to the A condition ($F(1,16)=120.3, p < .001$). There were no significant interactions with group (all F 's < 1.6 , all p 's $> .22$). However, planned comparisons contrasting the two groups at each level of context and modality showed significantly poorer performance of the MCI patients than the controls for high context sentences in the AV modality.

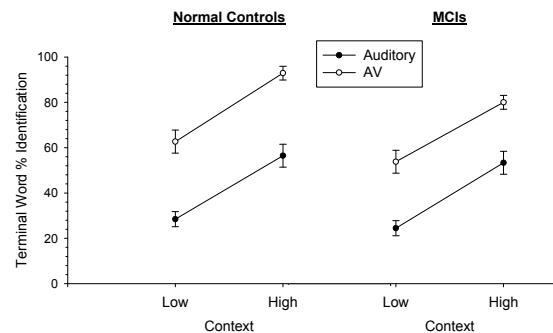


Figure 1. Mean (and SE) of the identification rate of sentence terminal words for Normal Control and MCI participants as a function of Modality (Auditory-alone or AV) and sentence context.

4 Discussion

Our findings to date indicate a number of notable results. First, as expected, healthy older adults showed significant improvement in word identification when there was a supportive sentence context and when sentences were presented in the AV modality compared to the auditory-only modality. These factors did not interact. There was a non-significant trend towards overall poorer word identification in the MCI patients than the normal controls; however, patients showed a significant improvement in performance in the AV modality compared to A alone and with increasing sentence context. Nevertheless, planned comparisons indicated that the MCI patients showed lower levels of performance than the controls in the one condition where benefits should have been greatest, namely in the identification of terminal words from high constraint sentences during the AV modality. This result must be interpreted with caution since the Group X Context X Modality interaction was not significant; however, it suggests that a larger sample size may reveal deficits in the patients' ability to benefit from complementary visual speech cues and supportive context to the same extent as older adults. This is consistent with findings that patients with MCI demonstrate at least mild deficits in lexical-semantic aspects of language function (e.g., [43]). Additional testing of patients with AD will determine whether these abilities decline in dementia. These preliminary data must be interpreted with caution as current the sample of MCI participants are older and have fewer years of education than the NECs.

In the long run, the research outlined here will allow a deeper understanding of the nature and scope of speech processing deficits in AD and MCI, and will indicate possible strategies to improve communication with these individuals. Establishing and maintaining eye contact during conversation is one of the top 10 recommended communication strategies with AD patients [44]. Although not stated explicitly, it is likely that the success of this strategy lies in its ability to engage the attention of the patient and to communicate important AV speech cues. Although caregivers report using this strategy only occasionally, its use is highly correlated with its rated efficacy [44]. The evidence contributed by our study of AV speech perception in AD and MCI could pinpoint where in the communication process AV facilitation takes place and could be translated into improved awareness and training of communication strategies for caregivers.

References

- [1] M. L. Hummert and J. F. Nussbaum, *Aging, Communication, and Health: Linking Research and Practice for Successful Aging*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 2001.
- [2] J. B. Orange, "Perspectives of family members regarding communication change," in *Dementia and Communication*. R. Lubinski, Ed. Philadelphia: B.C. Decker Publishing, 1991, pp. 168-186.
- [3] J. M. Richter, D. Bottenberg, and K. A. Roberto, "Communication between formal caregivers and individuals with Alzheimer's disease," *The American Journal of Alzheimer's Care and Related Disorders & Research*, pp. 20-26, 1993.
- [4] J. M. Richter, K. A. Roberto, and D. J. Bottenberg, "Communicating with persons with Alzheimer's disease: Experiences of family and formal caregivers," *Archives of Psychiatric Nursing*, vol. IX(5), pp. 279-285, 1995.
- [5] J. B. Orange, R. Lubinski, and D. J. Higginbotham, "Conversational repair by individuals with dementia of the Alzheimer's type," *Journal of Speech and Hearing Research*, vol. 39, pp. 881-895, 1996.
- [6] J. B. Orange, K. M. Van Gennep, L. Miller, and A. M. Johnson, "Resolution of communication breakdown in dementia of the Alzheimer's type: A longitudinal study," *Journal of Applied Communication Research*, vol. 26, pp. 120-138, 1998.
- [7] J. A. Small, K. Geldart, and G. Gutman, "Communication between individuals with dementia and their caregivers during activities of daily living," *American Journal of Alzheimer's disease and other Dementias*, vol. 15(5), pp. 291-302, 2000.
- [8] K. W. Grant, B. E. Walden, and P. F. Seitz, "Auditory-visual speech recognition by hearing-impaired subjects: consonant recognition, sentence recognition, and auditory-visual integration," *Journal of Acoustical Society of America*, vol. 103 (5), pp. 2677-2689, 1998.
- [9] W. H. Sumby and I. Pollack, "Visual contributions to speech intelligibility in noise," *Journal of the Acoustical Society of America*, vol. 26, pp. 212-215, 1954.
- [10] N. P. Erber, "Interaction of audition and vision in the recognition of oral speech stimuli," *Journal of Speech and Hearing Research*, vol. 12, pp. 423-425, 1969.
- [11] B. E. Walden, R. A. Prosek, and D. W. Worthington, "Auditory and audio-visual feature transmission in hearing-impaired adults," *Journal of Speech and Hearing Research*, vol. 18, pp. 272-290, 1975.

- [12] R. C. Petersen, R. Doody, A. Kurz, R. C. Mohs, J. C. Morris, P. V. Rabins, K. Ritchie, M. Rossor, L. Thal, and B. Winblad, "Current concepts in mild cognitive impairment," *Archives of Neurology*, vol. 58, pp. 1985-1992, 2001.
- [13] H. Chertkow, "Mild cognitive impairment," *Current Opinion in Neurology*, vol. 15, pp. 401-407, 2002.
- [14] D. W. O'Connor, "Mild dementia: A clinical perspective," in *Dementia and Normal Aging*. F. A. Huppert, C. Brayne, and D. W. O'Connor, Eds. Cambridge: Cambridge University Press, 1994, pp. 91-117.
- [15] R. C. Petersen, G. E. Smith, S. C. Waring, R. J. Ivnik, E. G. Tangalos, and E. Kokmen, "Mild cognitive impairment. Clinical characterization and outcome," *Archives of Neurology*, vol. 56, pp. 303-308, 1999.
- [16] E. H. Rubin, J. C. Morris, E. A. Grant, and T. Vendegna, "Very mild senile dementia of the alzheimer type," *Archives of Neurology*, vol. 46, pp. 379-382, 1989.
- [17] K. W. Grant and P. F. Seitz, "Measures of auditory-visual integration in nonsense syllables and sentences," *Journal of Acoustical Society of America*, vol. 104 (4), pp. 2438-2450, 1998.
- [18] L. D. Braida, "Crossmodal integration, in the identification of consonant segments," *The Quarterly Journal of Experimental Psychology*, vol. 43(A), pp. 647-677, 1991.
- [19] D. W. Massaro, *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, NY: Lawrence Erlbaum, 1987.
- [20] B. E. Walden, D. A. Busacco, and A. A. Montgomery, "Benefit from visual cues in auditory-visual speech recognition by middle-aged and elderly persons," *Journal of Speech and Hearing Research*, vol. 36, pp. 431-436, 1993.
- [21] M. S. Sommers, N. Tye-Murray, and B. Spehar, "Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults," *Ear and Hearing*, vol. 26, no. 3, pp. 263-275, 2005.
- [22] N. A. Phillips, J. P. Gagne, M. Basu, L. Copeland, P. Gosselin, A. Saint-Pierre, and A. H. Winneke, "Audio-visual speech perception in younger and older adults: Effects on word identification and memory," Presented at the 10th Annual meeting of the *International Multisensory Research Forum (IMRF)*, New York City, June 29 - July 2, 2009.
- [23] D. C. Garstecki, "Auditory, visual, and combined auditory-visual speech perception in young and elderly adults," *Journal of the Academy of Rehabilitative Audiology*, vol. 106, pp. 221-233, 1983.
- [24] K. S. Helfer, "Auditory and auditory-visual recognition of clear and conversational speech by older adults," *Journal of the American Academy of Audiology*, vol. 9, pp. 234-242, 1998.
- [25] W. J. Tippet and L. E. Sergio, "Visuomotor integration is impaired in early stage Alzheimer's disease," *Brain Research*, vol. 1102, pp. 92-102, 2006.
- [26] E. K. Festa, R. Z. Insler, W. C. Heindel, D. P. Salmon, J. Paxton, and J. M. Hamilton, "Neocortical disconnectivity disrupts sensory integration in Alzheimer's disease," *Neuropsychology*, vol. 19, no. 6, pp. 728-738, 2005.
- [27] J. A. Small, E. S. Andersen, and D. Kempler, "Effects of working memory capacity on understanding rate-altered speech," *Aging, Neuropsychology, and Cognition*, vol. 4, no. 2, pp. 126-139, June 1997.
- [28] J. A. Small, S. Kemper, and K. Lyons, "Sentence comprehension in Alzheimer's disease: effects of grammatical complexity, speech rate, and repetition," *Psychol Aging*, vol. 12, no. 1, pp. 3-11, Mar. 1997.
- [29] G. E. Smith, R. C. Petersen, J. E. Parisi, R. J. Ivnik, E. Kokmen, E. G. Tangalos, and S. Waring, "Definition, course, and outcome of mild cognitive impairment," *Aging, Neuropsychology, and Cognition*, vol. 3, pp. 141-147, 1996.
- [30] H. Akutsu, G. E. Legge, J. A. Ross, and K. J. Schuebel, "Psychophysics of reading--X. Effects of age-related changes in vision," *Journal of Gerontology*, vol. 46, no. 6, pp. 325-331, 1991.
- [31] A. Subramanian and S. Pardhan, "The repeatability of MNREAD acuity charts and variability at different test distances," *Optometry and Vision Science*, vol. 83, no. 8, pp. 572-576, 2006.
- [32] S. J. Leat and G. C. Woo, "The validity of current clinical tests of contrast sensitivity and their ability to predict reading speed in low vision," *Eye*, vol. 11, pp. 893-899, 1997.
- [33] B. E. Dougherty, R. E. Flom, and M. A. Bullimore, "An evaluation of the Mars Letter contrast sensitivity test," *Optometry and Vision Science*, vol. 82, no. 11, pp. 970-975, 2005.

- [34] Z. S. Nasreddine, N. A. Phillips, V. Bedirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, and H. Chertkow, "The Montreal Cognitive Assessment: A brief screening tool for mild cognitive impairment," *Journal of the American Geriatric Society*, vol. 53, pp. 1-5, 2005.
- [35] R. C. Bilger, M. J. Nuetzel, W. M. Rabinowitz, and C. Rzeczkowski, "Standardization of a test of speech perception in noise," *Journal of Speech and Hearing Research*, vol. 27, pp. 32-48, 1984.
- [36] M. K. Pichora-Fuller, B. Schneider, and M. Daneman, "How young and old adults listen to and remember speech in noise," *Journal of Acoustical Society of America*, vol. 97, pp. 593-608, 1995.
- [37] T. J. Schwartz, K. D. Federmeier, C. Van Petten, D. P. Salmon, and M. Kutas, "Electrophysiological analysis of context effects in Alzheimer's disease," *Neuropsychology*, vol. 17, no. 2, pp. 187-201, 2003.
- [38] R. D. Nebes and C. B. Brady, "The effect of contextual constraint on semantic judgements by Alzheimer patients," *Cortex*, vol. 27, pp. 237-246, 1991.
- [39] R. D. Nebes, F. Boller, and A. Holland, "Use of semantic context by patients with Alzheimer's Disease," *Psychol Aging*, vol. 1, pp. 261-269, 1986.
- [40] M. J. Hamberger, D. Friedman, W. Ritter, and J. Rosen, "Event-related potential and behavioral correlates of semantic processing in Alzheimer's patients and normal controls," *Brain and Language*, vol. 48, pp. 33-68, 1995.
- [41] I. J. Hirsh, H. Davis, R. S. Silverman, S. Reynolds, E. Eldert, and R. Benson, "Development of materials for speech audiometry," *Journal of Speech and Hearing Research*, vol. 17, pp. 321-337, 1952.
- [42] V. Taler and N. A. Phillips, "Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review," *J Clin Exp Neuropsychol*, vol. 30, no. 5, pp. 501-556, 2008.
- [43] J. A. Small and G. Gutman, "Recommended and Reported Use of Communication Strategies in Alzheimer Caregiving," *Alzheimer's Disease and Associated Disorders*, vol. 16(4), pp. 270-278, 2002.