Speech Perception in Individuals with Dementia of the Alzheimer’s Type (DAT)¹

Mitchell S. Sommers

Introduction

Considerable evidence is now available to suggest that dementia of the Alzheimer’s type (DAT) is associated with impairments in speech perception beyond those observed in healthy older adults (Ferraro, 1995; Kempler, 1991; McNamara, Obler, Au, Durso, & Albert, 1992; Sommers, 1998; Strouse, Hall, & Burger, 1995; Tomoeda, Bayles, Boone, Kaszniak, & Slauson, 1990). What remains unclear, however, is the extent to which these DAT-related declines are attributable to age-related hearing loss, DAT-related declines in cognitive abilities, or the interaction between sensory and cognitive impairments. The current study provides one approach to addressing this issue by comparing performance of healthy older adults and individuals with (DAT) on a number of abilities important for understanding speech. Specifically, we compared three abilities necessary for accurate speech perception – understanding speech in noise, accommodating variations in talker characteristics, and discriminating phonetically similar words – in healthy older adults and DAT patients with similar auditory sensitivity. Any differences between these two groups would implicate DAT-related declines in cognitive abilities as the principal contributor to the observed differences. In addition, we compared individuals with and without DAT on the same three measures after selective amplification of signal components to compensate for hearing loss (i.e., spectral shaping). Evidence that individuals with DAT and healthy older adults differ even after compensating for hearing loss would provide additional support for the importance of cognitive declines as the principal factor contributing to impaired speech perception in DAT patients.

As noted, it is also possible that the interaction of sensory and cognitive declines contributes to impaired speech perception in DAT. According to this proposal, the comparable hearing ability of healthy older adults and DAT patients would produce similar degradations to the auditory signal. In the case of individuals with DAT, however, these degraded auditory inputs would be processed by a cognitive system that is compromised relative to that of healthy older adults. Within this framework, the additional declines in speech perception observed in DAT patients would be a consequence of added cognitive impairments preventing them from recovering the intended message from the degraded auditory input.

Clinically, the findings from the current study can serve to specify how best to improve speech perception and spoken communication in individuals with DAT. Evidence that healthy older adults and DAT patients exhibit comparable improvements from spectral shaping would advocate for similar audiological rehabilitation, including hearing aids and other amplification devices, in the two groups. Conversely, evidence for differential improvements following spectral shaping would suggest that amplification alone is unlikely to provide equivalent benefits for cognitively impaired and unimpaired groups of older patients. Such findings would argue, instead, for tailoring rehabilitation strategies to minimize cognitive demands as the most effective way of improving spoken communication in DAT patients.

¹ Presented at the conference on Hearing Care for Adults 2009 – The Challenge of Aging. Chicago, Il.
Methods Common to All Experiments

In this section, I describe methods common to all of the experiments. Stimuli and procedures specific to individual experiments are described in separate sections.

Participants

A total of 145 participants took part in the study and these same participants were tested in all of the experiments. All participants were recruited from the Alzheimer’s Disease Research Center (ADRC) at Washington University. Each participant was screened by the Clinical Core of the Center for depression, severe hypertension, Parkinson’s disease and other disorders that might affect cognitive functions. The severity of dementia was rated according to the Washington University Clinical Dementia Rating (CDR) scale (Hughes, Berg, Danzinger, Coben, & Martin, 1982; Morris, 1993). The CDR is based on a 90-min interview conducted by a board-certified neurologist assessing cognitive function in six areas: memory, orientation, judgment, community affairs, home and hobby, and personal care. CDR levels of 0, 0.5, 1, 2, and 3 represent no, very mild, mild, moderate and severe dementia, respectively. The accuracy and reliability of the CDR and the research diagnostic criteria for DAT have been well established (Berg & Morris, 1994; Burke et al., 1988). All participants were tested within 6–8 weeks of their most recent clinical evaluation to minimize the likelihood of status changing before experimental testing. Only participants from the first three CDR categories (0, 0.5, and 1) were recruited for the present study because prior experience indicated that individuals with more severe dementias would have difficulty understanding and performing the required speech and auditory perception tasks. Table 1 displays demographic characteristics and scores from participants’ most recent neuropsychological assessments in the areas of working memory, processing speed and language. As indicated, age and education did not differ significantly as a function of CDR status. However, with the exception of forward digit span, significant differences were observed in all three cognitive domains (memory, processing speed, and language) between the healthy older adults (CDR 0) and the very mildly impaired DAT patients (CDR 0.5) and between the CDR 0.5 and the mildly impaired participants (CDR 1).

Audiological Assessment

Although previous results (Gates et al., 1995) suggest that DAT does not produce declines in auditory sensitivity beyond those associated with normal aging, we wanted to ensure that this was true for our sample of participants. To examine changes in auditory sensitivity (hearing loss) as a function of CDR status, all participants received an abbreviated audiological examination. All participants had normal middle-ear function (normal

<table>
<thead>
<tr>
<th>Demographic measures</th>
<th>CDR 0 (n = 53)</th>
<th>CDR 0.5 (n = 47)</th>
<th>CDR 1 (n = 45)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>78.2</td>
<td>75.9</td>
<td>74.3</td>
</tr>
<tr>
<td>Education</td>
<td>13.9</td>
<td>13.3</td>
<td>14.1</td>
</tr>
<tr>
<td>Memory measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit span forward</td>
<td>6.4</td>
<td>6.4</td>
<td>5.8</td>
</tr>
<tr>
<td>Digit span backward</td>
<td>4.8</td>
<td>4.4*</td>
<td>3.5**</td>
</tr>
<tr>
<td>Paired associates</td>
<td>14.4</td>
<td>9.4*</td>
<td>7.1**</td>
</tr>
<tr>
<td>Processing speed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit-symbol</td>
<td>40.1</td>
<td>36.8*</td>
<td>23.8**</td>
</tr>
<tr>
<td>Language measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS vocabulary</td>
<td>53.4</td>
<td>43.7*</td>
<td>35.5**</td>
</tr>
<tr>
<td>Boston naming</td>
<td>53.9</td>
<td>44.6Z*</td>
<td>35.0**</td>
</tr>
</tbody>
</table>

*Significant difference between CDR 0 and CDR 0.5  
**Significant difference between CDR 0.5 and CDR 1

Table 1. Demographic and Neuropsychological Data for Study Participants
tympanograms and presence of acoustic reflex for contralateral stimulation with a 1000-Hz pure tone at 100 dB HL. Pure-tone air-conduction thresholds for octave frequencies from 250 to 4000 Hz were obtained from both ears and mean hearing loss for the better ears (American National Standards Institute [ANSI], 1989) in each participant group are displayed in Figure 1. All three groups exhibited a sloping high-frequency hearing loss that is characteristic of older adults (Committee on Hearing and Bioacoustics [CHABA] Working Group on Speech Understanding and Aging, 1988). Comparisons across the three groups at each of the frequencies revealed no significant differences in hearing loss as a function of CDR status. Thus, mean absolute sensitivity was approximately equivalent for all three CDR ratings.

Spectral Shaping

In all three experiments participants were tested using both unmodified and spectrally shaped speech signals. Spectral shaping was intended to amplify portions of the speech signal such that all components in the speech sounds up to 4000 Hz were obtained from both ears and mean hearing loss for the better ears (American National Standards Institute [ANSI], 1989) in each participant group are displayed in Figure 1. All three groups exhibited a sloping high-frequency hearing loss that is characteristic of older adults (Committee on Hearing and Bioacoustics [CHABA] Working Group on Speech Understanding and Aging, 1988). Comparisons across the three groups at each of the frequencies revealed no significant differences in hearing loss as a function of CDR status. Thus, mean absolute sensitivity was approximately equivalent for all three CDR ratings.

Experiment 1 – Speech Reception Thresholds

In Experiment 1, we measured speech reception thresholds (SRTs) – defined as the signal-to-noise ratio (SNR) yielding approximately 50% correct word identification – to investigate whether DAT produced deficits in the ability to identify words presented in background noise. One of the consistent findings from cognitive assessments in DAT patients is that relative to healthy older adults, they have a decreased ability to inhibit irrelevant information (Duchek, Balota, & Thessing, 1998; Spieler, Balota, & Faust, 1996; Sullivan, Faust, & Balota, 1995). Moreover, deficits in inhibitory control generally increase as a function of dementia severity (Sommers, 1998). Based on these earlier findings, we expected to find systematically higher SRTs as a function of dementia severity.

Method

Stimuli

Stimuli were 100 monosyllabic words taken from the revised Speech Perception in noise test (Bilger, Nuetzel, Rabinowitz, & Rzeczkowski, 1984). The 100 words were divided into two lists of 50 items each. Lexical characteristics including word frequency and word familiarity were matched as closely as possible across the two lists.

Design and Procedure

On each trial participants viewed the word “ready” presented on a CRT and pressed the space bar once they were ready to listen to the next item. Words were presented in a background of 6-talker babble. Presentation of the unmodified and spectrally shaped stimuli were blocked, with half the participants in each CDR group receiving the spectrally-shaped stimuli first and the other half receiving the unmodified stimuli first. The SNR for the first presentation in each block was set to a level considerably below the estimated SRT. This item was repeated at gradually increasing SNRs until it was identified correctly. Step size for the initial item was 4 dB (i.e., SNR improved by 4 dB for each presentation of the item until it was identified correctly). Once this initial item was correctly identified, subsequent words were presented using an adaptive one-down, one-up procedure with a 2-dB step size.
This procedure tracks the 50% correct identification point and final SRT values were taken as the mean of the final 25 reversals.

Results

Figure 2 displays mean SRTs obtained for the shaped and unshaped stimuli as a function of CDR status. Unless otherwise noted, all results in this and the remaining experiments reported as significant were reliable at the p < .001 level. As expected, all three groups had significantly lower SRTs in the shaped versus the unshaped condition. In addition, performance differed significantly as a function of CDR status, with CDR 0 participants exhibiting significantly lower SRTs than CDR 0.5 individuals and CDR 1 participants exhibiting higher SRTs than the CDR 0.5 participants. Of particular interest to the current study is that the effects of spectral shaping interacted significantly with CDR status such that CDR 0 individuals showed significantly greater differences between shaped and unshaped stimuli than did either of the CDR groups. Differences between the two CDR groups were also significant. Thus, despite similar auditory sensitivity, CDR status significantly affected individuals’ ability to benefit from spectral shaping when listening to speech in noise.

Experiment 2 – Lexical Discrimination

One ability that is necessary for recognizing spoken words is lexical discrimination. In the present study, lexical discrimination refers to the capacity to match representations derived from incoming speech signals to one of the tens of thousands of items stored in long-term lexical memory (the mental lexicon). Although a number of models have been proposed to account for this process (Luce & Pisoni, 1998; Marslen-Wilson, 1990; McClelland & Elman, 1986), the present study will focus on one, the Neighborhood Activation Model (NAM).

The NAM (Luce & Pisoni, 1998) proposes that words in the mental lexicon are organized into similarity neighborhoods. A similarity neighborhood, according to the model, consists of items that are phonetically similar to a given target word. Operationally, neighborhoods are defined as all words that can be created from a target item by adding, deleting or substituting a single phoneme. Thus, the neighborhood of the word “CAT”, for instance, would contain items such as “HAT”, “BAT”, “COT” and “CAB” (as well as all other items that can be created using the addition, deletion, and substitution rules). The focus of the present study was on one key structural characteristic of similarity neighborhoods, neighborhood density, that has been demonstrated to influence the speed and accuracy of spoken word recognition (Luce & Pisoni, 1998; Sommers, 1996). Neighborhood density refers to the number of phonetically similar words (neighbors) of a specified target item. Words with many neighbors are considered to reside in “dense” neighborhoods, whereas those with relatively few neighbors are said to be located in “sparse” neighborhoods.

According to the NAM, spoken word recognition requires activation levels on a target (the stimulus) to be increased and activation levels on similar sounding competitor items (neighbors) to be reduced. This proposal predicts that certain words (easy words) should be identified more accurately than others (hard words) because of differences in their neighborhood characteristics. Specifically, the NAM predicts that target words from high-density neighborhoods should be identified less accurately (i.e., are hard words) than items from low-density neighborhoods (easy words) because hard words require listeners to reduce activation levels on a greater number of neighbors.

As noted, considerable evidence is available to suggest that DAT impairs the ability to inhibit irrelevant information. In the context of the current experiment, therefore, we predicted that the ability to reduce activation on competing neighbors would decline as a function of CDR status, leading to bigger differences between easy and hard words as the severity of dementia increased. Furthermore, to the extent
that lexical discrimination improves as a result of improved audibility, we would expect the effects of CDR status on differences between easy and hard words to be reduced for spectrally shaped compared with unmodified stimuli.

**Method**

**Stimuli**

Stimuli were monosyllabic words selected using the web-based neighborhood search database (neighborhoodsearch.wustl.edu). Seventy-six hard words and seventy-six easy words were selected from the database. Hard words had a mean neighborhood density of approximately 26 (i.e., 26 other words could be created from the target item by adding, deleting or substituting a single phoneme) and easy words had a mean neighborhood density of approximately 10.

**Design and Procedure**

On each trial participants saw the word “ready” appear on the CRT and pressed the space bar when they were ready to hear the next item. All items were presented in a background babble at an SNR ratio of approximately +2. Easy and hard words were presented randomly within a block, but spectral shaping was manipulated across blocks. Within each CDR group approximately half the participants heard spectrally shaped items first and the other half heard unmodified version of the stimuli first. Individual items were also counterbalanced such that across participants in a given CDR status each word was presented spectrally shaped and unmodified to approximately an equal number of participants.

**Results**

As in Experiment 1, all results reported as significant were reliable to the p < .001 level, unless otherwise noted. The top panel in Figure 3 displays percent correct identification for lexically easy and hard words as a function of CDR status. Performance for easy words did not differ across the three groups and this was true for both shaped and unshaped stimuli. These findings are in good agreement with previous results (Sommers, 1996) that impaired inhibitory demands have little or no effect on identification of easy words because they have relatively few competitors that need to be inhibited. For the hard words, no differences were observed between the CDR 0 and CDR 0.5 participants, but the more severely impaired CDR 1 individuals exhibited significant declines for both shaped and unshaped stimuli.

The bottom panel in Figure 3 shows differences in performance between the shaped and unshaped stimuli. For lexically easy words spectral shaping improved performance approximately equally for all three CDR groups. In contrast, the effects of spectral shaping varied significantly for lexically hard words, with the benefits of spectral shaping decreasing as a function of dementia severity. CDR 0 participants showed significantly greater benefits from spectral shaping than did the 0.5 individuals and this group exhibited significantly more benefit from shaping than did the CDR 1 group.

**Experiment 3 – Talker Normalization**

Talker normalization has been proposed (Johnson, 1990; Nearey, 1989) as one of the principal mechanisms that listeners use to maintain perceptual constancy during speech perception. This mechanism is critical for accurate recognition of spoken words because a number of factors, including changes in talkers, speaking rate, and stress pattern (Johnson, 1990) can alter the acoustic properties of speech sounds. For example, productions of the same word by a man, a woman, and a child will have distinct physical characteristics due to vocal-tract differences between the three talkers. In most environ-
ments, however, listeners have little difficulty identifying these different signals as instances of the same word (produced by different talkers) despite the dramatic variations in acoustic properties.

Traditionally, theories of speech perception have accounted for listeners' ability to maintain perceptual constancy, despite the high degree of acoustic variability in speech signals, by positing a stage of processing referred to as talker normalization. According to this account, normalization is an early stage in the speech perception system during which acoustic differences resulting from changes in talker characteristics are re-scaled or normalized to standardized forms (Johnson, 1990). Sommers (1997) reported that talker normalization abilities were significantly poorer in healthy older adults than in young listeners matched for overall hearing levels. One explanation that was proposed for this age-related impairment in talker normalization was reduced processing speed. Salthouse (1994) suggested that reduced processing speed can affect performance on a variety of perceptual tasks because stimulus information degrades over time. This degradation can cause representations derived from incoming stimuli to become less stable and more susceptible to disruption as general slowing increases. Sommers (1997) proposed that general slowing might contribute to age-related declines in spoken word recognition because delays in completing talker normalization could result in older listeners relying on increasingly degraded representations to make phonetic decisions. Recent evidence indicates that processing speed is significantly slower in DAT than in healthy older adults (Myerson, Lawrence, Hale, Jenkins, & Chen, 1998). Furthermore, Marshall et al. (1996) reported that DAT listeners were slowed, relative to non-demented older adults in completing several early stages of spoken word recognition. One goal of the current investigations, therefore, was to assess perceptual normalization in DAT listeners and to establish whether any observed impairment varied as a function of dementia severity. In addition, Experiment 3 examined the extent to which changes in talker normalization as a function of DAT status could be overcome by improving audibility through spectral shaping. On the basis of findings from previous investigations (Marshall et al., 1996; Myerson et al, 1998), the working hypothesis was that DAT listeners would exhibit greater impairments in perceptual normalization than age-matched healthy controls.

**Method**

**Stimuli**

The stimuli for Experiment 3 were 150 monosyllabic words recorded by six different talkers (3 male and 3 female). Stimuli were divided into two lists equated for word frequency and neighborhood density.

**Design and Procedure**

On each trial participants saw the word “ready” appear on a CRT screen and pressed the space bar when they were ready to have the next word presented. Stimuli were presented in a six-talker background babble at an SNR of approximately +2. Participants heard one block of 75 items all spoken by one talker and another block of 75 words with items spoken by all six talkers. The specific talker used in the single talker condition was counterbalanced such that within each CDR group approximately equal numbers of listeners heard each of the six talkers as the single talker. Half of the individuals in each group heard the multiple talker condition first and half heard the single talker condition first. Finally, within each block half the words were unmodified and half were spectrally shaped as described in the general method section.

**Results**

The top panel of Figure 4 displays percent correct performance for the three CDR groups in both the single- and multiple-talker conditions. As expected based on previous studies comparing single- and multiple-talker conditions, performance for all three groups of participants was significantly better when items were produced by the same talker than when there was trial-to-trial variations in talker. Also as expected performance in the spectrally shaped conditions was better than when listeners heard unmodified versions of the stimuli. Critically, however, the three-way interaction of spectral shaping x CDR status x single vs. multiple was significant, suggesting that the benefits of spectral shaping in the single and multiple talker conditions varied based on CDR status. Simple effects analyses indicated that for the single-talker conditions, spectral shaping had comparable benefits across the three groups of CDR participants. This result can be seen on the left side of the bottom panel of Figure 4, which displays differences between performance in the shaped and un-
shaped conditions. In contrast to the similar benefits of spectral shaping in the single-talker condition, the benefits of spectral shaping for multiple-talker presentations varied significantly as a function of CDR status. Post-hoc analyses indicated that CDR 0 individuals were the only ones who exhibited significant benefits from spectral shaping in the multiple talker condition.

**General Discussion**

Taken together the findings from the current set of experiments suggest that the cognitive impairments associated with DAT affect several of the abilities needed for successful speech perception, including understanding speech in noise, discriminating lexical competitors, and normalizing for talker differences. Consistent with previous findings (Gates et al., 1995) healthy older adults and individuals with DAT had similar auditory thresholds, suggesting that the observed impairments were largely a consequence of cognitive declines (or the interaction of cognitive and sensory impairments) rather than any additional cochlear pathology attributable to DAT. The effects of spectral shaping also point to a primarily cognitive, rather than sensory basis for the difficulties that DAT patients have in more natural listening situations, such as those with difficult lexical discriminations and multiple talkers. Specifically, under less demanding listening conditions, such as single talkers and lexically easy words all three groups demonstrated comparable benefits of spectral shaping. Under more difficult listening conditions, such as multiple talkers or lexically difficult words, DAT patients were less able than their healthy older counterparts to benefit from spectral shaping.

In addition to providing important new theoretical information about the nature of the deficits underlying spoken communication in DAT, the findings from the current study have several important clinical implications. One finding that was particularly encouraging is that although individuals with DAT did not benefit as much from spectral shaping as healthy controls, they nevertheless benefited from selective amplification of the speech signal. This result suggests that hearing aids and other forms of amplification are likely to be reasonably effective in this population.

A second important clinical implication of the current findings is that they provide guidance on environmental modifications that are likely to improve speech perception in individuals with DAT. First, the SRT results suggest that DAT increases susceptibility to masking noise, making it essential to minimize background interference during spoken communication with this population. Second, the findings from the study on talker normalization indicate that when DAT individuals have an opportunity to adapt to a single speaker they perform as well as healthy older adults in understanding speech and that they also benefit as much from selective high-frequency amplification. Thus, noisy multiple talker environments, such as those often encountered in hospitals or nursing homes, are likely to be particularly problematic for individuals with DAT. Moreover, without additional modifications to the environment, such as having as much information as possible conveyed by a single familiar talker, amplification is unlikely to produce significant improvement in speech understanding for individuals with DAT.

One additional implication of the current findings is that measures of speech perception abilities might be usefully incorporated into diagnostic assessments for DAT. All three measures of speech perception, resistance to noise, lexical discrimination, and talker normalization were significantly impaired in DAT patients relative to healthy older controls. If these findings are confirmed by additional research, then these measures may provide an additional tool to distinguish healthy aging from DAT. Of particular importance, annual assessments of these abilities would allow clinicians to identify significant longitudinal changes and provide an early indicator of possible cognitive declines.

![Figure 4](image-url)
Acknowledgments

This research was supported by the Brookdale Foundation and the National Institute of Aging. The author would like to thank the Washington University Alzheimer’s Disease Research Center and Dr. John Morris for their invaluable help in recruiting and assessing our Alzheimer’s participants.

References


