Washington University in St. Louis
Washington University Open Scholarship

All Theses and Dissertations (ETDs)

Winter 12-1-2013

The Effect of Age-related Declines in Inhibitory Control on Audiovisual Speech Intelligibility

Avanti Dey
Washington University in St. Louis

Recommended Citation

Dey, Avanti, "The Effect of Age-related Declines in Inhibitory Control on Audiovisual Speech Intelligibility" (2013). *All Theses and Dissertations (ETDs)*. 1206.
https://openscholarship.wustl.edu/etd/1206

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Psychology

The Effect of Age-related Declines in Inhibitory Control on Audiovisual Speech Intelligibility

by

Avanti Dey

A thesis presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the
degree of Master of Arts

December 2013

St. Louis, Missouri

TABLE OF CONTENTS

List of Figures and Tables

FIGURES

TABLES

ACKNOWLEDGEMENTS

Much appreciation and thanks are due to Mitch Sommers for his invaluable assistance on the development of this project, as well as revisions of this manuscript. Thanks are also due to Dave Balota and the Cognitive Psychology Lab, and Sandy Hale for their immensely helpful suggestions during various stages of this project. Finally, Karl Wassman and Niki Runge were involved in most of the participant recruitment and data collection, and I owe them many thanks as well for their commitment.

ABSTRACT OF THE THESIS

The Effect of Age-related Declines in Inhibitory Control on Audiovisual Speech Intelligibility

by

Avanti Dey

Master of Arts in Psychology

Washington University in St. Louis, 2013

Professor Mitchell S. Sommers, Chair

Audiovisual (AV) speech perception is perception in which both auditory and visual information is available in order to understand a talker, compared to an auditory signal alone, during face-to-face communication. This form of communication yields significantly higher word recognition performance as compared to either sensory modality alone, constituting a general AV advantage for speech perception. Despite an overall AV advantage, older adults seem to receive less benefit from this bimodal presentation than do younger adults. However, there is evidence to suggest that not all age-related deficits in AV speech perception are of a sensory nature, but are also influenced by cognitive factors (e.g. Pichora-Fuller et al., 1995). In the current study, I extend an existing model of spoken-word recognition to the AV domain and refer to the new model as the Auditory-Visual Neighborhood Activation Model (AV-NAM). The primary goal of the current study was to examine the cognitive factors that contribute to age-related and individual differences in AV perception of words varying in lexical density (i.e. easy and hard words). Forty-nine younger and 50 older adults completed a series of cognitive inhibition tasks and several spoken word identification tasks. The words were presented in auditory-only, visual-only, and AV conditions. Overall, younger adults demonstrated better inhibitory abilities and higher word identification performance than older adults. However, whereas no relationship was observed between inhibitory measures and word identification performance in younger adults, there was a significant relationship between inhibition, as measured by Stroop interference, and intelligibility of lexically difficult words in older adults. These results are interpreted within the framework of the newly adapted AV-NAM and the implications for inhibitory deficits in older adults that contribute to impairments in speech perception.

Introduction

I. Sensory Contributions to Audiovisual Speech Perception

During face-to-face speech communication, listeners use information from the speech signal available in both the auditory and visual modalities in order to facilitate perception. This combined sensory signal is known as audiovisual (AV) speech perception, and successful AV perception involves successfully extracting and integrating auditory (A) and visual (V) information present in the speech signal. Numerous studies have shown a benefit for AV speech compared with A or V alone (e.g., Green, 1998; Summerfield, 1987).

One particularly striking demonstration of the visual contribution to speech perception is the *McGurk effect* (McGurk & MacDonald, 1976). This phenomenon occurs when an auditory signal specifying a particular phoneme (e.g., /ba/) is dubbed with a visual signal specifying a different phoneme, usually at a different place of articulation (e.g., /ga/). Most often, listeners will report hearing a different phoneme (typically /da/ or /tha/, for an auditory /ba/ paired with a visual /ga/) when the visual signal is added. For example, superimposing the voiced utterance /ba-ba/ over lip movements for /ga-ga/ results in subjects reporting the utterance /da-da/ (McGurk & MacDonald, 1976). Numerous studies (e.g. Green, Kuhl, Meltzoff, & Stevens, 1991; Green & Norrix, 1997; Rosenblum, Yakel, & Green, 2000) have since replicated this effect, validating the claim that information from the acoustic signal is combined with information from the visual signal and can significantly affect speech perception.

While the McGurk effect focuses on AV perception of phonemes, studies have also examined the effect of an additional visual signal to more complex acoustic speech stimuli. In a study by Behne et al. (2007) which compared syllable identification under audio-only (A), visual-only (V), and AV conditions between younger adults (19-30 years) and mid-aged adults

(49-60 years), the investigators observed that older adults made a higher proportion of correct responses in V-only presentation, as compared to younger adults. From these findings, the authors suggested that with age comes an increasing use and reliance on the visual signal for additional sensory information. Such results suggest that the use of visual cues in AV speech perception may be a function of the experience accompanying advancing age, such that the effects of age-related hearing loss, as well as greater linguistic knowledge (Schneider & Pichora-Fuller, 2000; Wingfield, Alexander, & Cavigelli, 1994) which may produce a greater benefit from the visual signal in older adults. Most models of AV speech perception incorporate at least three independent contributors to AV speech performance – 1) the ability to encode auditory information; 2); the ability to encode visual information (i.e. lipread) and 3) the ability to integrate information obtained from the two modalities (A and V).

These factors of unimodal encoding and sensory integration have been particularly important in the investigation of age-related differences in AV speech perception, particularly in how older individuals are able to recognise speech. Advancing age is typically accompanied by some degree of hearing loss in the form of cochlear hair cell degeneration (Hull, 1995; Liu & Yan, 2007), as well as vision loss in the form of lens stiffening and macular degeneration (Liu, White, & LaCroix, 1989; Kasper, 1978). In particular, studies consistently demonstrate that age-related sensorineural hearing loss (*presbycusis*) is the primary contributor to impaired speech perception (e.g. Halling & Humes, 2000, Humes et al., 1994). Reductions in sound sensitivity, particularly in the higher frequencies, as well as other changes in the auditory periphery can all affect older listeners' ability to perceive speech (Morrell, Gordon-Salant, Pearson, Brant, & Fozard, 1996; Pichora-Fuller, 2003). Visual speech cues, such as lip-reading have also been found to be impaired in older adults (Honnell, Dancer, & Gentry, 1991; Sommers, Tye-Murray,

& Spehar, 2005), such that older individuals are less adept at recognising speech in a V condition. Taking these findings together, it is clear that sensory loss in older adults negatively affects unimodal encoding. However, this deficit can be overcome by an alternative means of obtaining additional information to the speech signal, available in the form of combined audiovisual (AV) speech.

The benefit of bimodal (i.e. AV) presentation of speech perception has been studied across a range of age groups. In one such study, Walden, Busacco, and Montgomery (1993) examined the differences between middle-aged (35 to 50 years of age) and older (65 to 80 years of age) adults with comparable hearing on recognition of speech stimuli in A, V, and AV presentation. For both groups, AV percent correct was significantly higher than that of A-only performance, which in turn exceeded V-only performance. However, due to large differences between the two groups in V-only performance, interpretable findings independent of V-only performance are difficult. As such, the question of focus has shifted to how AV perception is differentially affected by age while minimising the influence of unimodal perception. Sommers et al. (2005) tested younger and older adults on consonant, word, and sentence identification in all three sensory modalities (A, V, and AV), while additionally equating the age groups for 50% audibility in the A condition (ASHA, 1978). The authors observed a similar pattern of results across the three conditions as did Walden et al, such that performance for both younger and older adults was highest in the bimodal AV condition, followed by the unimodal conditions. Other studies (e.g. Hay-McCutcheon, Pisoni, & Kirk, 2005; Tye-Murray, Sommers, & Spehar, 2007) have also observed a similar pattern of AV performance as a function of sensory modality, in which individuals perform significantly better under AV conditions, as opposed to unimodal conditions.

One reason for this improved performance in AV, compared with A or V conditions, is the complementary nature of AV speech perception (e.g. Blamey, Cowan, Alcantara, Whitford, & Clark, 1989; Grant & Seitz, 1998), in the sense the two sensory signals provide complementary information about speech. This advantage derives from the fact that when the information conveyed by one sensory signal is masked (i.e. through perceptual degradation), the introduction of the alternate signal provides additional information (Grant, Walden, & Seitz, 1998). Thus, in order to obtain the maximum amount of information in noisy conditions, it is beneficial to take advantage of the two speech signals together. For example, voicing and phrasing cues are well conveyed by the auditory signal, but not by the visual signal. Conversely, cues conveying information about place of articulation are better conveyed by the visual signal than by the acoustic signal, particularly in noise. Moreover, the AV advantage occurs even when both the A and V signals are fully available (Massaro & Light, 2004). This complementarity of available information from two speech signals in perception becomes especially true in the presence of noise, wherein perception of the A-only signal is more prone to effects of noise (Sumby & Pollack, 1954; Erber, 1969). That is to say, the AV advantage is greatest under adverse listening conditions in which the auditory signal is masked (Macleod & Summerfield, 1987) in both younger and older adults (Girin, Schwartz & Feng, 2001).

The visual contribution to speech perception has been formalised as the term *visual enhancement* (VE) from an idea presented by Sumby and Pollack (1954). VE is a measure of the improvement obtained in an AV condition compared with an A-only condition, normalised to A-only performance Specifically, VE refers to the degree to which perception of an auditory speech stimulus benefits from adding an additional visual signal. VE is calculated as follows:

$$VE = (AV-A)/(1-A)$$

This formula directly compares an individual’s A and AV performance, such that differences in baseline A-only performance are normalised in order to calculate the overall improvement (relative to the maximum possible improvement) as a result of additional visual information. This measure has been used in a number of studies investigating AV speech perception (Rabinowitz, Eddington, Delhorne, & Cuneo, 1992; Grant et al., 1998; Sommers, 2005) in order to observe the obtained benefit of additional visual information. Sommers et al. (2005) specifically investigated age-related differences in AV perception with an emphasis on examining the effects of VE. The authors observed that while older adults appeared to be benefitting slightly less than younger adults from an additional visual signal (an average proportion of 0.63 in younger adults, compared to an average proportion of 0.57 in older adults), this difference did not reach significance. Moreover, while A-only performance was equivalent for both groups, older adults achieved significantly lower identification accuracy than younger adults in V-only performance. Thus, despite significant impairments in V performance, older adults were still able to benefit from an additional visual signal at a comparable level to younger adults. In a later study, Tye-Murray et al. (2007) assessed VE in normal-hearing and hearing-impaired older adults, and found that after accounting for differences in unimodal A and V performance, the normal-hearing and hearing-impaired groups also experienced comparable levels of visual enhancement. Together, these findings suggest that older adults have an intact ability to integrate multiple sources of sensory information into single bimodal percept; that is, despite overall lowered levels of AV performance, they are still receiving the same amount of measurable benefit from the additional visual signal.

II. Cognitive Contributions to Spoken-Word Recognition

Overall, evidence from the literature demonstrates that speech perception is facilitated by information obtained simultaneously from auditory and visual speech signals, and that this form of bimodal perception is superior to speech perception presented within either modality in isolation. Older adults also benefit from AV speech such that recognition is improved in comparison to A-only and V-only speech, although not to the same degree as younger adults. In addition to age-related sensory impairments being the cause of this poorer recognition performance (Pichora-Fuller & Souza, 2003), there also appears to be a significant cognitive component as well. Studies have demonstrated that age-related cognitive declines in processing speed (van Rooij & Plomp, 1990) and working memory capacity (Pichora-Fuller, Schneider, & Daneman, 1995) also contribute to auditory speech perception ability. Moreover, it has been shown that these cognitive deficits exist independently of sensory decline; Sommers (1996) found evidence to support age-related deficits in lexical retrieval that persisted despite controlling for hearing loss. It is these cognitive changes that will be the focus of the following sections, with a particular focus on the access and retrieval of lexical representations during auditory spoken word recognition.

a. The Neighborhood Activation Model (NAM).

In the current study we adapt a well-investigated model of auditory spoken-word recognition to the case of AV presentation. According to such models, auditory word recognition occurs when an incoming speech stimulus is compared to existing representations in the mental lexicon to establish a best-match between the target and items in the lexicon. Luce and Pisoni (1998) posited that the underlying structural relations of words in the lexicon are also important in the process of spoken word recognition in their Neighborhood Activation Model (NAM) of spoken-word recognition. According to the NAM, lexical items are arranged based on their phonetic similarity to a particular target word, such that phonetically similar items are grouped close to the target, while phonetically dissimilar items are situated further away from the target. Operationally this arrangement of lexical ‘neighbourhoods’ is derived from words that can be formed from the target word through deletion, addition, or substitution of a single phoneme (Greenberg & Jenkins, 1964). For instance, the words CAN, SCAT, and AT are all lexical neighbours of the word CAT. Subsequently, the term *neighbourhood density* (ND) refers to the number of phonetically similar items to a target word; accordingly, *high*-density neighbourhoods are comprised of many phonetically similar items to a target word, while *low*-density neighbourhoods are comprised of a small number of phonetically similar items to a target word. In addition to structural relations, word recognition can also be affected by frequency relations, such that the ease with which spoken words are recognised is directly related to the objective frequency of these words in experienced language.

Although the NAM considers a combination of these factors together as playing a role in word recognition, the effects of ND and frequency can exist independently of one another. Indeed, Sommers (1996) found that ND effects exist after equating frequency across lexically easy and hard words; it is the ND parameter that will be of primary interest here. In particular, the NAM hypothesises that this neighbourhood-based structural organisation of lexical representations has consequences for spoken-word recognition, such that a certain word – depending upon its neighbourhood size – will be either easy or difficult to recognise. Specifically, high ND words – which intrinsically comprise more neighbours –, ought to be more difficult to perceive in comparison to words with low ND words (i.e. words with fewer confusable items). In addition to observing classic word frequency effects, Luce and Pisoni (1998) observed lexical density effects in speech intelligibility. Specifically, the authors observed that response accuracy of high ND words was worse in comparison to low ND words, findings indeed consistent with the NAM. The word frequency and ND parameters were then combined with neighbourhood frequency to designate words based on lexical difficulty. The authors defined lexically *easy* words as a high frequency word in a low density neighbourhood whose neighbourhood frequency was also low. In contrast, a lexically *hard* word was a low frequency word in a high density, high frequency neighbourhood. Through a series of experiments testing frequency and density effects on perceptual identification in noise, auditory lexical decision time, and word naming time, they observed that listeners indeed recognised lexically easy words more quickly and accurately than lexically hard words. Collectively, the findings from the experiments support the hypothesis that a neighbourhood organisational structure significantly affects the discriminability and facilitation of access and retrieval during spoken-word recognition. The active dynamics of this structure will be discussed in the next section.

b. The Role of Inhibition in the NAM.

The underlying mechanisms of the NAM entail a dynamic interaction between activation and competition among word candidates within the lexicon as the basis for spoken-word recognition. This activation-competition framework is not necessarily unique to the NAM (e.g. Lahiri & Marslen-Wilson, 1991; Norris, 1994), but represents a range of speech perception models that are characterised by the relationship between the target speech stimulus and competing candidates. According to these models, the process of auditory word recognition proceeds as follows: 1) perceiving a word activates the corresponding lexical representation in the lexicon; 2) perceiving the word involves parallel activation of acoustically similar representations which results in competition amongst the representations; 3) in order to isolate and identify the target, the activated competitors must be suppressed; this inhibitory mechanism subsequently leads to lowered activation levels of the competitors and elevated activation of the target. It is also important to note that competition between candidates continues until the level of one candidate exceeds the average within the entire lexicon, thus emphasising the need for continuing inhibition. Moreover, the level of inhibition is determined directly by the level of activation, such that more strongly activated candidates will exert greater inhibition. This inhibition of competitors typically occurs in a graded fashion based on the phonetic similarity of the target word to neighbouring candidates, such that more similar items initially receive higher levels of activation. These activation levels are monitored by *word decision* units that correspond to the specific acoustic-phonetic pattern of a candidate item. As the speech stimulus continues to be received and processed, levels of activation are reduced for those candidates that are no longer consistent with the target, while those items that display continued levels of acoustic similarity exhibit increased levels of activation, accruing in the word decision unit. The final stage of successful word recognition occurs when all the irrelevant items have been eliminated through active inhibition, and levels of activation converge upon the target word, and its word decision unit exceeds some criterion threshold. That is, word recognition is achieved once sufficient perceptual evidence has become available that matches an existing lexical representation.

This mechanism of inhibition in spoken-word word recognition has been demonstrated in several paradigms including word-to-picture matching (e.g., Magnuson, Dixon, Tanenhaus, & Aslin, 2007), gating (e.g., Garlock, Walley, & Metsala, 2001), and priming (Dufour & Peereman, 2003; Goldinger, Luce, & Pisoni, 1989). Together, these findings demonstrate that the effect of competing neighbours act to inhibit word recognition, consistent with the original findings of Luce and Pisoni (1998). Moreover, these effects have been computationally modelled (Chen & Mirman, 2012), substantiating the claim that coactivated representations produce a net inhibitory effect upon the target word if they are strongly activated, and also consistent with modifications to the NAM proposed by Sommers (1996).

Given this inhibitory mechanism which appears to play an important role in lexical access during auditory word recognition, it is possible to make certain predictions for accuracy in word recognition performance. First, lexically easy words should be associated with a relatively high level of accurate recognition. By definition, these words belong to a sparse neighbourhood in which there are only a few words that compete for recognition during initial activation. Therefore, during recognition, the four-stage process outlined above should proceed in a relatively rapid manner – because there are fewer items in the neighbourhood, fewer items compete for activation, thus making the process of eliminating competitors faster and easier by reducing inhibitory demands relative to hard words. Secondly, lexically hard words should be associated with a relatively lower level of accurate recognition as compared to lexically easy words. Such words belong to a dense neighbourhood with many words potentially competing for activation. In this case, the word recognition process should be more difficult – now with a greater number of words within the auditory neighbourhood, inhibitory demands are increased as more competing neighbours must be inhibited in order to identify the target word.

In sum, lexically easy words should yield higher and more accurate recognition performance than lexically hard words. The findings of Luce and Pisoni (1998) are consistent with these hypotheses, providing evidence from both recognition accuracy speed – lexically easy words yielded higher accuracy and faster recognition, while lexically hard words yielded poorer accuracy and slower recognition. As Luce and Pisoni also observed, a number of studies have confirmed density effects on recognition accuracy to persist in the presence of perceptual noise added during presentation of the speech stimulus, wherein words from low density neighbourhoods are recognised more accurately than those from high density neighbourhoods

(Bradlow & Pisoni, 1999; Dirks, Takayana, & Moshfegh, 2001; Takayanagi, Dirks, & Moshfegh, 2002, Vitevich & Luce, 1999).

Collectively, these findings are consistent with the framework put forth by the NAM, the hypothesis of which is an activation-competition framework mediates the process of auditory spoken-word recognition, in which recognition of target words is facilitated by inhibition of competing neighbours. The next section will discuss this framework as it applies to theories of cognitive aging, as well as direct evidence for the role of inhibition in spoken-word perception.

III. Cognitive Aging & the NAM

There is limited work investigating how age-related changes in cognitive processing affect the architecture and associated processing within the speech perception system. Given the prior discussion on age-related sensory impairments in speech perception, the current discussion now turns to age-related cognitive changes that may also be involved, particularly the mechanisms of spoken-word recognition as put forth by the NAM. As with the effects of lexical difficulty upon word recognition, predictions can also be made concerning the effects of age, and specifically, age-related cognitive declines in inhibitory control.

A major theory in the field of cognitive aging has been the *inhibitory deficit hypothesis* (Hasher & Zacks, 1988; McDowd & Filion, 1992; McDowd & Oseas-Kreger, 1991). The hypothesis asserts that there are age-related declines in the ability to inhibit or suppress irrelevant information from working memory (WM). Specifically, three types of inhibitory control have been proposed: *access*, *deletion*, and *restraint* (Hasher & Zacks, 1988). Access refers to the inhibitory control of preventing items from entering WM. Deletion refers to the removal of items from WM which are no longer relevant. Finally, restraint refers to the inhibition of strongly activated responses which could incorrectly influence cognition/behaviour. Evidence for all three types of inhibition have been observed in a number of paradigms, such as the Stroop task (Houx, Jolles, & Vreeling, 1993; McDowd, 1997), off-target verbosity (Arbuckle & Gold, 1993), directed forgetting (Conway, Harries, Noyes, Racsma'ny, & Frankish, 2000; Zacks, Radvansky , & Hasher, 1996), the Flanker task (Verbruggen, Liefooghe, & Vandierendonck, 2004), and many others. Declines in inhibitory control for adults are manifest in many forms such as poorer recall and higher proportions of false memories (Balota et al., 1999; Craik & McDowd, 1987). Moreover, inhibition has been implicated to play an important role in speech and language processing (e.g. Berg & Schade, 1992; Wurm & Samuel, 1997), and is considered here within the context of the NAM.

As previously described, inhibition is posited to be one of the mechanisms involved in the NAM's conception of auditory spoken-word recognition. Thus, further predictions can be made concerning age differences in auditory speech perception of lexically easy and hard words, with particular focus on the contribution of age-related declines in inhibitory control. First, older adults should achieve a similar level of recognition accuracy of easy words as younger adults (if audibility is equated). The rationale for this prediction is based on the relative inhibitory demands – fewer competitors experience activation and consequently require less inhibition in order to be eliminated as candidates. That is, for easy words deactivation of competitors requires relatively minimal inhibitory demands. Without this burden of inhibitory demands and all other factors held constant, younger and older adults should be able to perform similarly in their ability to recognise words with few competitors.

In contrast, older adults should be disproportionately worse at recognising lexically hard words. For these types of words, total activation levels within the system are heightened overall, and so greater inhibition is required to reduce activation of these competitors to below threshold in order to identify the target word. In cases, such as older adults, where inhibitory control may be impaired, this would result in continued (partial) activation of competing lexical candidates, thus reducing the ability of the system to have activation that exceeds threshold on the unique target word. Thus, age-related decreases in recognition accuracy of hard words would result in part from impairments in inhibitory control that are necessary to suppress activation of competing words.

The first study testing these age-related differences in the effects of lexical difficulty on recognition was conducted by Sommers (1996). A series of experiments compared younger and older adults in their ability to identify words varying in lexical difficulty (i.e. easy vs. hard). Sommers observed that both younger and older adults achieved a higher level of recognition accuracy for easy words than for hard words, a finding consistent with the first set of predictions from the NAM. Moreover, it was observed that older adults exhibited a significantly greater difference between easy and hard items than did younger adults, such that recognition accuracy for easy words between age groups was roughly equivalent, while older adults' accuracy for hard items was significantly lower than for young. It is worth noting that these age differences could not be accounted for by differences in sensitivity. These findings were replicated in a follow-up experiment which presented the words in a background of white noise; while overall performance was expectedly lower for both groups as compared to Experiment 1, the exaggerated difference between younger and older adults' identification of hard words remained. The interaction of age and lexical difficulty was also observed in a later study by Sommers and Danielson (1999). Using a similar experimental manipulation in which younger and older listeners were instructed to identify lexically easy and hard words presented in a background of speech-shaped noise. An identical pattern of results was observed here, such that younger and older adults performed comparably in identifying easy words, but older adults performed significantly worse in identifying hard words.

To investigate the contribution of inhibition in accounting for the age differences in identifying hard words, Sommers and Danielson (1999) examined the pattern of errors for incorrect responses: if competing candidates for hard words fail to be inhibited and remain active, then they should exhibit a level of activation similar to the intended target word and thus be more likely to be produced during identification. That is, incorrect responses to hard words should be phonetic neighbours of the target word. Moreover, if older adults have more difficulty reducing activation levels on competitors due to inhibitory deficits, then they should have an increased probability of incorrectly responding with a lexical neighbour. The findings confirmed this, demonstrating that age differences in the probability of producing neighbours varied as a function of lexical difficulty, such that older adults were approximately 30% more likely to produce neighbours as errors to hard words than to easy words, while younger adults were only 8% more likely.

A second experiment directly examined the role of inhibition in accounting for age-related differences in lexical difficulty, using an index of inhibition (composed of an auditory Garner speeded classification task [Garner, 1974], and an auditory Stroop task [Jerger et al., 1993]) to predict the ability to recognize lexically easy and hard words on the basis of individual and age differences in inhibitory control. Not only did older adults perform more poorly on the inhibition tasks, the investigators found that the index of inhibition was significantly correlated with identification performance of lexically hard words, but not easy words. Indeed, a regression analysis revealed that inhibition accounted for roughly 28% of unique variance in hard word identification, independent of the effects of education, vocabulary and age. Moreover, age was also a unique contributor to performance, independent of inhibitory ability (~10%). These results were interpreted to suggest that due to older adults' reduced inhibitory ability; they are less able than younger adults to inhibit lexical neighbours during word perception. Taler, Aaron, Steinmetz, and Pisoni (2010) also demonstrated the relationship between inhibition and identification performance for high ND words in older adults, observing a significant negative correlation between Stroop interference and the accuracy for high ND words. Thus, age-related impairments in inhibitory control appear to be a significant contributor to difficulties in recognising words with many lexical neighbours (i.e. hard words).

Overall, the body of work to date concerning the relationship between age-related impairments in cognitive inhibition and identification accuracy of lexically easy and hard words suggest a clear relationship between the two, and provides evidence to support the role inhibitory mechanisms in the NAM. Despite critiques of this hypothesis (Burke, 1997), the predictions regarding the activation-competition framework of this model remain well-supported by the literature.

IV. AV Speech Perception & the NAM

Thus far, the discussion concerning lexical density has been exclusively in the domain of auditory spoken-word recognition. The final section introduces the concept of AV lexical density, and its significance in the current experiment.

As in the auditory domain where words are recognised in the context of other words, i.e. neighbourhoods, research has extended this to the domain of visual speech perception. Mattys, Bernstein, and Auer (2002) found that visually similar words appear to operate according to principles similar to those observed for auditory presentation, such that words presented visually produce a pattern of identification similar to that observed with auditory presentations; lexically easy words have relatively higher rates of identification than lexically hard words. These visually similar words are defined on the basis of containing certain groups of phonemes, such as /f/ and /v/, that are indistinguishable based on V-only information because they appear similar on the face and lips. These phonemic units, based on the addition, substitution, or deletion of a single visual unit, are called *visemes*, and are the visual analog of the auditory unit phoneme, defined as a visual motor pattern that distinguishes speech sounds (Jackson, 1988). Accordingly, similar to auditory neighbourhoods, visual neighbourhoods can be either sparse or dense, depending upon the number of visual neighbours of a target item that are activated during visual word recognition. In a further examination of visual speech perception, Auer (2002) extended the NAM to predicting visual spoken-word identification accuracy, examining whether ND principles would also apply to speech stimuli when presented visually. The investigators reported that after controlling for segmental intelligibility (i.e. the acoustic-phonetic stimulus properties), low ND were more readily identified than high ND words. These findings were interpreted as providing support for similar operations underlying visual and auditory word recognition

Tye-Murray et al. (2007) subsequently proposed and created an auditory-visual (AV) neighbourhood structure combining both auditory and visual neighbourhoods. These neighbourhoods constitute the region of overlap between the auditorily and visually defined neighbourhoods for a particular word – that is, words that are both auditory and visual neighbours of the target item. This overlap between unimodal neighbourhoods can be referred to as the *intersection density* (ID); candidate items for words presented in AV are thus items that are both A and V neighbours, i.e. those that fall within the ID. These can be further placed within the context of the earlier discussion about lexically easy and lexically hard words, in that easy words are those that have relatively low ID because of only few competing items within the region of neighbourhood overlap, and hard words are those with a higher ID (i.e., more items competing within the region of overlap). Figure 1 displays a schematic of the relationship between auditory and visual neighbourhoods and their overlap. The top set of circles indicate auditory (A) and visual (V) neighbours for the words *fork* (a) and *fish* (b). The bottom set of circles depicts the overlap between the unimodal neighbourhoods when both are activated simultaneously during AV presentation. Although both words have comparable unimodal neighbourhood sizes, the intersection density is much larger for *fork* than it is for *fish*, thus making it more likely that an individual will recognise *fish* than *fork* in AV presentation, given the fewer number of competitors.

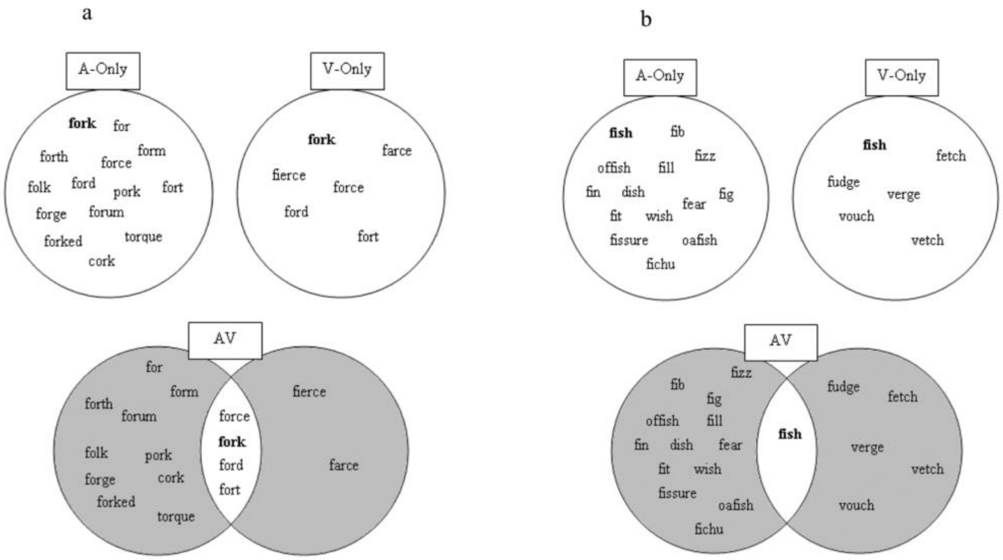

Figure 1. Borrowed from Tye-Murray et al. (2007). A schematic depiction of the intersection density (ID) created as a result of simultaneous activation of unimodal sensory neighbourhoods.

According to Tye-Murray et al. (2007), the process of spoken-word identification in AV presentation proceeds in a nearly identical manner to the unimodal case; a speech stimulus presented both acoustically and visually activates the target word, as well as modality-specific candidate words in the region of overlap that are also coactivated by the two speech signals. These competitors must be inhibited in order to activate the target word, whereas those words outside of the intersection receive minimal or no activation[1]. Therefore, words with few competitors should be more easily identified than words with more competitors. Tye-Murray et al. tested this hypothesis using a list of words from the Children's Audiovisual Enhancement Test (CAVET). They divided words on this test into groups of low (i.e. easy) and high (i.e. hard) intersection density. As predicted, items with low IDs were recognised significantly more often than those with high IDs, supporting both the existence of an AV neighbourhood structure in spoken-word recognition, and suggesting that similar mechanisms may underlie unimodal and bimodal word recognition.

Using these principles set forth by the NAM, it is thus possible to conceptualise an adapted model which considers the dynamics of activation-competition in the AV speech system. This model will henceforth be referred to as the Auditory-Visual Neighborhood Activation Model (AV-NAM).

The AV advantage becomes especially apparent in the framework of this bimodal framework. By perceiving a word through both an acoustic and visual signal, the number of activated words in the intersection density is immediately reduced in comparison to the speech signal presented in either modality alone. For example, the word *fork* has 13 auditory neighbours and 5 visual neighbours, as depicted in Figure 1. When the word *fork* is presented under AV conditions however, the number of neighbours is reduced to three, such that there are only three common candidates between the auditory and visual neighbourhoods. Moreover, the additional

---

[1] Words that are outside the intersection density area also receive activation, but are not candidates for recognition because they do not reside within the intersection density overlap.

benefit of the visual signal to the existing acoustic signal also becomes apparent. Visual enhancement within the context of the AV-NAM specifically refers to the reductions of items in the ID relative to unimodal neighbourhoods, such that a large number of A-only candidates can be eliminated quickly as potential candidates because they are not also visual neighbours and therefore do not enter into the region of overlap.

One prediction from the AV-NAM with respect to aging is that, as was the case with A presentation, older adults should have a disproportionately greater difficulty in identifying hard words than easy words, as compared to younger adults. However, contrary to this hypothesis, Tye-Murray et al (2007) observed no Age x Lexical Difficulty interaction for AV presentations. The authors tested both younger and older participants and although both had lower performance for lexically hard words, there was no interaction with age, although the pattern was trending in the predicted direction. In addition, the focus of this study was on the lexical characteristics of the AV neighbourhood, and therefore did not assess the contributions of inhibitory mechanisms.

V. Overview of the Current Experiment

The present study was designed to extend the findings of Tye-Murray et al. (2007) by investigating the role of inhibition in AV speech perception using the framework of the AV-NAM. The current experiment tested younger and older adults on three separate tasks of cognitive inhibition, as well as a spoken-word identification task in A, V, and AV conditions. Individuals were also assessed for vocabulary, processing speed, and hearing ability. In order to determine the degree to which inhibitory ability predicts age-related and individual differences in identification of words varying in lexical difficulty, regression techniques were used to examine word identification performance as a function of inhibitory ability. If AV perception of spoken-words is affected by the number of competing words that are simultaneously activated by unimodal input, older adults’ AV perception may also suffer because their inhibitory deficits result in larger numbers of competing word candidates common to both modalities remaining active and interfering with correct identification. Moreover, because inhibitory demands are increased for perception of high ID words, a stronger relationship between inhibition and intelligibility of high ID words is expected. Additionally, this experiment also examined the relationship between inhibitory ability and visual enhancement to determine whether the ability to benefit from visual information added to the acoustic signal differs as a function of age-related or individual differences in inhibitory ability. Together, these findings will further extend our current knowledge of the role of cognitive mechanisms involved in AV spoken-word identification, and how this process may elucidate the changes in speech perception that accompany the aging process.

Methods

Participants

One hundred and seven total participants were recruited for this study. Fifty-four young adults (42 women) were recruited from the Washington University (WU) Psychology Subject Pool. Fifty-three community-dwelling older adults (30 women) were recruited from the subject pool maintained by the WU Aging and Development Program.

All participants were tested for hearing ability, processing speed, and vocabulary knowledge. To assess hearing ability, participants received a hearing test measuring pure-tone air-conduction thresholds (PTA) for octave frequencies from 250 Hz to 4,000 Hz. The PTA average (the average threshold at 500, 1000, and 2000 Hz) was obtained for each individual. Participants also completed two tasks of processing speed. In the dot-location task with 40 trials, (Chen, Myerson, & Hale, & Simon, 2000), participants were to respond as quickly as possible as to which of two dots was closer to a central target dot. In a category judgement task with 80 trials, (Chen, Hale, & Myerson, 2007), participants were to respond as quickly as possible as to which of two words presented side by side in the centre of the screen was an animal (the alternate choice was a non-animal, such as an inanimate object or food item). All participants also completed the Vocabulary subtest of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955). Additionally, older adults completed the Mini-Mental Status Examination (MMSE) (Folstein, Folstenin, & McHugh, 1975). Table 1 depicts this information, in addition to demographic variables. Younger and older adults did not significantly differ on WAIS Vocabulary scores, but did predictably differ on hearing ability and speeded response times. However, both groups exhibited thresholds that were within clinically normal limits (less than 20 dB hearing loss) for all but the highest frequency, and response times were consistent with age-related slowed responding (e.g. Salthouse, 1996).

Table 1

Demographic, Hearing Loss, and Processing Speed Information for Younger and Older Adults

Variable	Younger Adults	Older Adults
N	54	53
Mean Age	20.36 (.21)	71.98 (0.99)
MMSE	-	28.98 (0.15)
WAIS-IV Vocabulary	63.36 (.69)	63.42 (1.25)
PTA Left dB HL	4.92 (.58)	21.16 (1.59)**
PTA Right dB HL	5.16 (.56)	19.73 (1.39)**
Dot Location Task (in ms)	635.43 (17.23)	987.04 (38.08)**
Dot Location Accuracy (%)	98.40 (0.55)	97.47 (0.97)
Category Judgment Task (in ms)	596.13 (15.84)	742.63 (24.45)**
Category Judgment Accuracy (%)	100.00 (0.00)	96.60 (1.91)

* $p < .001$

Note. Standard error of the mean shown in parentheses.

Tasks

Visual colour-word Stroop task.

Three measures of inhibitory control were assessed in this experiment. All tasks were performed in the same room, where participants were seated approximately 50 cm away from the computer screen. The first task was the Stroop (1935) colour-naming paradigm, which has been used extensively in prior literature to examine age-related differences in inhibitory control (e.g., Earles et al., 1997; Hasher, Stoltzfus, Zacks, & Rypma, 1991; West & Alain, 2000).

Stimuli. The stimuli for color Stroop task consisted of four colour words (*red, blue, green, yellow*), and four neutral words (*deep, legal, poor, bad*) presented randomly in either a red, blue, green, or yellow hue (Spieler, Balota, & Faust, 1996). A total of 80 trials were presented, consisting of 32 congruent (the colour word and hue are identical), 24 incongruent (the hue differs from the colour word), and 24 neutral (a neutral word) trials.

Procedure. This task was presented in E-Prime 2.0 (Psychological Software Tools, Inc., Pittsburgh, PA). In the task, participants were told to name the colours in which the word appeared as quickly and as accurately as possible. Participants were given 16 practise trials before beginning the experimental trials, and the experimenter remained in the room for the duration of the testing session. Each trial began with a fixation stimulus consisting of three plus signs displayed for 700 ms. The screen was then blank for 50 ms, followed by the appearance of the stimulus word. The word was displayed on a black background and remained on the screen until the participant responded vocally into a microphone. Once the voice-operated relay for response latency was triggered, the experimenter pressed one of four keys to code the participant's colour response.

Eriksen Flanker task

The second task of inhibitory control was the Eriksen and Eriksen (1974) Flanker task, adapted from a version used by Friedman and Miyake (2004), in which the task was used as an assessment of how well participants were able to select targets that were simultaneously presented with irrelevant distractors. In this task, participants must respond as quickly and accurately as possible to the identity of a central target that is flanked by distracting items.

Stimuli. The stimuli consisted of four different targets: the letters H, K, S, and C. The flanker letters consisted of these same letters, but presented in a string of four (two on each side of the central target letters). Participants were instructed to press a button on the right side of the keyboard when the target letter was H or K, and a button on the left side when the target letter was S or C. A congruent trial consisted of a target letter flanked by identical distractors (e.g., *HHHHH*); an incongruent-compatible trial consisted of a target letter flanked by different distractor letters but corresponding to the *same* response (e.g., *KKHKK*); an incongruent-incompatible trial consisted of a target letter flanked by different distractor letters and corresponding to the *opposite* response (e.g., *SSHSS*); and a neutral condition consisted of a target letter presented in isolation (e.g., *H*). There were 160 total trials, consisting of 35 congruent trials, 35 compatible incongruent trials, 40 incompatible incongruent trials, and 50 neutral trials.

Procedure. This task was presented in E-Prime. Participants were given 16 practise trials before beginning the experimental trials. Each trial began with a fixation stimulus consisting of three plus signs displayed for 700 ms. The screen was then blank for 50 ms, followed by the appearance of the stimulus. The stimulus was displayed in black letters on a white background and remained on the screen until the participant pressed a key which triggered the next stimulus.

Phonological DRM Task

The final task of inhibitory control included in the experiment was the phonological variant of the Deese-Roediger-McDermott (DRM; Roediger & McDermott, 1995) task developed to examine false memories by Sommers and Lewis (1999). False phonological memories occur when a participant falsely recognises a critical word which was not previously presented, but is a phonologically-related associate of all words in a given list (e.g. an ‘old’ response to the critical word CAT, whereas CAB, THAT, and MAT were presented previously).

Stimuli. Eight lists were randomly selected from the 24 lists generated by Sommers and Lewis (1999). Each list consisted of 15 phonological neighbours of a critical item (CI) specific to that list. All stimuli were recorded by a male talker with a Midwestern American dialect and presented auditorily to the participants through headphones.

Procedure.

Study phase. The task was presented in SuperLab 4.5.4 (Cedrus Corporation, 1999). Participants heard one list at a time, separated by a 500-Hz tone signifying the beginning of the next list. Within each list, words were separated by a 500-ms interstimulus interval.

Test phase. After all the lists had been presented, an old/new auditory recognition task began in which participants were presented with a test list of 24 old and 24 new items. Of the 24 new words, 8 were CI's of the previously presented lists and the remaining 16 new items consisted of words randomly selected from non-presented lists. The 24 old items were words pseudo-randomly selected from the presented lists. Participants were instructed to press a key corresponding to OLD if they had heard the item in the study phase, or a key corresponding to NEW if they had not heard the item in the study phase.

Spoken-Word Identification Task.

Stimuli. One hundred and eighty words were obtained from the English Lexicon Project (Balota et al., 2007). Words were selected on the basis of words that are both A neighbours and V neighbours, i.e., intersection density (ID), in order to obtain sufficient numbers of both high ID and low ID items in order to obtain the strongest density effects. High and low ID values were determined using the 10C4V method, in which V neighbourhoods are created by categorising all of the phonemes in each word as one of 10 consonants or 4 vowels (from Iverson, Bernstein, &

Auer, 1998). These words consisted of both monosyllabic and disyllabic nouns and verbs (see *Appendix* for the complete stimulus list). Hyperspace Analogue to Language (HAL) log-transformed frequency norms (Lund & Burgess, 1996) across low and high words were also obtained and equated as closely as possible in order to assess neighbourhood properties independently of word frequency. Table 2 depicts mean ID and HAL log frequency for both easy and hard words.

Table 2
Lexical Characteristics of Low Intersection Density Words and High Intersection Density Words

	Easy Words	Hard Words
Mean Intersection Density (SD)	1.0 (0.0)	11.1 (2.4)
Mean HAL Log Frequency (SD)	9.1 (2.3)	8.8 (2.1)

SD = standard deviation.

All stimuli were recorded by a young female talker who had a Midwest American dialect. During recording, only the head and shoulders of the talker were visible, filmed with high-intensity studio lighting before a neutral background. Each stimulus presentation began with the talker saying the carrier phrase, “Say the word”, followed by the target word, and lasted for approximately 5 s.

Procedure. Participants were tested in a sound-attenuated room. Stimuli were presented via a PC (Dell 420) equipped with a Matrox (Millennium G400 Flex) 3D video card and through headphones (Sennheiser HD 265). Video configuration was in dual screen mode so that stimuli could be presented to participants in the test booth while the experimenter monitored participants’ responses in a control room. The participant was instructed to listen to the entirety of the stimulus (“Say the word _____”), and then verbally repeat the terminal word into a microphone. Each sentence stimulus was cued by a prompt on the screen indicating the mode of presentation: an eye for V-only (“watch”), an ear for A-only (“listen”), and both an eye and ear for AV (“watch and listen”). During presentation, a 6-talker babble noise was added to the auditory stimulus using a sound processing delivery system (RP 2, Tucker-Davis Technologies System 3, Gainesville, FL). The signal-to-babble (SNB) was set at -1 dB across all participants to prevent ceiling performance during identification. The stimuli were randomly distributed equally across A-only (the auditory stimulus mixed with the babble without seeing the talker), V-only (the talker’s head and shoulder’s also presented in babble without the auditory stimulus), and AV (seeing and hearing the talker overlaid with babble) presentation. The same presentation order was used for all participants to eliminate presentation order as a source of variance. Participants were encouraged to provide a response on every trial, regardless of certainty. A correct response was only counted if the participant repeated the word verbatim. The first three trials were considered practise trials.

The entire experimental session lasted approximately 2 hours. Participants were compensated $10 or 1.0 course credit per half-hour of participation. All procedures were approved by and adhered to guidelines outlined by the Human Research Protection Office at Washington University in St. Louis.

Results

Several participants had to be eliminated from analyses due to a combination of incomplete/missing data from one or more tasks and/or inappropriate inclusion in the age group (i.e. a few of the older adult participants had signed up through the PSYC 100 subject pool), leaving a total of 99 participants (younger, n = 49, 38 women; older, n = 50, 30 women).

Stroop task.

Response latencies exceeding two standard deviations of the mean for each condition (congruent, neutral, and incongruent) were excluded from the analyses. This resulted in the exclusion of approximately 6% of the data from younger adults, and 8% from older adults. A small proportion of these excluded trials contained no-response trials. The mean response latencies and accuracy are displayed in Table 3, as well as the Stroop effect for each group and z scored (standardized) means. A mixed analysis of variance (ANOVA) with age as a between-subjects variable and Stroop condition as a repeated-measures variable revealed that older adults had significantly slower response latencies, $F(1,97) = 89.99$, $\eta^2 = .57$, $p < .001$. Response latencies also significantly differed across the three conditions, $F(2,96) = 234.03$, $\eta^2 = .43$, $p < .001$. Critically, there was a Age x Condition interaction, $F(2,194) = 37.78$, $\eta^2 = .35$, $p < .001$. The interaction remained significant when analysed using the standardised means, which were calculated by transforming each reaction time to a z score based on each participant's mean and standard deviation (Faust et al., 1999), $F(2,194) = 12.43$, $\eta^2 = .21$, $p = .039$. The Stroop effect, or the Stroop interference score, was calculated for each participant by subtracting the mean response latency in the congruent trials from the mean response latencies in the incongruent trials. Stroop interference was significantly greater for older adults than for younger adults, $t(98) = 7.22$, Cohen's $d = 1.45$, $p < .001$.

As with response latencies, ANOVAs were also conducted on the mean accuracy data. Given that both older and younger participants achieved near ceiling performance in all conditions, there were no main effects or interactions. Thus, the results from the Stroop task suggest a reliable age-related increase in interference that is unrelated to a possible speed-accuracy tradeoff.

Table 3

Response Latencies (in Milliseconds) and Accuracy (in Percentage) in the Stroop Task

	Young (*n* = 49)	Old (*n* = 50)
Condition	Mean response latencies (ms)	
Congruent (SE)	584.6 (8.5)	705.82 (14.1)
Neutral (SE)	617.4 (8.3)	769.53 (13.6)
Incongruent (SE)	680.2 (10.9)	898.21 (19.1)
	Mean standardised response latencies (ms)	
Congruent (SE)	-.61 (.04)	-.77 (.07)
Neutral (SE)	-.13 (.03)	-.21 (.04)
Incongruent (SE)	.52 (.03)	.67 (.04)
Stroop Interference (SE)	95.6 (5.4)	192.42 (12.2)
	Accuracy (%)	
Congruent (SE)	96.8 (0.9)	98.3 (0.2)
Neutral (SE)	96.6 (0.7)	99.9 (0.1)
Incongruent (SE)	97.7 (0.6)	93.3 (0.3)

SE = standard error of the mean.

Flanker task

Similar to the Stroop task, response latencies exceeding two standard deviations of the mean for each condition were excluded from analyses, resulting in the removal of approximately

5% and 7% of responses from younger adults and older adults, respectively. One older participant and one younger participant were excluded for significantly poor accuracy. The remaining average response latencies within each condition and accuracy are displayed in Table 4. A 2 (age) x 3 (condition) mixed ANOVA revealed that older adults were slower than younger adults in all conditions, $F(1,95)= 45.91$, $\eta^2 = .49$, $p < .001$. There was also an effect of condition, $F(3,93)= 33.15$, $\eta^2 = .14$, $p < .001$. However, there was no Age x Condition interaction, $F(3,285)= 2.12$, $\eta^2 = .01$, $p = .09$. ANOVAs were also conducted on the mean accuracy data, and revealed no main effects or interactions. Thus, in contrast to the Stroop task, there did not appear to be a significant effect of age-related interference.

Table 4
Response Latencies (in Milliseconds) and Accuracy (in Percentage) in the Flanker Task

	Young (*n* = 49)	Old (*n* = 49)
Condition	Response latencies	
Congruent (SE)	560.1 (16.6)	753.6 (26.8)
Compatible incongruent (SE)	599.1 (20.1)	830.3 (31.6)
Incompatible incongruent (SE)	573.1 (13.2)	802.0 (30.9)
Neutral (SE)	528.4 (14.9)	743.4 (27.2)
Interference (SE)	75.8 (8.2)	94.7 (8.9)
	Accuracy (%)	
Congruent (SE)	99.6 (0.1)	98.6 (0.1)
Compatible incongruent (SE)	99.5 (0.1)	98.8 (0.1)
Incompatible incongruent (SE)	99.1 (0.1)	97.4 (0.2)
Neutral (SE)	99.0 (0.1)	98.5 (0.2)

SE = standard error of the mean.

Phonological DRM task

Table 5 displays the mean proportion of ‘OLD’ responses to old items, new items, and critical lure items. Old responses to critical lure items represent the false alarm rate. A 2 (age) x 3 (condition) repeated-measures ANOVA revealed a significant effect of age, $F(1,97) = 3.10$, $\eta^2 = .09$, $p = .049$; and condition, $F(2,96) = 323.67$, $\eta^2 = .694$, $p < .001$. There was also a significant Age x Condition interaction, $F(2,194) = 4.07$, $\eta^2 = .01$, $p = .02$. Post-hoc pairwise comparisons with Bonferroni corrections revealed that older adults performed similarly to younger adults in the proportion of ‘OLD’ responses to old and new items, but had a significantly higher proportion of ‘OLD’ responses (i.e. false alarms) to critical lures than younger adults, $p < .01$.

Table 5

Proportion of ‘OLD’ Responses to Items in the Phonological DRM Task

	Young (*n* = 49)	Old (*n* = 50)
Condition		
Old Items	71.8 (1.6)	71.4 (1.0)
New Items	26.7 (1.7)	29.7 (2.4)
Critical Lures (False Alarm Rate)	45.1 (2.6)	56.7 (3.0)

Note. Parentheses indicate standard error of the mean.

Relationship Between Inhibition Measures

To assess whether the three tasks shared a common construct, Pearson product-moment correlations were computed. The results are presented in Table 6. None of the correlations were significant, indicating that the three tasks may not be tapping into the same aspect of inhibition.

Table 6

Pearson Product-Moment Correlations between the interference measures from the Stroop, Flanker, and DRM task

	Stroop interference	Flanker interference	DRM interference
Stroop interference	—	.08	.13
Flanker interference		—	.17
DRM interference			—

Spoken-word Identification Task

The practise (i.e. the first three) trials from each participant were excluded from analyses. The exclusion of these trials, which happened to be all in A-only presentation, did not affect the overall pattern of the data.

Figure 2 depicts overall identification in each modality (A-only, V-only, and AV) between the two age groups. A 2 (age) x 3 (modality condition) factorial ANOVA revealed a significant interaction, F (2,194) = 37.42, $\eta^2 = .02$, $p < .001$. Pairwise comparisons with Bonferroni corrections indicated that younger adults demonstrated superior identification performance of words presented in A and AV conditions compared to older adults (both $p < .001$), but demonstrated comparable performance to older adults in the V condition ($p = .179$); this absence of age differences may be due to near floor level performance.. The results from AV will be the focus of the rest of the analyses.

Because participants were not equated for audibility, it was particularly important to examine whether the age differences in AV performance were due to reduced audibility in older adults. A one-way ANCOVA was conducted to compare AV performance between young and older adults, using hearing ability as a covariate. The age difference was still significant with a mean difference of 11.33 percent, $F(1,96) = 37.63$, $\eta^2 = .21$, $p < .001$, suggesting a reliable difference between younger and older participants, even after controlling for hearing ability. A plot depicting the range in identification performance for younger and older adults across the three conditions is shown in the Appendices as Figure A2 and A3.

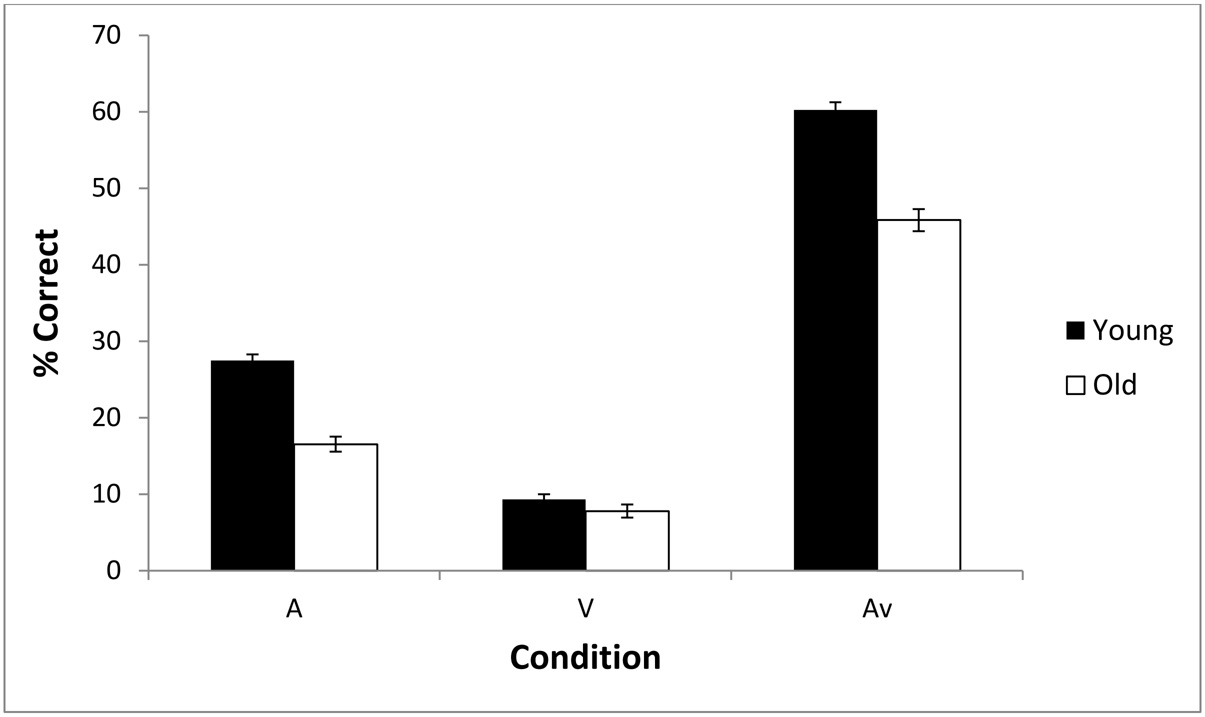

Figure 2. Percentage correct word identification in auditory-only (A), visual-only (V), and audiovisual (AV) conditions. Error bars represent the standard error of the mean.

Figure 3 depicts identification performance as a function of age and intersection density (ID, i.e. low vs high) in AV performance only. A 2 (Age) x 2 (Intersection Density) ANOVA revealed a main effect of Age, $F(1,97) = 65.74$, $\eta^2 = .42$, $p < .001$, in which younger adults performed significantly better than older adults; and a main effect of ID, $F(1,97) = 46.93$, $\eta^2 = .79$, $p < .001$, in which low ID items were recognised more often than high ID items. There was also a significant Age x ID interaction, $F(1,97) = 8.21$, $\eta^2 = .03$, $p = .005$, indicating that while older adults are generally poorer at recognising AV words overall, they are significantly impaired at recognising high ID words in particular.

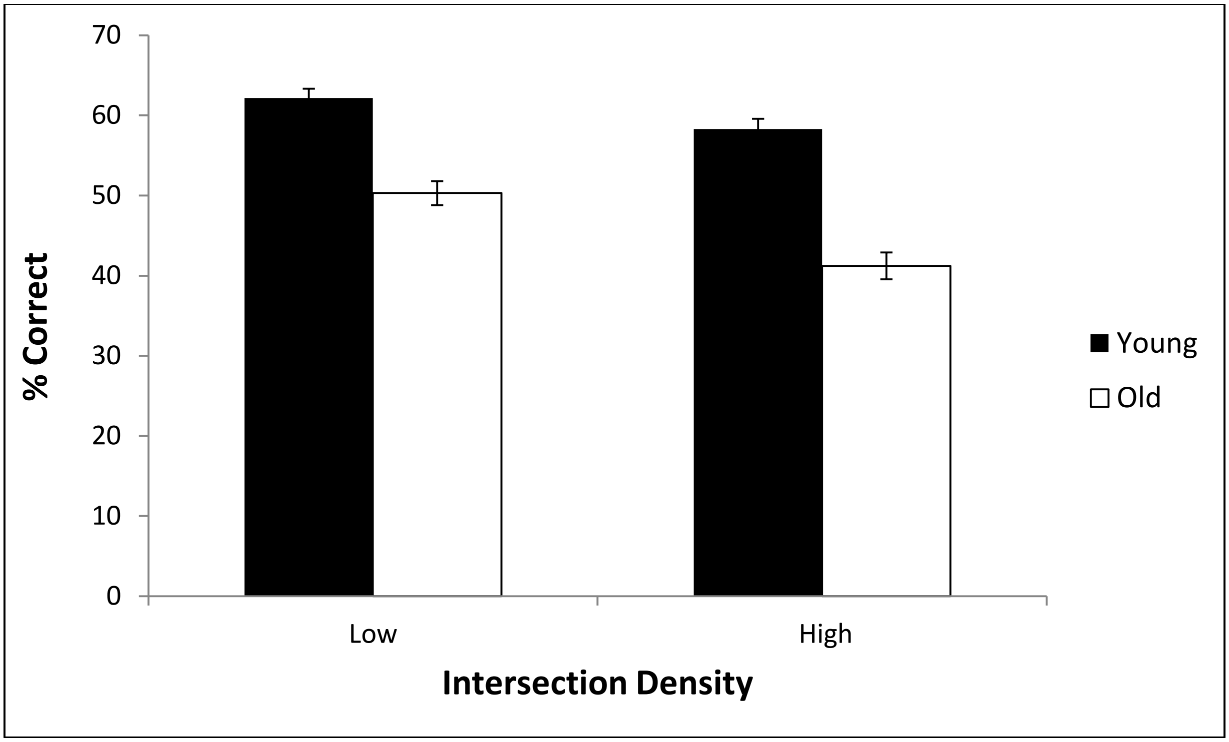

Figure 3. Percentage correct word identification in AV as a function of low ID words and high ID words. Error bars represent the standard error of the mean.

Error Analysis

An error analysis was then conducted to examine what types of errors participants were producing. The rationale for this stems from the activation-competition mechanism of the NAM. During perception of high ID words, a large number of competitors receive activation, and may remain active for some time before being inhibited. Low ID words do not necessitate the same degree of inhibition due to fewer competitors. Therefore, if deficits in inhibiting lexical neighbours are partly responsible for age-related differences in recognising high ID words, then older adults ought to be more likely to produce neighbours (i.e. persistently active candidates) as incorrect responses.

The percentage of errors which were lexical neighbours to the target item was computed (i.e. neighbourhood errors) in order to address this prediction. Table 7 depicts the overall percentage of neighbourhood errors in AV, as well as neighbourhood errors made in response to low ID items and high ID items. An independent-samples t-test indicated that older adults produced significantly more neighbourhood errors overall, $t(97) = 3.54$, Cohen's $d = .71$, $p < .001$. A repeated-measures ANOVA was then conducted to determine whether the percentage of ID errors differed depending on whether the target item was a low ID or high ID item. The results revealed a significant effect of ID, $F(1,97) = 527.03$, $\eta^2 = .84$, $p < .001$, in which a significantly higher proportion of errors were made to high ID items. The interaction between Age x Intersection Density was only marginally significant, $F(1,97) = 3.73$, $\eta^2 = .01$, $p = .056$. Despite not reaching strict criterion for significance, the trend indicates that older adults appear to produce more neighbourhood errors than do younger adults, particularly to high ID words.

Table 7

Percentage of Error Responses Classified as Lexical Neighbours to the Target Item in Younger and Older Adults

Young (n = 49)			Old (n = 50)		
Tot. Errors	Low	High	Tot. Errors	Low	High
14.92 (.84)	13.43 (3.14)	86.56 (3.25)	19.14 (.85)	6.83 (1.26)	93.17 (1.25)

Note. Standard error of the mean in parentheses.
Tot. Errors = the percentage of total errors that are neighbourhood errors (i.e. incorrectly producing a lexical neighbour)
Low = the percentage of incorrectly-produced lexical neighbours to low ID items
High = the percentage of incorrectly-produced lexical neighbours to high ID items

Visual Enhancement

The degree to which adding a visual signal improves word intelligibility (i.e. visual enhancement, or VE) was also compared between the age groups. VE is calculated as follows:

$$VE = (AV-A) / (1-A)$$

This formula (Sumby & Pollack, 1954) directly compares an individual's A and AV performance, such that differences in baseline A-only performance are normalised in order to calculate the overall improvement (relative to the maximum possible improvement)

as a result of additional visual information. Figure 4 depicts overall VE for both younger and older adults, while Figure 5 depicts VE as a function of age group and lexical difficulty. Using an independent-samples t-test analysis, a significant difference in VE was found between younger and older adults, such that younger adults obtained higher VE scores than older adults, $t(97) = 4.84$, SE = .02, Cohen's $d = .95$, $p < .001$. A repeated-measures ANOVA was then conducted to determine whether younger and older adults differed in VE scores, depending upon whether the item was easy or hard. The interaction was significant, $F(1,97) = 4.97$, $\eta^2 = .02$, $p =$

.028, indicating that older adults receive significantly less benefit than younger adults from an additional visual signal to hard words, as compared to easy words. This analysis was also repeated using V-only performance as a covariate, and a similar pattern of results was obtained As with the results from overall AV identification, an ANCOVA was conducted to compare VE scores between young and older adults, using hearing ability as a covariate. The difference was still significant with a mean proportion difference in VE of .08, $F(1,96) = 12.54$, $\eta^2 = .10$, $p < .001$, suggesting that reliable differences between younger and older participants persist after controlling for hearing ability. A scatterplot depicting the range of VE scores for young and old adults is shown in *Appendix A4*.

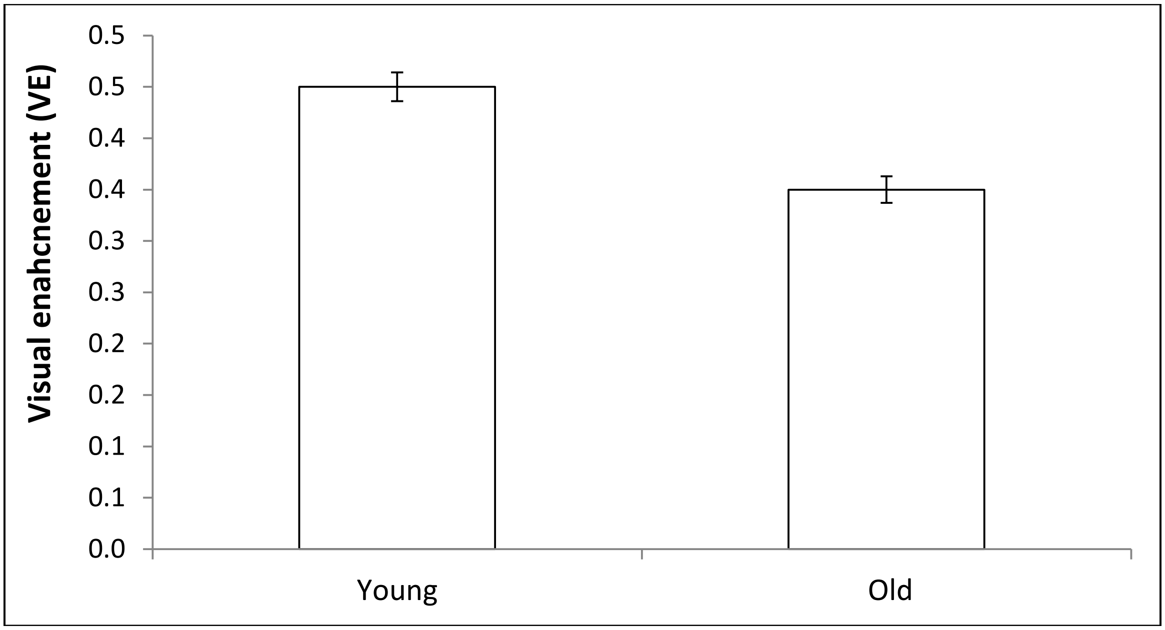

Figure 4. Overall visual enhancement (VE) scores for younger and older adults shown in proportion. Error bars represent the standard error of the mean.

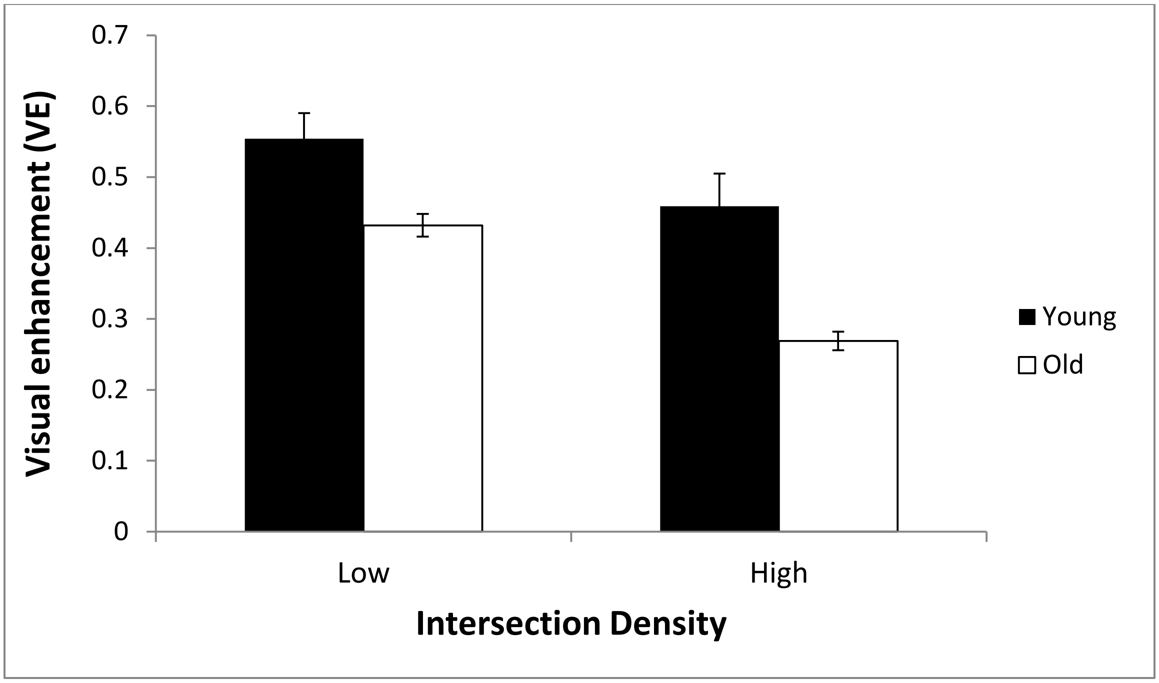

Figure 5. Visual enhancement (VE) scores for younger and older adults as a function of low ID words and high ID words. Error bars represent the standard error of the mean.

Inhibition and AV Word Identification

As demonstrated above, none of the three tasks designed to assess inhibition were correlated. However, two of the tasks – the Stroop and DRM task – produced age-related effects in interference. Nevertheless, all three tasks will be examined with respect to their relationship to AV word identification.

However, in order to properly isolate the effect of age-related inhibition, general slowing must be taken into account (Salthouse, 1996). Because the two measures of processing speed were significantly correlated (r = .55 at the p < .001 level), the two measures were combined into a speed index, created by standardizing response latencies for the two tasks and computing an average z-score for each participant (Faust et al., 1999).

Table 8 displays the Pearson zero-order correlations between the demographic variables (age and WAIS vocabulary), hearing ability (as measured by PTA in the better ear of each participant), processing speed, the three inhibition tasks, AV identification scores (overall, as well as for low ID items and high ID items), and VE scores (overall, as well as for low ID items and high ID items). The table clearly demonstrates correlations with age in a number of different factors, particularly the AV and VE scores, such that increases in age are associated with poorer AV and VE scores. However, there was no correlation between age and the Flanker or DRM task. In fact, the only inhibition task to be correlated with age was the Stroop task, indicating an age-related increase in Stroop interference as a measure of inhibition. Moreover, higher AV and VE scores were associated with lower Stroop interference. Together, these findings may suggest that the relationship between Stroop inhibition and AV word identification may be a result of the strong relationship Stroop interference has with age.

Table 8.

Pearson Product-Moment Correlations Between Demographic, Speed, Inhibitory, and AV Measures

Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
1. Age	—	.01	.67**	.37**	.59**	.16	-.16	-.64**	-.52**	-.64**	-.44**	-.25*	-.44**
2. Vocabulary		—	-.18	.02	-.05	-.20*	-.02	.01	.06	-.03	-.02	.09	.01
3. Hearing Ability			—	.13	.36**	.13	.07	-.63**	-.55**	-.61**	-.44**	-.24	-.46**
4. Speed Index				—	.43**	.11	.02	-.35**	-.35**	-.30**	-.16	-.13	-.18
5. Stroop interference					—	.08	-.01	-.56**	-.21*	-.57**	-.45**	-.15	-.36**
6. Flanker interference						—	.04	-.08	-.08	-.07	-.07	.11	.06
7. DRM interference							—	.02	-.02	.05	-.05	-.04	-.04
8. AV Total								—	.90**	.93**	.89**	.42**	.63**
9. AV (Low ID)									—	.69**	.81**	.50**	.47**
10. AV (High ID)										—	.82**	.28**	.67**
11. VE											—	.47**	.68**
12. VE (Low ID)												—	.78**
13. VE (High ID)													—

Note: AV = Audiovisual; ID = Intersection density; VE = Visual enhancement

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

Predictors of AV Word Identification

In order to investigate the unique contributions of age and Stroop inhibition to AV word intelligibility, a series of hierarchical linear regressions were conducted. Given the strong correlation between age and AV, the regressions were conducted separately for each age group. For each analysis, the hearing ability, speed index, as well as the measure of Stroop inhibition were used to predict AV and VE performance (because of the lack of significant correlations between the Flanker and DRM task and the AV and VE scores, these variables were omitted from further analysis). Hearing ability was entered into the first step, followed by the speed index in the next step, and Stroop interference in the final step. It is particularly important to note that processing speed was entered into the regression analysis in the step preceding Stroop inhibition – this allows for the element of speed to be controlled for when considering Stroop interference, as it important to do so when considering age differences in the Stroop task. Three separate analyses were conducted for AV performance – overall identification, as well as performance for low ID items and high ID items.

Table 9 displays overall AV identification scores for both younger and older adults, as predicted by hearing, speed, and Stroop interference. Table 10 displays AV intelligibility scores for younger and older adults, separately for identification of low ID and high ID items as predicted by the three variables. From Table 9, it appears that Stroop interference is not significantly predictive of overall AV word identification in younger adults (1.8%), however AV word identification in older adults is significantly predicted by Stroop interference (12.3%), in addition to hearing ability. In Table 10 where the results are not parsed into identification of low and high ID items, neither hearing nor inhibitory ability is predictive of AV word identification performance of low or high ID items in younger adults. In contrast, hearing ability in older adults was observed to be predictive of both low and high ID word identification; however, inhibitory ability as measured by Stroop accounts for an additional 10.3% unique variance in predicting AV word identification of high ID words, controlling for hearing ability and processing speed. In contrast, Stroop interference was not significantly predictive of low ID word identification (2.6%). Therefore, inhibitory ability as measured by Stroop appears to be strongly predictive of high, but not low, ID word intelligibility in older adults only, while inhibitory ability is not significantly predictive of AV word intelligibility in younger adults. In addition to controlling for the effects of speed via partial correlation, individual effect sizes were computed for each individual (Faust et al., 1999), such that the effect of processing speed was removed for each individual based on the difference between his/her mean RTs in the incongruent and congruent conditions, divided by the standard deviation of their combined incongruent and congruent RTs. Entering these effect sizes into the regressions analyses converged on the same pattern of data, further providing support for the effect of age-related interference independent of processing speed (see Table 11 for R^2 change and *F* values).

Table 9.

Results of Hierarchical Regression Analyses Predicting Audiovisual (AV) Identification

Variable		Coefficient	R^2	R^2 Change	*F* Change
		Younger Adults			
Step 1					
	Hearing Ability	-.106	.011		.519
Step 2					
	Hearing Ability	-.111	.036	.025	1.162
	Speed Index	-.158			
Step 3					
	Hearing Ability	-.158	.054	.018	.885
	Speed Index	-.157			
	Stroop interference	-.143			
		Older Adults			
Step 1					
	Hearing Ability	-.439	.193		11.208**
Step 2					
	Hearing Ability	-.453	.228	.035	2.110
	Speed Index	-.189			
Step 3					
	Hearing Ability	-.453	.351	.123	8.531**
	Speed Index	-.061			
	Stroop interference	-.373			

Note. Coefficients are standardized.

** $p < .01$

Table 10.

Results of Hierarchical Regression Analyses Predicting AV Identification of Low ID Words and High ID Words

	Low ID Words				High ID Words			
Variable	Coefficient	R^2	R^2 Change	F Change	Coefficient	R^2	R^2 Change	F Change
	Younger Adults							
Step 1								
Hearing Ability	-.154	.024		1.117	-.032	.001		.047
Step 2								
Hearing Ability	-.159	.046	.022	1.028	-.036	.015	.014	.624
Speed Index	-.148				-.117			
Step 3								
Hearing Ability	-.145	.047	.002	.075	-.119	.072	.058	2.739
Speed Index	-.148				-.115			
Stroop interference	.043				-.254			
	Older Adults							
Step 1								
Hearing Ability	-.274	.075		.415	-.431	.186		10.731**
Step 2								
Hearing Ability	-.245	.146	.071	.438	-.440	.200	.014	.798
Speed Index	-.276				-.118			
Step 3								
Hearing Ability	-.345	.172	.026	.656	-.440	.303	.103	6.650*
Speed Index	-.315				.001			
Stroop interference	.174				-.342			

Note. Coefficients are standardized. ID = Intersection Density.

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

Table 11.

Results of Hierarchical Regressions Models Using Stroop Effect Sizes as Predictor Variable

Dependent Variable	R^2 Change	*F*
Older Adults		
AV – Low ID Words	.014	0.833
AV – High ID Words	.093	5.689*
Younger Adults		
AV – Low ID Words	.034	1.610
AV – High ID Words	.008	0.368

* $p < .05$

Predictors of Visual Enhancement (VE)

A similar series of hierarchical regressions were conducted in order to investigate the independent contributions of hearing, speed, and Stroop inhibition on VE for each age group. Again, three separate analyses were conducted for AV identification scores – overall VE, as well as VE of low ID items and high ID items.

Table 12 displays overall VE scores for younger and older adults, as predicted by hearing, speed, and Stroop interference. Table 13 displays VE scores from younger and older adults, separated by identification of low and high ID items as predicted by the three variables. In a similar pattern to the AV word identification data, inhibitory ability in younger adults is not significantly predictive of VE. In older adults however, inhibitory ability significantly accounts for 13.9% of the variance in accounting for VE, in addition to hearing ability. In Table 13 where the data is separated into VE of low and high ID items, Stroop inhibitory ability is still not predictive of VE for either low or high ID items in younger adults. In a similar pattern to the AV results, Stroop interference significantly accounts for 16.7% of the variance in predicting VE of high ID items, but not low ID items, in older adults.

Thus, in a similar pattern to AV word identification, the benefit obtained from additional visual information in high ID word identification in older adults is strongly predicted by inhibitory ability, while there is no inhibitory contribution to VE of low ID words.

Table 12.

Results of Hierarchical Regression Analyses Predicting VE Scores for Words in Audiovisual (AV) Presentation

Variable		Coefficient	R^2	R^2 Change	*F* Change
		Younger Adults			
Step 1					
	Hearing Ability	-.106	.011		.518
Step 2					
	Hearing Ability	-.111	.044	.033	1.554
	Speed Index	-.182			
Step 3					
	Hearing Ability	-.146	.054	.010	.472
	Speed Index	-.181			
	Stroop interference	-.107			
		Older Adults			
Step 1					
	Hearing Ability	-.292	.085		4.388*
Step 2					
	Hearing Ability	-.300	.097	.012	.604
	Speed Index	-.109			
Step 3					
	Hearing Ability	-.301	.236	.139	8.190**
	Speed Index	.027			
	Stroop interference	-.397			

Note. Coefficients are standardized. VE = Visual Enhancement.

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

Table 13.

Results of Hierarchical Regression Analyses Predicting VE of Low ID Words and High ID Words in Audiovisual (AV) Presentation

Variable	Low ID Words				High ID Words			
	Coefficient	R^2	R^2 Change	*F* Change	Coefficient	R^2	R^2 Change	*F* Change
Younger Adults								
Step 1								
Hearing Ability	-.002	.000		.000	-.130	.017		.792
Step 2								
Hearing Ability	-.006	.009	.009	.392	-.136	.047	.030	1.439
Speed Index	-.093				-.175			
Step 3								
Hearing Ability	.019	.014	.005	.233	-.197	.079	.031	1.504
Speed Index	-.093				-.173			
Stroop interference	.076				-.188			
Older Adults								
Step 1								
Hearing Ability	-.203	.041		2.018	-.299	.090		4.627*
Step 2								
Hearing Ability	-.218	.083	.042	2.096	-.299	.090	.000	.003
Speed Index	-.205				.007			
Step 3								
Hearing Ability	-.219	.140	.057	2.993	-.300	.257	.167	10.102**
Speed Index	-.118				.156			
Stroop interference	-.255				-.435			

Note. Coefficients are standardized. VE = Visual Enhancement; ID = Intersection Density.

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

Discussion

The goal of the current study was to assess the role of inhibitory control on audiovisual (AV) spoken-word perception, independent of age-related sensory declines and processing speed. Measures of hearing ability, processing speed, inhibitory control, and word intelligibility were obtained from younger and older participants. The basic pattern of results revealed that younger adults demonstrated superior identification performance in AV presentation, compared to older adults, consistent with previous findings (e.g. Sumby and Pollack, 1954; Erber, 1969; Sommers et al., 2005). Low intersection density (ID) words, as defined by words with few neighbours that are both auditory and visual neighbours were also identified at a higher rate than high ID words, i.e. words with a high number of neighbours. Critically, it was observed that inhibitory ability as measured by the Stroop task was observed to be correlated with high ID word identification, but only for older adults.

A more direct investigation of the relationship between age-related inhibitory control and AV word intelligibility was obtained through hierarchical regression analyses, which revealed that a significant proportion of the variance in AV word identification of high ID words in older adults was accounted for by Stroop interference, independent of hearing loss or processing speed. This relationship did not reach significance for low ID words, nor in younger adults for either high or low ID words. These findings largely conform to an inhibitory deficit account of age-related declines in word intelligibility which exist independent of sensory impairments or general slowing. In addition to the significant independent contributions of hearing ability and processing speed which have been implicated in previous speech perception research (e.g. Humes & Roberts, 1990; Cienkowski & Carney, 2002; van Rooij & Plomp, 1990), these findings are consistent with an existing body of research that describes the effect of age-related declines in cognition on speech perception.

Aging and the AV-NAM

The results of this study are consistent with previous research and predictions of the original NAM (Luce & Pisoni, 1998; Sommers, 1996; Vitevitch & Luce, 1999; Sommers & Danielson, 1999; Taler et al., 2010): words with many lexical neighbours generate greater activation and subsequently greater competition in spoken-word recognition, resulting in greater inhibitory demands which are necessary to isolate the target word. Although explicit claims about inhibition *per se* were not made in the original conception of the NAM, Sommers (1996) proposed additional tenets to the NAM which account for inhibition-driven competition amongst lexical items. Specifically, Sommers proposed that during word recognition, active inhibition is the mechanism by which the activation of irrelevant lexical candidates must be lowered.

The validity of this proposal is supported by the present results considered within the context of the adapted AV-NAM which demonstrated intelligibility differences for low ID and high ID words. By definition, high ID words have many lexical neighbours and therefore many lexical items competing for activation; consequently, inhibitory demands are relatively high, in contrast to words with fewer lexical competitors. Indeed, the finding that Stroop interference was only predictive of high ID words and only for older adults supports this activation-competition framework of the AV-NAM, and specifically, the inhibitory mechanism which mediates AV speech intelligibility.

The age-based interaction of performance and ID also speaks to the inhibitory mechanisms of the AV-NAM, as well as to age-related decreases in inhibitory function. The results are also consistent with previous research suggesting inhibitory declines in older adults (e.g. Hasher & Zacks, 1988; Hasher et al., 1991; Kramer, Humphrey, Larish, & Logan, 1994) – age-related reductions in inhibitory abilities directly and independently contribute to intelligibility difficulties, particularly of words with many lexical neighbours, as was observed in Table 10. From these results, it would appear that as a result of impaired inhibitory function as measured by increased Stroop interference, older adults are less able to meet the increased inhibitory demands required for high ID word identification, and intelligibility consequently suffers. This also demonstrates individual differences within older adults with regard to those with either intact or diminished inhibitory ability, who may be described as ‘good’ and ‘poor’ inhibitors, respectively. Those who are good inhibitors, i.e. those who experience less Stroop interference, will demonstrate higher AV identification of high ID words; in contrast, ‘poor’ inhibitors, i.e. those who experience more Stroop interference, will exhibit poorer high ID word identification performance.

These findings also extend the results of Sommers (1996) who found that younger and older adults performed similarly in identification of words with few lexical neighbours, but diverged significantly in identification of words with many neighbours (note, however, that this study examined A-only lexical difficulty). If the difference between younger and older adults were solely due to age-related hearing impairments – which were observed to be significantly correlated with both low ID and high ID word identification (Table 8) – then the difference between the age groups should be roughly similar for both easy and hard words. However, the greater difference between the age groups for high ID words indicates that audibility alone is not the determining factor.

Younger adults also exhibited higher VE scores compared to older adults, indicating that they are benefiting to a greater degree from an additional visual signal than older adults, even when controlling for hearing ability. This is the first study to demonstrate a significant differences in VE between the two age groups, as previous studies examining VE have found directional, but not statistically significant, differences between younger/older adults (Sommers et al., 2005) and normal-hearing/hearing-impaired older adults (Tye-Murray et al., 2007). It is worth noting, however, that these previous studies used different sets of speech stimuli; further investigations of comparing VE between younger and older adults would help to specify the conditions under which the benefits of VE occur for different age groups and/or materials.

Stroop interference also predicted obtained visual benefit (i.e. VE) for intelligibility of high ID words in older adults, but not younger adults. The pattern of results echo those in AV word intelligibility, such that inhibitory ability strongly predicts overall VE for older but not younger adults, and when separated into easy and hard words, is only predictive of VE of high ID words. These findings suggest that older adults' ability to benefit from the addition of a visual signal in AV word intelligibility is at least partially determined by inhibitory ability. Considered within the context of the AV-NAM, this appears to make sense. This adapted model supposes that during AV word presentation, the unimodal neighbourhoods are perceived and integrated simultaneously into a bimodal percept, thereby eliminating unimodal neighbours as potential competitors. However, impairments in the inhibitory mechanism might lead to inefficient integration of the sensory neighbourhoods, resulting in the unimodal candidates remaining at least somewhat active, thereby preventing a clearly delineated intersection density. Therefore, age-related declines in inhibitory ability may lead to unimodal neighbours (in either A-only or V-only) incorrectly remaining active, thereby increasing the pool of competitors to the target word.

Essentially, older adults may not be benefitting from additional visual information because they are unable to integrate the two sensory neighbourhoods and inhibit those unimodal candidates which are irrelevant.

However, it is important to note that general claims about inhibitory function must be interpreted with caution. Age-related interference effects were only observed in two of the tasks (Stroop and DRM), and only the Stroop task was correlated with AV word identification. Therefore, it is possible that the three tasks may be measuring different aspects of the same construct (inhibition). Sommers and Danielson (1999) also found that Stroop was predictive of auditory word identification of hard but not easy words, although a different measure of Stroop interference was used (the RTs from the incongruent condition, as opposed to the present study which measured Stroop interference as a differences score between incongruent and congruent RTs). There thus seems to be a common mechanism between AV speech intelligibility and the Stroop task in particular. According to the Hasher and Zacks (1988) conception of inhibition, the Stroop task would be considered a *restraint* function, in that it involves a strongly activated but incorrect response (i.e. word naming) which must be *restrained* in order to proceed with the correct response (colour naming). This idea can be considered similar to that which occurs with lexical activation in the AV-NAM – during AV word presentation, strongly activated but irrelevant competitors must be inhibited (i.e. restrained) in order to activate the target item. Accordingly, when words possess a greater number of lexical competitors, more items must be inhibited, thus paralleling its similarity with the restraint mechanism of the Stroop task. Although a full discussion of the inhibition construct is beyond the scope of this thesis, it is clear that some aspect of age-related increases in susceptibility to interference is involved in the obtained findings.

Burke (1997) has contested the involvement of age-related inhibition in hard word recognition. Burke argues that the inhibitory relationship between the target item and competitors is bidirectional, such that inhibition of the competitors would also result in inhibition of the target; in the case of older adults with impaired inhibitory abilities, there would be less inhibition of the competitors *as well as* the target. Subsequently, the effect of age-related declines in inhibition would be counteracted by reduced inhibition of targets by their neighbours. The difficulty with this argument, as Sommers and Danielson (1999) point out, is that it does not consider *overall* activation and inhibition, merely relative levels. Overall activation levels are accompanied by proportional increases in inhibitory effects, which are altered with additional sensory evidence to identify to the *target* word, not to the competitors. As this additional information is received, it is thus the target-to-competitor inhibitory relationship which has greater effects than the reverse situation. As Sommers and Danielson note, "it is this asymmetry in the inhibitory influences of targets and competitors that provides a mechanism whereby age-related inhibitory deficits could impair recognition of hard words" (pg 468).

Finally, the Pearson correlations in Table 8 showed that processing speed was highly correlated with AV performance. Although this relationship disappeared in the regression analyses when Stroop interference was entered as a predictor variable, it suggests that speed may be an important factor to consider in speech perception. Contrary to the current findings, Sommers (1998) and Sommers and Danielson (1999) both observed that tasks measuring speed of processing did not significantly predict (auditory) spoken-word recognition. Nevertheless, the influence of speed is an important one to consider, particularly when assessing age-related differences. General slowing of cognitive processing has been observed in a number of lexical tasks (Myerson, Ferraro, Hale, & Lima, 1992), and one could imagine within the framework of the NAM that general slowing could result in more items remaining active for a longer time, especially for high ID words, because of delays in eliminating more competitors. Naming times were not obtained in the present study, and so further investigations may be necessary to conclude whether individual differences in processing speed contribute to individual differences in AV speech intelligibility.

Future Research

One theoretical aspect of the AV advantage involves the reduction in competing lexical candidates, although we have no direct evidence for this from the current study. The mechanistic explanation for the AV advantage arises from the fact that the AV intersection density eliminates a large number of unimodal competitors, and so inhibitory ability should predict a smaller amount of variance in explaining AV word identification (because there are fewer candidates) than it should for A-only word identification (because of more candidates). This was unable to be addressed in the current study because all stimuli were selected based on AV characteristics (i.e. those based on intersection density values), and even examining a select number of A-only presented stimuli based on extreme density values would be an unbalanced and unfair comparison to AV-presented stimuli. Future investigations into the AV advantage would benefit from examining the differences between intelligibility of words with high- and low-density A neighbours and words high- and low-density AV neighbours in order to determine whether a decrease in lexical competitors alters the amount of variance explained by inhibitory ability.

Although word-based stimuli are necessary to examine the phenomenon of speech intelligibility, the words were placed within identical carrier phrases. Placing the words within more naturalistic sentences may allow for a more realistic investigation of how spoken-word recognition works. For example, Sommers and Danielson (1999) placed easy and hard words within the context of high- and low-predictability sentences, from which target word identification scores were easily obtained. As this study was only the first to examine the role of inhibition in easy/hard AV word intelligibility, future studies would do well to examine lexical difficulty effects within the context of more naturalistic speech.

Clinical Implications and Conclusions

The clinical implications of this work concerning the interplay between aging, inhibitory function, and speech perception largely lie in the practice of aural rehabilitation. It is a well-established finding that age-related impairments in hearing ability affect the intelligibility and perception of spoken language (e.g. CHABA, 1988; see Introduction). Much of the previous work examining aural rehabilitation strategies in clinical populations (i.e. individuals with hearing impairments, older adults) has focussed on auditory processing. To this end, hearing aids have been implemented in clinical settings, primarily targeting the higher frequency amplification (Hogan & Turner, 1998; Turner & Cummings, 1999) and sound localisation (Byrne & Noble, 1998; Seeber, Baumann, & Fastl, 2004). However, the success of hearing aids has been mixed. While some studies have reported some benefits of aided listening (Gatehouse, Naylor, & Elberling, 2006) and perception in modulated noise (Gatehouse, Naylor, & Elberling, 2003), the evidence to suggest robust effectiveness of hearing aids – particularly over a long-term period – remains largely limited (e.g. Mulrow, Tuley, & Aguilar, 1992; Surr, Cord, & Walden, 1998).

The findings from the present study, along with results of previous studies (e.g. CHABA, 1988; Tun, Wingfield, Stine, & Mecsas, 1992) suggest that a focus on auditory processing alone may limit the scope of clinical utility. Instead, an alternative strategy would be to incorporate both cognitive *and* sensory factors in an integrated approach to rehabilitation. In the present study, both factors clearly demonstrated independent contributions to speech perception in older adults, suggesting that it is a combination of both processes which cause impairments to perceiving speech in noise. This strategy is especially relevant in the context of the NAM, which indeed makes specific claims about the interplay between low-level sensory mechanisms and higher-level cognitive mechanisms. Luce and Pisoni (1998) specifically state that, “when evaluating hearing-impaired listeners, the [NAM] model emphasizes that simple tests of speech pattern discrimination and phonetic feature discrimination will grossly underestimate the complex task that faces hearing impaired and normal listeners in understanding spoken-words in naturalistic settings” (pg 41). This statement directly speaks to the interaction of sensory and cognitive mechanisms, such that assessments which solely target acoustic mechanisms will be at a clinical disadvantage, due to the exclusion of the punitive effect of cognitive mechanisms. In some cases, these cognitive effects may be detrimental to lexical discrimination, as in the case of inhibitory decline. In other instances however, the wide range of cognitive mechanisms present in the speech system may actually promote compensatory strategies in lieu of other declines. Indeed, Sommers and Danielson (1999) observed that older listeners are able to use additional semantic information provided by high predictability context in order to compensate for inhibitory declines.

The current study illustrates the importance of considering spoken-word recognition as a set of dynamic perceptual processes, as delineated by the NAM and adapted to a bimodal framework of the model called the AV-NAM, which will be of great utility in proceeding forward with strategies to improve speech perception in older adults. The results provide evidence as to inhibitory mechanisms underlying some aspects of speech perception, which also interact with sensory processes. A better understanding of this interaction will inform future work examining both the theoretical nature of the relationship, as well as training strategies for older adults.

REFERENCES

Arbuckle, T. Y., & Gold, D. P. (1993). Aging, inhibition, and verbosity. *Journal of Gerontology*, *48*(5), 225-232.

Auer, E. T. (2002). The influence of the lexicon on speech read word recognition: Contrasting segmental and lexical distinctiveness. *Psychonomic Bulletin & Review*, *9*(2), 341-347.

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger III, H. L., McDermott, K. B., & Yerys, B. E. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology*, *16*(3-5), 361-384.

Balota, D. A., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., & Treiman, R. (2002). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445-459.

Behne, D., Wang, Y., Alm, M., Arntsen, I., Eg, R., & Valsø, A. (2007). Changes in audio-visual speech perception during adulthood. In *International Conference on Auditory-Visual Speech Processing*.

Berg, T., & Schade, U. (1992). The role of inhibition in a spreading-activation model of language production. I. The psycholinguistic perspective. *Journal of Psycholinguistic Research*, *21*(6), 405-434.

Blamey, P. J., Cowan, R. S., Alcantara, J. I., Whitford, L. A., & Clark, G. M. (1989). Speech perception using combinations of auditory, visual, and tactile information. *Journal of Rehabilitation Research and Development*, *26*(1), 15.

Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *The Journal of the Acoustical Society of America*, *106*, 2074.

Burke, D. M. (1997). Language, aging, and inhibitory deficits: Evaluation of a theory. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *52*(6), P254-P264.

Byrne, D., & Noble, W. (1998). Optimizing sound localization with hearing aids. *Trends in Amplification*, *3*(2), 51-73.

Cedrus Corporation. (1992). *SuperLab*, Version 4.5.4. Silver Spring: Cedrus Corporation.

CHABA (Committee on Hearing, Bioacoustics and Biomechanics). (1988). Speech understanding and aging. *Journal of the Acoustical Society of America, 83,* 859-895.

Chen, J., Hale, S., & Myerson, J. (2007). Predicting the size of individual and group differences on speeded cognitive tasks . *Psychonomic bulletin & review*, *14*(3), 534-541.

Chen, J., Myerson, J., Hale, S., & Simon, A. (2000). Behavioral evidence for brain-based ability factors in visuospatial information processing. *Neuropsychologia*, *38*(4), 380-387.

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological review*, *119*(2), 417.

Cienkowski, K. M., & Carney, A. E. (2002). Auditory-visual speech perception and aging. *Ear and hearing*, *23*(5), 439-449.

Conway, M. A., Harries, K., Noyes, J., Racsma'ny, M., & Frankish, C. R. (2000). The disruption and dissolution of directed forgetting: Inhibitory control of memory. *Journal of Memory and Language*, *43*(3), 409-430.

Craik, F. I., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 474.

Dirks, D. D., Takayana, S., & Moshfegh, A. (2001). Effects of lexical factors on word recognition among normal-hearing and hearing-impaired listeners. *Journal of the American Academy of Audiology*, *12*(5), 233.

Dufour, S., & Peereman, R. (2003). Inhibitory priming effects in auditory word recognition: When the target's competitors conflict with the prime word. *Cognition*, *88*(3), 33-44.

Earles, J. L., Connor, L. T., Frieske, D., Park, D. C., Smith, A. D., & Zwahr, M. (1997). Age differences in inhibition: Possible causes and consequences. *Aging, Neuropsychology, and Cognition*, *4*(1), 45-57.

Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech, Language and Hearing Research*, *12*(2), 423.

Eriksen, B.A., & Eriksen, C.W. (1974). Effects of noise letters upon the identification of a target in a nonsearch task. Perception & Psychophysics, 16 , 143–149.

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: implications for group differences in response latency. *Psychological bulletin*, *125*(6), 777.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189-198.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and language*, *45*(3), 468-492.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Lawrence Erlbaum.

Gatehouse, S., Naylor, G., & Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *International Journal of Audiology*, *42*, S77-S85.

Gatehouse, S., Naylor, G., & Elberling, C. (2006). Linear and nonlinear hearing aid fittings-1. Patterns of benefit. *International Journal of Audiology*, *45*(3), 130-152.

Girin, L., Schwartz, J. L., & Feng, G. (2001). Audio-visual enhancement of speech in noise. *The Journal of the Acoustical Society of America*, *109*, 3007.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, *28*(5), 501-518.

Grant, K. W., & Seitz, P. F. (1998). Measures of auditory–visual integration in nonsense syllables and sentences. *The Journal of the Acoustical Society of America*, *104*, 2438.

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.

Green, K. P., Kuhl, P. K., Meltzoff, A. N., & Stevens, E. B. (1991). Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics*, *50*(6), 524-536.

Green, K. P., & Norrix, L. W. (1997). Acoustic cues to place of articulation and the McGurk effect: The role of release bursts, aspiration, and formant transitions. *Journal of Speech, Language and Hearing Research*, *40*(3), 646.

Green, K. P. (1998). The use of auditory and visual information during phonetic processing: Implications for theories of speech perception. *Hearing by eye*, *2*, 3-25.

Greenberg, J. H., & Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, *20*, 157–177.

Halling, D. C., & Humes, L. E. (2000). Factors affecting the recognition of reverberant speech by elderly listeners. *Journal of Speech, Language and Hearing Research*, *43*(2), 414.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of experimental psychology: Learning, memory, and cognition*, *17*(1), 163.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation,* (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.

Hay-McCutcheon, M. J., Pisoni, D. B., & Kirk, K. I. (2005). Audiovisual speech perception in elderly cochlear implant recipients. *The Laryngoscope*, *115*(10), 1887-1894.

Hogan, C. A., & Turner, C. W. (1998). High-frequency audibility: Benefits for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *104*, 432.

Honnell, S., Dancer, J. & Gentry, B. (1991). Age and speechreading performance in relation to percent correct, eyeblinks, and written responses. *Volta Review*, *93*(4), 207-13.

Houx, P. J., Jolles, J., & Vreeling, F. W. (1993). Stroop interference: aging effects assessed with the Stroop Color-Word Test. *Experimental aging research*, *19*(3), 209-224.

Hull, R. H. (1995). *Hearing in aging*. Singular Publishing Group.

Humes, L. E., Watson, B. U., Christensen, L. A., Cokely, C. G., Halling, D. C., & Lee, L. (1994). Factors associated with individual differences in clinical measures of speech recognition among the elderly. *Journal of Speech, Language and Hearing Research*, *37*(2), 465.

Humes, L. E., & Roberts, L. (1990). Speech-recognition difficulties of the hearing-impaired elderly: The contributions of audibility. *Journal of Speech, Language and Hearing Research*, *33*(4), 726.

Jackson, P. L. (1988). The theoretical minimal unit for visual speech perception: Visemes and coarticulation. *The Volta Review*.

Jerger, S., Stout, G., Kent, M., Albritton, E., Loiselle, L., Blondeau, R., & Jorgenson, S. (1993). Auditory Stroop effects in children with hearing impairment. *Journal of Speech, Language and Hearing Research*, *36*(5), 1083.

Kasper, R. (1978). Eye problems of the aged. In W. Reichel (Ed.), *Clinical aspects of aging* (pp. 391-402). Baltimore, MD: Williams & Wilkins.

Kramer, A. F., Humphrey, D. G., Larish, J. F., & Logan, G. D. (1994). Aging and inhibition: beyond a unitary view of inhibitory processing in attention. *Psychology and aging*, *9*(4), 491.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, *38*(3), 245-294.

Liu, I. Y., White, L., & LaCroix, A. Z. (1989). The association of age-related macular degeneration and lens opacities in the aged. *American journal of public health, 79*(6), 765-769.

Liu, X. Z., & Yan, D. (2007). Ageing and hearing loss. *The Journal of pathology*, *211*(2), 188-197.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and hearing, 19*(1), 1.

MacLeod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *British Journal of Audiology*, *21*(2), 131-141.
masking, and speech understanding in noise in the elderly. J.Speech Hear. Res. *35*, 1410–1421.

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science, 31*(1), 133-156.

Massaro, D. W., & Light, J. (2004). Using visible speech to train perception and production of speech for individuals with hearing loss. *Journal of Speech, Language and Hearing Research, 47*(2), 304.

Mattys, S. L., Bernstein, L. E., & Auer, E. T. (2002). Stimulus-based lexical distinctiveness as a general word-recognition mechanism. *Perception & psychophysics*, *64*(4), 667-679.

McDowd, J. M. (1997). Inhibition in attention and aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *52*(6), 265-273.

McDowd, J. M., & Filion, D. L. (1992). Aging, selective attention, and inhibitory processes: a psychophysiological approach. *Psychology and aging, 7*(1), 65.

McDowd, J. M., & Oseas-Kreger, D. M. (1991). Aging, inhibitory processes, and negative priming. *Journal of Gerontology*, *46*(6), P340-P345.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review*, *88*(5), 375.

McGurk, H., & MacDonald, J. Hearing lips and seeing voices. *Nature,* 1976, 264, 746-748.

Morrell, C. H., Gordon-Salant, S., Pearson, J. D., Brant, L. J., & Fozard, J. L. (1996). Age-and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. *Journal of the Acoustical Society of America*, *100*(4), 1949-1970.

Mulrow, C. D., Tuley, M. R., & Aguilar, C. (1992). Correlates of successful hearing aid use in older adults. *Ear and Hearing*, *13*(2), 108-113.

Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. *Psychology and Aging,* 7, 257-270.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234.

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International journal of audiology*, *42*(S2), 11-16.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America, 97,* 593.

Pichora-Fuller, M. K. (2003). Cognitive aging and auditory information processing. *International Journal of Audiology, 42*, 2S26-2S32.

Rabinowitz, W. M., Eddington, D. K., Delhorne, L. A., & Cuneo, P. A. (1992). Relations among different measures of speech reception in subjects using a cochlear implant. *The Journal of the Acoustical Society of America*, *92,* 1869.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803.

Rosenblum, L. D., Yakel, D. A., & Green, K. P. (2000). Face and mouth inversion effects on visual and audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 806.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological review*, *103*(3), 403-427.

Schneider, B. A., & Pichora-Fuller, M. K. (2000). Age-related changes in temporal processing: implications for speech perception. *Seminars in hearing*, *22*(3), 227-240.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime (Version 2.0. *Pittsburgh, PA: Psychology Software Tools Inc*.

Seeber, B. U., Baumann, U., & Fastl, H. (2004). Localization ability with bimodal hearing aids and bilateral cochlear implants. *The Journal of the Acoustical Society of America*, *116*, 1698.

Sommers, M. S. (1996). The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychology and aging*, *11*(2), 333.

Sommers, M. S. (1998). Spoken word recognition in individuals with dementia of the Alzheimer's type: Changes in talker normalization and lexical discrimination. *Psychology and aging*, *13*(4), 631.

Sommers, M. S., & Lewis, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory and Language*, *40*(1), 83-108.

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*(3), 458.

Sommers, M. S., 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*(3), 263–75.

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

Summerfield, Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In B. Dodd and R. Campbell (Eds.), *Hearing by Eye: The Psychology of Lip-reading*, Lawrence Erlbaum, London, 3-51.

Surr, R. K., Cord, M. T., & Walden, B. E. (1998). Long-term versus short-term hearing aid benefit. *Journal of the American Academy of Audiology*, *9*(3), 165-171.

Takayanagi, S., Dirks, D. D., & Moshfegh, A. (2002). Lexical and talker effects on word recognition among native and non-native listeners with normal and impaired hearing. *Journal of Speech, Language and Hearing Research*, *45*(3), 585.

Taler, V., Aaron, G. P., Steinmetz, L. G., & Pisoni, D. B. (2010). Lexical neighborhood density effects on spoken word recognition and production in healthy aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *65*(5), 551-560.

Tun, P. A., Wingfield, A., Stine, E. A., & Mecsas, C. (1992). Rapid speech processing and divided attention: Processing rate versus processing resources as an explanation of age effects. *Psychology and Aging*, *7*(4), 546.

Turner, C. W., & Cummings, K. J. (1999). Speech audibility for listeners with high-frequency hearing loss. *American Journal of Audiology*, *8*(1), 47.

Tye-Murray, N., Sommers, M. S., & Spehar, B. (2007). Audiovisual integration and lipreading abilities of older adults with normal and impaired hearing. *Ear and hearing*, *28*(5), 656-668.

van Rooij, J. C., & Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. *The Journal of the Acoustical Society of America*, *87*, S25.

Verbruggen, F., Liefooghe, B., & Vandierendonck, A. (2004). The interaction between stop signal inhibition and distractor interference in the flanker and Stroop task. *Acta Psychologica*.

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *68*(1), 306-311.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*(3), 374-408.

Walden, B. E., Busacco, D. A., & Montgomery, A. A. (1993). Benefit from visual cues in auditory-visual speech recognition by middle-aged and elderly persons. *Journal of Speech, Language and Hearing Research*, *36*(2), 431.

Wechsler, D. (1955) Wechsler Adult Intelligence Scale Manual. New York: Psychological Corporation.

West, R., & Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology, 37*(2), 179-189.

Wingfield, A., Alexander, A. H., & Cavigelli, S. (1994). Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging. *Language and Speech*, *37*(3), 221-235.

Wurm, L. H., & Samuel, A. G. (1997). Lexical inhibition and attentional allocation during speech perception: Evidence from phoneme monitoring. *Journal of Memory and Language*, *36*(2), 165-187.

Zacks, R. T., Radvansky, G., & Hasher, L. (1996). Studies of directed forgetting in older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 143.

APPENDIX

A1

Audiovisual (AV) Stimulus List

Low Intersection Density Neighbourhood Words			High Intersection Density Neighbourhood Words		
shimmer	paid	audio	thorn	groan	lure
dance	daily	tape	ball	lows	ardour
stead	doily	cougar	sore	boar	warm
step	acme	sled	rail	torque	fort
better	honoured	valley	stoke	line	roar
brush	prod	kids	rain	rink	toll
dome	seeker	shaggy	where	score	sink
crumb	bones	shriek	core	filly	course
three	shoddy	tummy	sting	lower	shawl
outdo	cheek	moored	fork	fuse	round
office	little	hotter	dawn	root	more
mayor	glass	blush	airy	foal	bing
dream	live	cad	corn	while	saying
gas	nettle	boost	store	talk	coarse
pain	adding	hit	caught	ruse	wrought
gent	fizzle	god	load	pall	stone
pang	third	resin	wing	shore	force
person	slime	silo	offer	wick	tore
team	alloys	cob	fore	sync	yore
mud	van	dubbed	pour	taught	mark
attire	doing	peg	forge	form	fall
big	asset	much	rune	clone	wall
duo	base	manor	morn	gore	lawn
sighed	glove	ember	ward	ford	slow
nozzle	bin	void	mall	door	war
child	better	cheap	caller	tall	firm
zooms	kit	brass	ramp	fawn	thought
girl	judo	vase	bring	foul	poor
chateau	season	bottle	nor	sloan	lure
sheets	aside	drop	will	art	ardour
image	people		walk	cough	

A2

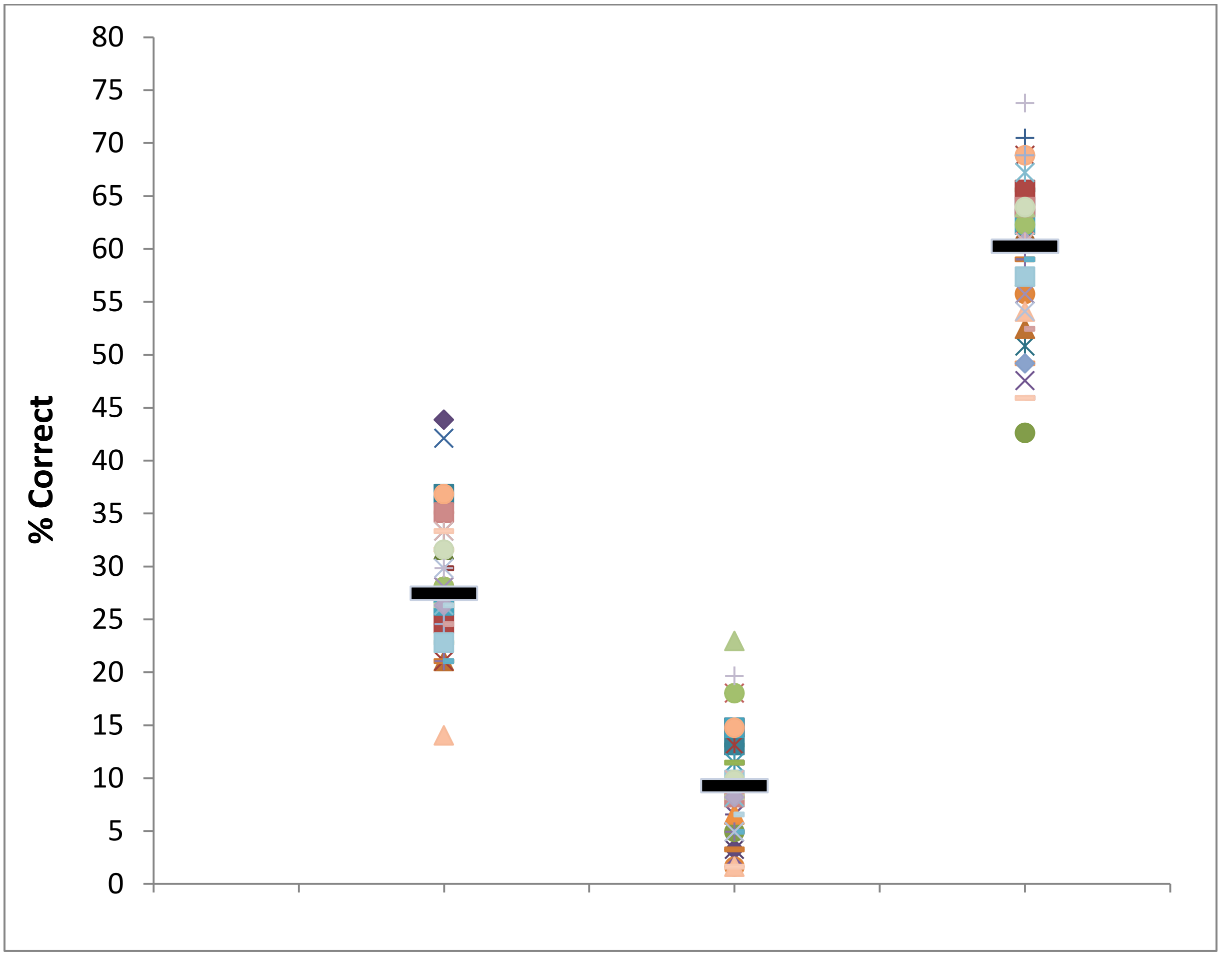

Figure I. A plot displaying the range of word recognition performance of younger adults across A-only, V-only, and AV conditions, depicted respectively across the horizontal axis. The black horizontal bar represents the mean within each condition.

A3

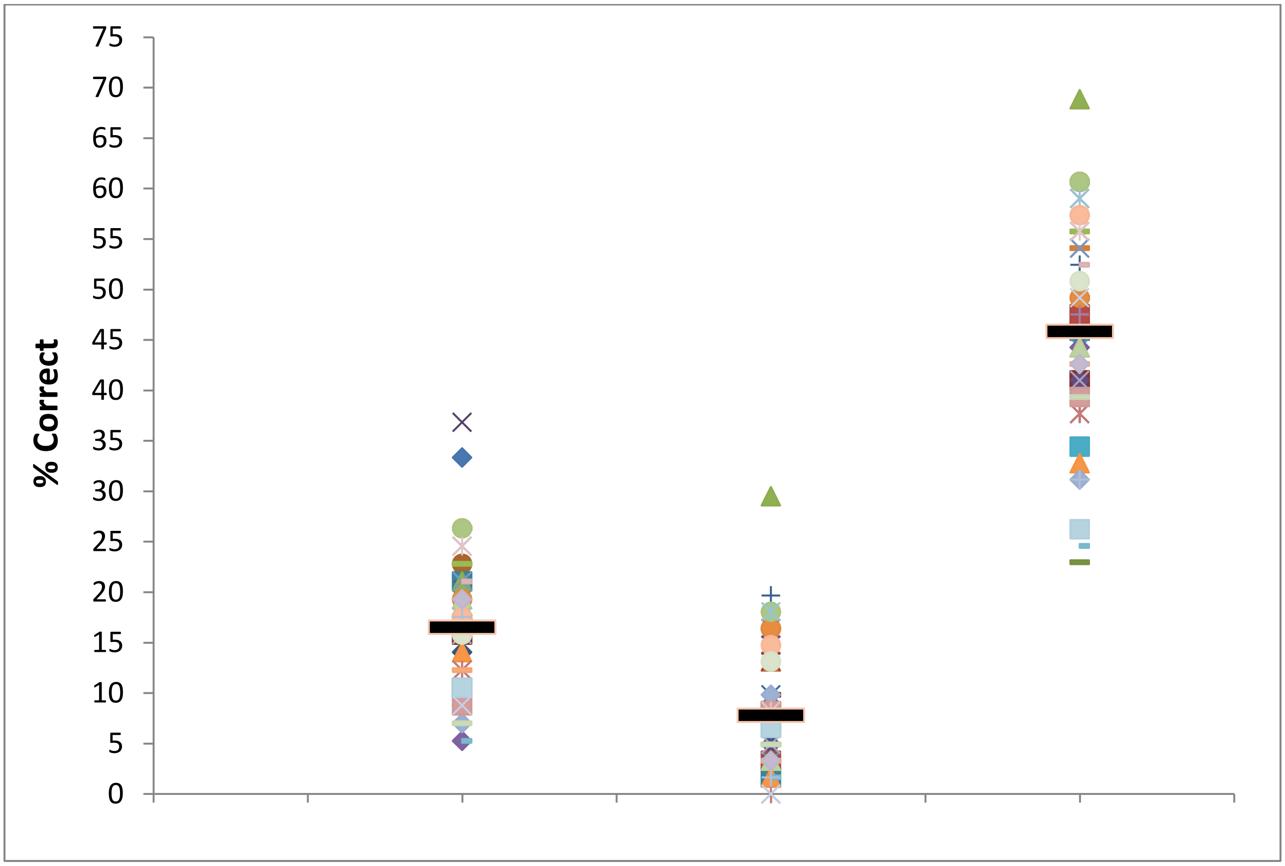

Figure II. A scatterplot displaying the range of word identification performance of older adults across A-only, V-only, and AV conditions, depicted respectively across the horizontal axis. The black horizontal bar represents the mean within each condition.

A4.

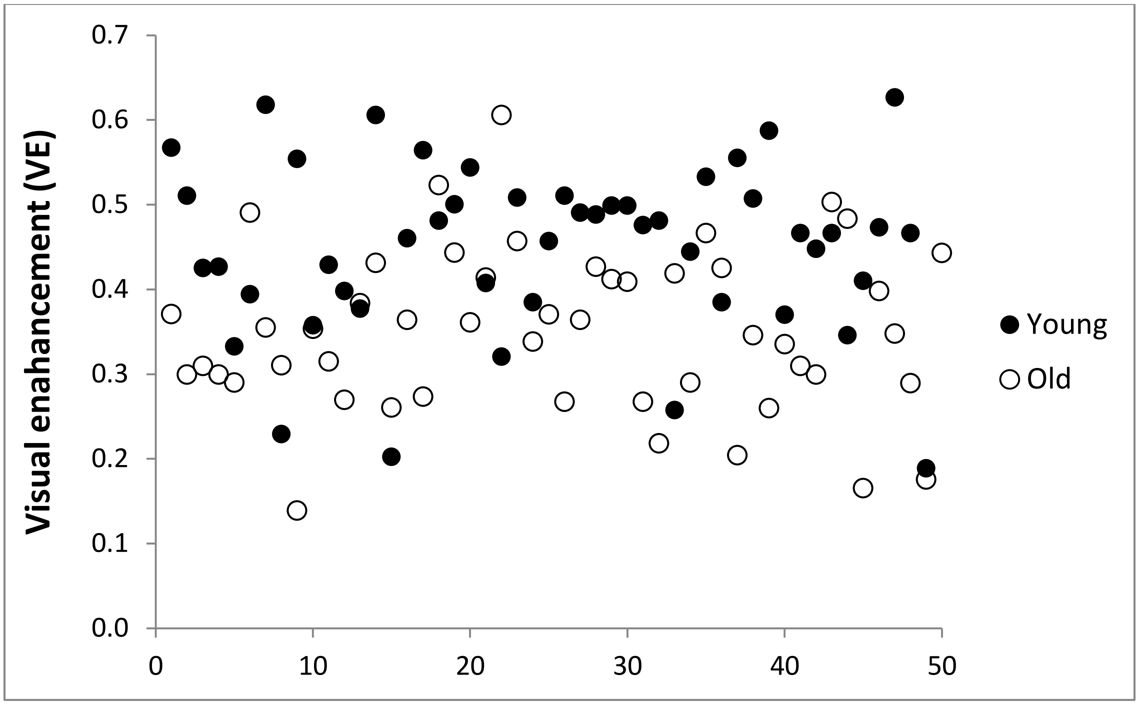

Figure III. A scatterplot depicting the range of visual enhancement (VE) scores obtained by younger (filled circles) and older (open circles) adults. The horizontal axis values represent each item.