Research Article

Lexical Influences on Errors in Masked Speech Perception in Younger, Middle-Aged, and Older Adults

Alexandra Jesse^a and Karen S. Helfer^b

Purpose: In situations with a competing talker, lexical properties of words in both streams affect the recognition of words in the to-be-attended target stream. In this study, we tested whether these lexical properties also influence the type of errors made by listeners across the adult life span.

Method: Errors from a corpus collected by Helfer and Jesse (2015) were categorized as phonologically similar to words in the target and/or masker streams. Younger, middle-aged, and older listeners had produced these errors when trying to identify key words from a target stream while ignoring a single-talker masker. Neighborhood density and lexical frequency of target words and masker words had been manipulated independently.

Results: Lexical properties of target words influenced all types of errors. With higher frequency maskers, the probability

of responding with a masker word increased and the phonological influence of target words decreased. Lower levels of lexical competition for maskers increased the probability that listeners reported a word phonologically related to both masker and target words. The influence of masker words increased across the adult life span, as evidenced by phonological intrusions into responses and the temporary failure in selectively attending to the target stream. The effects of lexical properties on error patterns, however, were consistent across age groups.

Conclusions: The ease of recognition of words in both attended and unattended speech influences the breakdown of speech perception. These influences remain robust across the adult life span.

Many of our conversations take place in noisy listening situations (e.g., with interference produced by other talkers, air-conditioning units, traffic, a dog barking). Over the course of the adult life span, listening situations with a particular type of noise—a single competing talker—become especially difficult (e.g., Duquesnoy, 1983; Helfer & Freyman, 2014; Tun, O’Kane, & Wingfield, 2002; Tun & Wingfield, 1999). Recognizing speech in situations with any type of noise can be challenging due to energetic masking, the acoustic blending of speech and noise that renders the signals less intelligible. An additional challenge in situations with competing speech as noise is the potential for

informational masking. In such situations, linguistic processing of the masker speech stream(s) can compete with the processing of the target speech stream. For example, a word that is phonologically similar to a target word (e.g., “candy” to “candle”) in the attended stream competes more strongly for recognition when it is simultaneously heard in the masker (Brouwer & Bradlow, 2016). The influence of informational masking on correctly recognizing words from the target stream depends to some extent on the degree to which competing speech is intelligible to the listener (e.g., Calandruccio, Dhar, & Bradlow, 2010; Van Engen & Bradlow, 2007). In addition, the extent of informational masking is modulated by the ease at which words in the target and masker streams can be recognized (Boulenger, Hoen, Ferragne, Pellegrino, & Meunier, 2010; Helfer & Jesse, 2015). In this study, we test whether the ease of lexical processing also predicts the pattern of breakdown in speech perception in situations with competing speech and investigate changes therein with aging.

The ease with which a spoken word is recognized is in part determined by its frequency of occurrence. Words are more reliably and more rapidly recognized, in situations with and without noise, to the degree that listeners encounter them often in their daily lives and thus often access their lexical

^aDepartment of Psychological and Brain Sciences, University of Massachusetts, Amherst

^bDepartment of Communication Disorders, University of Massachusetts, Amherst

Correspondence to Alexandra Jesse: ajesse@psych.umass.edu

Editor-in-Chief: Frederick (Erick) Gallun

Editor: David Eddins

Received March 14, 2018

Revision received July 2, 2018

Accepted October 11, 2018

https://doi.org/10.1044/2018_JSLHR-H-ASCC7-18-0091

Publisher Note: This article is part of the Special Issue: Select Papers From the 7th Aging and Speech Communication Conference.

Disclosure: The authors have declared that no competing interests existed at the time of publication.

representations (e.g., Dahan, Magnuson, & Tanenhaus, 2001; Goldinger, Luce, & Pisoni, 1989; Helfer & Jesse, 2015). All major models of spoken word recognition provide an account of this word frequency effect. For example, in mathematical models (Luce & Pisoni, 1998; Norris & McQueen, 2008), word frequency is a probability, acting as a weight. In connectionist models with localist representations, the activation level of a word representation at resting is proportionally related to the frequency of the representation's usage (Luce, Goldinger, Auer, & Vitevitch, 2000; Marslen-Wilson, 1987, 1990; McClelland & Elman, 1986). Higher frequency words have higher resting activation levels than lower frequency words and thus need to be activated less by input to be recognized. In connectionist models with distributed representations (Gaskell & Marslen-Wilson, 1997), connections between the units representing a word can be strengthened through usage, thus leading to better recognition of often-encountered words.

In addition to word frequency, the ease with which a word is recognized is also determined by how much lexical competition the input elicits, that is, by the word's phonological similarity to other word representations in the mental lexicon. Word representations stored in the listener's mental lexicon are considered as viable candidates to the degree that they are supported by the incoming acoustic information (e.g., Gaskell & Marslen-Wilson, 1997; Luce & Pisoni, 1998; McClelland & Elman, 1986; Norris & McQueen, 2008). Word representations compete for recognition by inhibiting the activation of competitors (Luce et al., 2000; McClelland & Elman, 1986; Norris, 1994; Norris, McQueen, Cutler, & Butterfield, 1997). If the acoustic information obtained from the target stream is similar to many representations rather than to just a few representations, more competition arises and the subsequent increase in inhibition renders the recognition of the target word more difficult. Phonological similarity to other lexical entries has been commonly operationalized as neighborhood density, that is, as the number of words that differ in one phoneme from a particular word (Luce, 1986). Target words are more difficult to recognize if they come from dense rather than sparse lexical neighborhoods (e.g., Luce, 1986; Luce & Pisoni, 1998; Luce, Pisoni, & Goldinger, 1990).

Recent work has shown that, in situations with competing speech, the lexical properties of words in the masker stream can modulate the accuracy of recognizing words in the target speech stream (Boulenger et al., 2010; Helfer & Jesse, 2015). In such situations with competing talkers, the frequency of masker words influenced target word recognition. Multitalker babble consisting of high-frequency masker words interfered more with younger listeners' lexical decisions on target words than multitalker babble consisting of low-frequency words (Boulenger et al., 2010). This increase in informational masking with higher levels of masker word frequency was, however, only found when the babble consisted of speech from two talkers; it was not found when the masker consisted of more than two talkers. That is, the influence of masker word frequency was only apparent in situations that allowed for masker words to be sufficiently

identified. The masker frequency effect was not observed when the number of masker talkers contributing to the babble was increased such that the linguistic content became inaccessible. However, if sufficient perceptual information about the masker word is needed for its frequency to modulate the degree of informational masking on target recognition, then this effect should also be observed in a listening situation with a single competing talker. A recent study by Helfer and Jesse (2015) nevertheless failed to observe an effect of masker frequency on accuracy of target identification in a situation with only one competing talker. In that study, target recognition accuracy by younger adults was not influenced by masker word frequency. The effect was also not found for groups of middle-aged and older listeners. The precise factors that may have caused this difference in results across studies (e.g., number of masker talkers, signal-to-noise ratios, and task) are currently undetermined.

Helfer and Jesse (2015) found, however, that the neighborhood density of the masker words affected target word recognition in a listening situation with a single competing speaker. Younger listeners were more likely to correctly recognize target words when the masker words came from dense rather than sparse lexical neighborhoods. That is, when lexical competition made the recognition of masker words more difficult, younger adults' ability to report the target words increased. This effect of masker neighborhood density on target recognition was only found when younger listeners were tested at the more favorable signal-to-noise ratio (SNR) of -1 dB, not when tested at -4 dB SNR. Middle-aged and older listeners did not show these effects. One possible explanation for this pattern is that these effects only emerge when processing of the target stream is easy enough such that individuals are able to dedicate remaining cognitive resources to the processing of the to-be-ignored masker stream.

Taken together, these prior studies provide evidence that the ease of lexical processing of words in the masker can influence target word recognition (Boulenger et al., 2010; Helfer & Jesse, 2015). Younger listeners can recognize words from the target stream more accurately when words in the masker stream are difficult to recognize because they have a lower lexical frequency and/or create more lexical competition. As typically done in speech perception studies, the effects were, however, evaluated on data scored on a dichotomous basis as correct or incorrect, yielding a score that reflects accuracy. The purpose of this study is to examine the possible influence of the ease of lexical processing of masker words by using another metric of speech perception: error patterns. Errors can inform about the breakdown of perceptual and linguistic processing during competing speech tasks, because different types of errors may stem from different underlying processes. For example, error responses that are neither the target nor the masker dominate at very high levels of energetic masking, likely reflecting guessing (Ihfeldt & Shinn-Cunningham, 2008). In tasks where two key words have to be reported on a given trial, answering with one target word and one masker word may reflect a problem in the

sequential organization of sound information over time into a larger coherent stream. In contrast, responding only with masker words (i.e., “source” errors) likely arises from a failure of selectively attending to the target stream (Ihfeldt & Shinn-Cunningham, 2008). Different processes thus seem to underlie the types of errors produced in different listening situations, each susceptible in its own way to energetic and informational masking.

In this study, we investigated the effect of lexical properties of words in the target and masker streams on various types of errors made by younger, middle-aged, and older listeners in a competing speech situation. These error analyses were conducted on the responses collected originally by Helfer and Jesse (2015). We critically expand on Helfer and Jesse’s study by providing a systematic analysis of how aging and lexical properties affect the breakdown in speech perception in the presence of competing speech. In Helfer and Jesse’s study, we focused on how the ease of lexical processing of words in the target and masker streams influenced whether listeners could achieve accurate target recognition. We also reported a preliminary analysis of lexical property effects on the probability of responding with words from the masker. The results of that analysis illustrate the importance of examining errors: Although masker word frequency did not affect accuracy for any age group, older listeners were more likely to erroneously respond with the masker word itself when that word was high rather than low in lexical frequency. Target word frequency affected the probability of responding with a masker word in younger adults who were tested at a more adverse SNR. The analyses in Helfer and Jesse considered, however, only responses that fully matched the target or the masker. In this study, we used a finer-grained approach by categorizing individual errors as phonologically similar to the target word, the masker word, or both, in addition to simply labeling masker errors. This finer-grained approach can help reveal more subtle phonological influences of target and masker words. We thus expanded on our prior work by focusing here on whether and how lexical characteristics of words in two competing streams influence error patterns (i.e., how speech perception breaks down) in participants across the adult life span.

Our research questions were threefold: The first goal was to test whether the phonological influence of the masker words on error responses would increase with aging. Few prior studies investigated the influence of aging on error patterns produced in a competing speech situation. Three of these studies support the contention that older adults respond more often with words from the masker, as compared to younger adults (Helfer, Merchant, & Freyman, 2016; Humes, Lee, & Coughlin, 2006; Meister et al., 2013). Four other studies found that older adults only produced more masker errors under certain conditions, such as when the to-be-ignored and to-be-attended speech streams were presented in a temporally asynchronous manner (Helfer, Mason, & Marino, 2013; Lee & Humes, 2012) or in the presence of a single, same-sex competing talker (Helfer & Freyman, 2014; Helfer & Jesse, 2015). Here, we build upon these studies by using an expanded categorization of error patterns that

allows us to analyze not just whether participants substitute a masker word for a target word in their responses but also whether their incorrect responses were more similar to the target word, to the masker word, or to both words. This detailed analysis can help reveal more subtle age-related changes in the phonological influence of masker words on error responses. We predicted that older adults would respond more often than younger adults with words that are phonologically related to only the masker words or to both masker and target words because of increased susceptibility to lexical interference. This increased masker influence could come, for example, from an age-related decline in inhibiting irrelevant information, in stream segregation, and/or in selectively attending to the target stream. By including data from middle-aged adults, the onset of this increased masker influence on the breakdown of speech perception can be assessed.

The second goal was to test whether lexical properties, that is, word frequency and neighborhood density, of words in the target stream and in the masker stream impact their influence on the phonological composition of errors. Whether a response resembles the target or the masker, or both, indicates the relative phonological influence of these words. The relative influence is modulated by the ease at which a word can be recognized and by its lexical properties. Boulenger et al. (2010) and Helfer and Jesse (2015) both provide evidence that lexical properties of masker words affect the accuracy of target word recognition. In addition, older participants in Helfer and Jesse’s study were more likely to respond with the masker word itself when it was of high rather than low frequency. Based on this prior work, we predicted that, when it is easier to recognize a target or a masker word (i.e., when the words have a higher frequency or a sparser neighborhood density), the word’s phonological influence on the error would increase.

Our third goal was to assess whether the influence of the ease of lexical processing on errors changes across the adult life span. The impact of the lexical properties of target words on their accurate recognition increases in older adults compared to younger adults (Dirks, Takayanagi, Moshfegh, Noffsinger, & Fausti, 2001; Revill & Spieler, 2012; Sommers & Danielson, 1999; Taler, Aaron, Steinmetz, & Pisoni, 2010). Older adults are more susceptible to the neighborhood density of target words than younger adults, in line with the notion that the ability to inhibit phonological competitors decreases with aging (Helfer & Jesse, 2015; Sommers & Danielson, 1999; Taler et al., 2010). In Helfer and Jesse’s (2015) preliminary analyses of lexical properties of target and masker words on masker responses, some age effects had also emerged: As reported above, target word frequency affected the probability of responding with a masker word, but only in younger adults tested in a situation with a more adverse SNR, and not in middle-aged and older adults. In addition, only older adults were more likely to respond with a masker if it was of high rather than low frequency. In this study, we conducted a detailed analysis in order to pinpoint more subtle influences of lexical aspects of target and masker words on the phonological makeup of errors.

Method

Participants

The data included in the error analyses reported here were responses collected by three of four participant groups tested in Helfer and Jesse's (2015) study at -1 dB SNR: older adults (aged 60–83 years, $M = 68$ years), middle-aged adults (aged 45–59 years, $M = 51$ years), and a group of younger adults (aged 19–22 years, $M = 21$ years). Data from a group of younger adults tested at a different SNR (-4 dB) were not included here. Altogether, responses from 45 native speakers of American English were analyzed (i.e., from all 15 participants per group). All participants reported having normal or corrected-to-normal vision and no history of otologic or neurologic disorders. Younger adults had pure-tone thresholds ≤ 25 dB HL in each ear from 250 to 8000 Hz. Older and middle-aged listeners had high-frequency pure-tone thresholds (averaged over 2–6 kHz) ≤ 65 dB HL and a score of at least 25 on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975). The mean high-frequency pure-tone averages were 17.27 dB HL ($SD = 7.65$) for the middle-aged group and 24.13 dB HL ($SD = 14.34$) for the older group. All participants had normal tympanograms on test days.

Materials

The stimulus materials in Helfer and Jesse's (2015) study consisted of pairs of sentences following the structure “*Cue name* discussed the ___ and the ___ today,” with monosyllabic key words in places indicated here by lines. Two female native speakers of American English had been audio-recorded producing these sentences. The cue name was “Theo” for target sentences and “Victor” or “Michael” for masker sentences. All four key words in a target–masker sentence pair were semantically unrelated to each other. Manipulations of spoken word frequency (high, low) and neighborhood density (sparse, dense; Washington University Neighborhood Activation Model database: Sommers, 2000) were manipulated independently for both masker and target words. That is, we did not use a full factorial design but rather manipulated only lexical frequency or neighborhood density in one stream or in the other stream. Although this design does not allow us to measure joint influences, it allows us to assess our question of interest, namely, whether a lexical property of words in the target or masker stream affects the error responses that participants make while trying to recognize targets. The manipulations of lexical frequency and neighborhood density for words in each stream resulted in eight sets of 26 sentence pairs (see Table 1). Words in each respective pair of sets manipulating a lexical property in a given stream were matched as a set on their values on the other lexical property. For example, sets used to manipulate high versus low frequency in the target stream were matched on their lexical neighborhood density. Words in the respective unmanipulated stream also were matched as a set on average to the mean of the set of words in the manipulated stream for the independent variable of interest.

Table 1. Schematic of the manipulations of neighborhood density (N) and lexical frequency in the analyzed eight sets.

Manipulations in data sets	Target category	Masker category
Target N	Sparse Dense	Mid-density Mid-density
Target frequency	Rare Frequent	Mid-frequency Mid-frequency
Masker N	Mid-density Mid-density	Sparse Dense
Masker frequency	Mid-frequency Mid-frequency	Rare Frequent

For example, for the two sets manipulating the frequency of key words in the target stream, the masker streams contained words in a mid-frequency range. In this study, we did not analyze the data collected in Helfer and Jesse's study in additional conditions with a nonspeech noise masker.

Procedure

In Helfer and Jesse's (2015) study, participants completed a test battery, consisting of audiometry, tympanometry, and behavioral tests assessing various cognitive abilities, before participating in the main study. For older and middle-aged participants, the main study took place in a separate, second test session. Target sentences were presented to all listeners with a single-talker competing speech message at a -1 dB SNR. Target sentences were presented at an average root mean square of 68 dBA. Stimuli were played from a single loudspeaker placed about 1 m in front of the listener. Fifteen practice trials were given beforehand. Listeners were instructed to repeat back as quickly as possible the two key words from the sentence starting with “Theo.” The response deadline was 4 s on half of the trials and unlimited on the other half. Because this manipulation did not affect the results, data were pooled for all analyses. Presentation order was fully randomized.

Results

The maximum number of possible responses per participant was 414, because participants were asked to identify two key words on each of the 207 critical trials. Due to an incorrect target–masker assignment in a target neighborhood density set, one trial per subject was removed. Any missing or incorrect response was counted as an error. The overall average error rate calculated out of all possible responses was 42.02% ($SD = 22.12\%$). Figure 1 shows the distributions of the total of errors by listener group. The mean error rates per group were 60.52% ($SD = 23.76\%$) for older adults, 38.34% ($SD = 14.5\%$) for middle-aged adults, and 27.17% ($SD = 10.96\%$) for younger adults. Participants thus produced a sufficient number of errors for our analyses.

Our first analyses investigated how the probability of making one (or two) error(s) changes as a function of age. For these analyses, we analyzed overall error rates, rather

Figure 1. Distributions of total error rates per participant by listener group. The “notch” in these box plots indicates the 95% confidence interval around the median. The size of the box indicates the interquartile range (1st–3rd quartiles), whereas the whiskers add 1.5 times of that range to these quartiles as a cutoff for defining outliers.

than error rates by error type. In Helfer and Jesse’s (2015) study, we had analyzed the effect of age on accuracy separately for each of the four manipulations (frequency vs. neighborhood density, manipulated in the masker vs. target stream). Furthermore, the analyses had focused on the (logit-transformed) proportion of accurate responses out of all possible responses. Our analyses here expand on that prior analysis, in that we assess age effects across a wider range of items (a) by pooling the data over the sets of all manipulations and (b) by analyzing the probability of committing an error on a given trial. To analyze the error data, generalized mixed-effects models were implemented in the glmer function (lme4 package; Bates, Mächler, Bolker, & Walker, 2015) of the R statistical program (Version 3.4.4; R Core Team, 2017) with a binomial link function. We first tested whether the probability of making at least one error on a trial changes across age groups. Trials were coded for one model as 0 or 1, depending on whether participants had made no error versus one or two errors. On average, participants made at least one error on 59.99% of the trials. The effect of age on errors was examined as an orthogonal polynomial contrast. The linear component thus compared error rates of the younger and older adults, and the quadratic component compared the performance of the middle-aged group to the combined performance of younger and older listeners. It therefore tested whether there is sufficient evidence suggesting that the middle-

aged group’s performance does not fall in between that of the younger and older adults. All models included participants as a random factor.

Results show that the probability of making an error on a given trial increases linearly with age (linear: $\beta = 1.392$, $SE = 0.26$, $p < .00001$; quadratic: $\beta = 0.346$, $SE = 0.25$, $p = .18$). Pairwise comparisons were conducted using Tukey tests with Holm–Bonferroni corrections. Older adults were more likely to produce errors on a trial ($M = 77.04\%$, $SD = 18.6\%$) than middle-aged adults ($M = 57.87\%$, $SD = 12.49\%$; $\beta = 1.41$, $SE = 0.37$, $p < .001$). Younger ($M = 45.06\%$, $SD = 11.85\%$) and middle-aged adults did not differ significantly from each other ($\beta = 0.56$, $SE = 0.36$, $p = .12$). The increased probability of committing at least one error on a given trial with aging is in line with the increase observed separately for analyses on each pair of data sets in Helfer and Jesse’s (2015) study for the probability of providing an error response. Although the probability of misreporting one of the target words on a trial only increased in older adults, our prior analyses showed that, overall, the proportion of errors made on a trial increases already by midlife.

In a second analysis, we tested whether the probability of making two errors versus one error on a given trial changes across age groups. On average, participants made one error on 35.93% of the trials and two errors on 24.06% of the trials. Trials were coded for the analysis as 0 or 1, depending on whether participants had made one error or two errors. The probability of giving two rather than one error response on a given trial also increased linearly with age (linear: $\beta = 1.18$, $SE = 0.19$, $p < .00001$; quadratic: $\beta = 0.151$, $SE = 0.19$, $p = .42$). Pairwise comparisons showed that, when a mistake was made on a trial, older listeners were more likely to make two errors rather than just one ($M = 52.52\%$ out of all trials with an error, $SD = 21.66\%$), compared to middle-aged adults ($M = 30.52\%$, $SD = 10.87\%$; $\beta = 1.02$, $SE = 0.26$, $p < .001$). Middle-aged adults were in turn more likely to make two errors rather than one error on a trial, as compared to younger adults ($M = 19.07\%$, $SD = 7.9\%$; $\beta = 0.649$, $SE = 0.27$, $p < .05$). The probability of committing two errors rather than just one error on a given trial thus steadily increased across the adult life span. Together, the results of these analyses suggest that, when participants were more likely to commit one error, they were not necessarily also more likely to commit a second error. Although middle-aged and younger adults were equally likely to make one error on a trial, when middle-aged adults made that error, they were more likely than younger adults to also commit a second error. Older adults had a higher chance than middle-aged adults of making either one or two errors.

Error Type Analyses

To analyze how particular types of errors change as a function of lexical characteristics and age, each error response was classified by a computer script (and checked by hand) as phonologically similar to the target word (target approximations), the masker word (masker approximations),

or both (blends); more detailed descriptions of these error types are provided below. The presented words and the responses were dissected into onset–nucleus–coda subunits. For example, in the word “mood,” /m/ is the onset, /u/ the nucleus, and /d/ the coda. Conversely, the word “ant” has no onset, with /æ/ as the nucleus and /nt/ as the coda. We opted for a liberal criterion of a one phoneme overlap between the target or masker word and the error response because subunits often only consisted of one phoneme and because some of the words only had two phonemes. The overlap was, however, less than a whole subunit in very few cases (in 2.24% of the target approximations and in 5.54% of the masker approximations). Furthermore, we were strict in that the overlap had to be unique to the stimulus word. That is, if the only overlap of an error response was a phoneme found in both the target word and the masker word (e.g., a response of “man” for the target “coin” and the masker “end”), then that response was not categorized. In addition, we ensured that the overlap occurred within the same subunit. For example, “limb,” but not “hole,” would have a match with the target word “slug.” As all target and masker words were monosyllabic, multisyllabic responses were not analyzed.

All other error responses were categorized as belonging to the following categories: *Masker approximation errors* were incorrect responses that were phonologically similar but not identical to the masker word (e.g., “lump” when hearing the masker word “lung”). *Target approximation errors* were incorrect responses that were phonologically similar to the target words (e.g., “boot” in response to the target word “booth”). For both types of approximations, responses had to overlap in at least one phoneme within a subunit with either the masker word or the target word, respectively, and share no phonemes with the respective other word. *Blend errors* were defined as error responses that were a combination of the target word and the masker word (e.g., “goat” in response to the target word “soap” and the masker word “gut”). In addition, we included *masker errors* in our analyses, defined as incorrect responses where the word from the masker stream had been reported. Masker errors had been analyzed before by Helfer and Jesse (2015). We reanalyzed them here as the focus of this error analyses was on the probability of committing each type of error on a given trial, whereas in Helfer and Jesse’s study, we had analyzed overall proportions of each error type.

Age Group Comparisons

To test our first hypothesis that the influence of the masker on the phonological composition of the error response increases with aging, we analyzed the effect of participant group on each error type. Similar generalized mixed-effects models as described above were implemented with a binomial link function. Trials were coded as containing no error (coded as 0) versus one or two errors (coded as 1). As participants made the same type of errors twice on very few trials, we did not conduct a second set of analyses comparing trials with one versus two errors of each type. The exception was masker errors, as they occurred relatively frequently twice per trial. As above, age group was coded

as a polynomial contrast, such that the linear contrast compared the error rates of younger versus older adults and the quadratic contrast compared the performance of the middle-aged group to the combined performance of younger and older listeners. Only significant results are reported in the text. All models included participants as a random factor. Pairwise comparisons were conducted using Tukey tests with Holm–Bonferroni corrections.

Figure 2 shows the proportion of trials with at least one error of each type by age group. The probability of responding with a word that was phonologically similar to the target (i.e., a target approximation) did not change across age groups (all p s > .05). The influence of the masker words on errors, however, increased with aging: The probability of responding with a word phonologically related to the masker word ($\beta = 0.898$, $SE = 0.22$, $p < .0001$) or with the masker word itself ($\beta = 0.818$, $SE = 0.25$, $p < .001$) both increased linearly across age groups. Comparisons across adjacent age groups showed only marginally significant increments in masker approximation responses (younger adults: $M = 3.03\%$, $SD = 1.57\%$; middle-aged adults: $M = 5.86\%$, $SD = 3.67\%$; $\beta = 0.629$, $SE = 0.31$, $p = .07$; middle-aged vs. older adults: $M = 11.66\%$, $SD = 10.35\%$; $\beta = 0.64$, $SE = 0.3$, $p = .07$), with no significant differences for masker errors (younger adults: $M = 9.89\%$, $SD = 6.36\%$; middle-aged adults: $M = 15.65\%$, $SD = 8.62\%$; $\beta = 0.639$, $SE = 0.35$, $p = .13$; middle-aged vs. older adults: $M = 26.63\%$, $SD = 22.16\%$; $\beta = 0.52$, $SE = 0.34$, $p = .13$).

In addition, we compared the probability of responding with one versus two masker words on a trial. Trials where listeners followed the wrong stream and thus only reported words that were from the masker or that were phonologically related to the masker can be regarded as “source” errors, likely arising from a failure of selectively attending to the target stream (Ihfeldt & Shinn-Cunningham, 2008). The likelihood of responding with two masker words rather than one masker word increased linearly with age ($\beta = 0.605$, $SE = 0.2$, $p < .01$), although none of the pairwise comparisons was significant (all p s > .05). Hence, the likelihood of failing to selectively attend to the speech stream of the target speaker increases across the adult life span.

The probability of responding with a phonological blend of both target and masker words followed both a linear trend ($\beta = 0.407$, $SE = 0.1$, $p < .0001$) and a quadratic trend ($\beta = -0.274$, $SE = 0.09$, $p < .01$) across age groups. The probability of a blend error increased with age but peaked at midlife, after which it stabilized. Pairwise comparisons thus showed an increase from younger ($M = 7.73\%$, $SD = 3.46\%$) to middle-aged adults ($M = 13.27\%$, $SD = 3.32\%$; $\beta = 0.623$, $SE = 0.14$, $p < .0001$), but no further change between middle-aged and older adults ($M = 12.79\%$, $SD = 3.79\%$; $\beta = -0.048$, $SE = 0.13$, $p = .71$).

Taken together, these results suggest that the extent to which the masker words, but not target words, influence errors increases gradually with aging, thus supporting our hypothesis. Furthermore, the propensity to produce errors that blend the phonological makeup of maskers and targets increases with age but stabilizes by midlife.

Figure 2. Counts of trials with each type of error by age group.

Lexical Properties

Next, we tested our two other hypotheses: that the ease of recognizing target words and masker words modulates their influence on error responses and that the influence of lexical characteristics of masker words would change across the adult life span. Analyses were conducted separately on the data collected for manipulations of frequency and neighborhood density in the target and masker speech streams. Generalized mixed-effects models with a binomial link function were implemented for each error type with age as a polynomial contrast and with the respective lexical property as a contrast-coded fixed factor (sparse = -0.5 , dense = 0.5 ; low frequency = -0.5 , high frequency = 0.5). The random-effect structure of all models included by-subject intercepts and within-subject factor slope adjustments. Alpha levels in post hoc tests following significant interaction effects were Bonferroni corrected. Again, trials were coded as containing no error versus one or two errors of each type. Figures 3–6 show the total number of trials with at least one error of each type, aggregated by age group, separately for neighborhood density and frequency manipulations in the two streams. Results of the statistical analyses are reported in Tables 2–5. Only significant effects of the lexical property or its interaction with age are reported in the text.

Lexical properties of target words. The frequency of the target word affected the probability of responding with each type of error (see Figure 3 and Table 2): When

participants were asked to identify high-frequency targets rather than low-frequency targets, the probability of all four types of errors decreased (target approximations: $\beta = -0.471$, $SE = 0.12$, $p < .0001$; masker approximations: $\beta = -0.592$, $SE = 0.30$, $p < .05$; masker errors: $\beta = -0.78$, $SE = 0.21$, $p < .001$; blends: $\beta = -1.127$, $SE = 0.19$, $p < .00001$). These effects were not modulated by age (all $p > .05$), except that the interaction between target frequency and the quadratic trend for age was significant for target approximations ($\beta = 0.446$, $SE = 0.19$, $p < .05$), as the frequency effect on target approximations was only observed in middle-aged adults ($\beta = -0.808$, $SE = 0.18$, $p < .0001$).

Neighborhood density of the target word also affected the probability of responding with each type of error (see Figure 4 and Table 3): When participants had to identify targets from dense rather than sparse phonological neighborhoods, the probability of all four types of errors increased (target approximations: $\beta = 0.197$, $SE = 0.10$, $p = .059$; masker approximations: $\beta = 0.99$, $SE = 0.29$, $p < .001$; masker errors: $\beta = 0.374$, $SE = 0.14$, $p < .01$; blends: $\beta = 0.298$, $SE = 0.15$, $p < .05$). These effects were also not modulated by age (all $p > .05$), although the interaction effect of target neighborhood density with the linear trend for age on masker approximations was marginally significant ($\beta = -0.785$, $SE = 0.45$, $p = .08$).

In summary, both lexical frequency and neighborhood density of target words influenced the probability that

Figure 3. Counts of trials with each type of error by target word frequency, aggregated by age group.

Figure 4. Counts of trials with each type of error by target neighborhood density (N), aggregated by age group.

Figure 5. Counts of trials with each type of error by masker word frequency, aggregated by age group.

Figure 6. Counts of trials with each type of error by masker neighborhood density (N), aggregated by age group.

Table 2. Linear mixed-effects models for error types in the target word frequency conditions.

Effects	Target approximations			Masker approximations		
	Estimate	SE	p	Estimate	SE	p
Frequency	−0.471	0.12	< .0001	−0.592	0.3	< .05
Linear contrast age	−0.035	0.12	.78	1.037	0.34	< .01
Quadratic contrast age	−0.259	0.12	<.05	0.177	0.34	.61
Frequency × Linear Contrast Age	−0.93	0.20	.64	0.6	0.39	.12
Frequency × Quadratic Contrast Age	0.446	0.19	< .05	0.229	0.38	.55

Effects	Blend responses			Masker responses		
	Estimate	SE	p	Estimate	SE	p
Frequency	−1.127	0.19	< .0001	−0.780	0.21	< .001
Linear contrast age	0.292	0.17	.08	1.089	0.32	< .001
Quadratic contrast age	−0.305	0.16	.06	0.086	0.31	.78
Frequency × Linear Contrast Age	−0.196	0.31	.52	0.344	0.33	.3
Frequency × Quadratic Contrast Age	0.209	0.28	.46	−0.232	0.32	.47

participants made errors of each type on a trial. All four types of errors became more likely when targets were more difficult to recognize because their frequency of occurrence in daily life was low or because the amount of lexical competition with phonologically similar words was high. All age groups were equally susceptible to these effects. The only exception was that the influence of target word frequency on target approximations was only found at midlife.

Lexical properties of masker words. The lexical properties of the masker words also influenced error responses. When participants had to identify target words while hearing high-frequency masker words compared to low-frequency masker words, target approximation responses became less likely ($\beta = -0.509$, $SE = 0.12$, $p < .00001$) and masker responses became more likely ($\beta = 1.034$, $SE = 0.14$, $p < .00001$; see Figure 5 and Table 4). These effects did not change with age. No main effects of masker word frequency were found for masker approximations and blend errors. For blend errors, an interaction between masker word frequency with the linear age trend was significant ($\beta = 0.613$, $SE = 0.25$, $p < .05$), as the linear trend with

age was only found for high-frequency words ($\beta = 0.877$, $SE = 0.2$, $p < .0001$). Testing for a frequency effect on blend responses for each age group revealed no significant results (all $p > .05$). For masker approximations, masker word frequency interacted with the quadratic age trend ($\beta = -0.715$, $SE = 0.32$, $p < .05$). Although masker word frequency had no significant effect in any age group, its numerical trend varied across groups: Younger adults showed no numerical difference, middle-aged adults showed a decrease, and older adults showed an increase in masker approximations with high- versus low-frequency masker words.

Masker neighborhood density also had an influence on errors (see Figure 6 and Table 5). When the masker words came from dense rather than sparse neighborhoods, participants made fewer target approximation errors ($\beta = -0.406$, $SE = 0.12$, $p < .001$) and fewer masker approximation errors ($\beta = -0.533$, $SE = 0.21$, $p < .011$) but showed a marginally significant increase in blends ($\beta = 0.264$, $SE = 0.15$, $p = .07$). No interaction with age was found (an interaction of the quadratic trend of age with masker

Table 3. Linear mixed-effects models for error types in the target neighborhood density conditions.

Effects	Target approximations			Masker approximations		
	Estimate	SE	p	Estimate	SE	p
Neighborhood density	0.197	0.10	.06	0.99	0.2	< .001
Linear contrast age	−0.036	0.12	.77	1.32	0.33	< .0001
Quadratic contrast age	−0.13	0.12	.27	−0.299	0.3	.32
Neighborhood Density × Linear Contrast Age	−0.155	0.18	.39	−0.785	0.45	.08
Neighborhood Density × Quadratic Contrast Age	0.003	0.17	.99	0.335	0.37	.37

Effects	Blend responses			Masker responses		
	Estimate	SE	p	Estimate	SE	p
Neighborhood density	0.298	0.15	< .05	0.374	0.14	< .01
Linear contrast age	0.362	0.13	< .01	0.939	0.26	< .001
Quadratic contrast age	−0.157	0.12	.18	0.028	0.26	.91
Neighborhood Density × Linear Contrast Age	−0.08	0.26	.76	−0.067	0.22	.76
Neighborhood Density × Quadratic Contrast Age	0.15	0.24	.54	0.141	0.22	.51

Table 4. Linear mixed-effects models for error types in the masker word frequency conditions.

Effects	Target approximations			Masker approximations		
	Estimate	SE	p	Estimate	SE	p
Frequency	−0.509	0.12	< .0001	0.113	0.22	.60
Linear contrast age	−0.007	0.14	.96	0.808	0.24	< .001
Quadratic contrast age	−0.067	0.14	.63	0.140	0.24	.56
Frequency × Linear Contrast Age	−0.072	0.2	.72	−0.284	0.32	.37
Frequency × Quadratic Contrast Age	−0.047	0.2	.81	−0.715	0.32	< .05

Effects	Blend responses			Masker responses		
	Estimate	SE	p	Estimate	SE	p
Frequency	−0.047	0.15	.75	1.034	0.14	< .00001
Linear contrast age	0.566	0.14	< .0001	0.701	0.24	< .01
Quadratic contrast age	−0.353	0.13	< .01	−0.168	0.24	.49
Frequency × Linear Contrast Age	0.613	0.25	< .05	−0.02	0.23	.93
Frequency × Quadratic Contrast Age	−0.151	0.22	.5	−0.346	0.22	.11

neighborhood density on target approximations was only marginally significant, $p = .07$). In contrast, the probability of responding with the masker word itself was not affected by whether the masker was phonologically similar to many other words or to few other words ($p > .05$). A marginally significant interaction with the linear trend of age was found ($\beta = 0.479$, $SE = 0.28$, $p = .08$), as only older adults were sensitive to masker neighborhood density ($\beta = 0.489$, $SE = 0.21$, $p < .05$).

In summary, when masker words were more easily recognizable because they were higher rather than lower in lexical frequency, participants were more likely to respond with those words but were not more likely to use parts of their sounds in an error response. At the same time, the phonological influence of target words (i.e., target approximations) on error responses was reduced. When masker words were more easily recognizable due to lesser degrees of lexical competition with neighbors, participants were more likely to respond with errors that included phonemes from the masker or the target but were somewhat less likely to blend targets and maskers. The probability of

responding with the masker word itself was not influenced by masker density. For the most part, these effects remained stable across the adult life span. The only apparent group difference was that blend errors increased linearly with age, but only when masker words were high in lexical frequency.

Discussion

Recognizing spoken words in situations with a competing speech stream can be challenging due to energetic masking and because the processing of the masker may compete with the processing of the target stream. Lexical properties of words in both speech streams modulate the extent of this informational masking on target word recognition (Boulenger et al., 2010; Helfer & Jesse, 2015). However, accuracy of word recognition is not the only dimension that can be used to quantify speech perception. Analyses of errors can inform about the breakdown of perceptual and linguistic processing in recognizing words in situations with competing speech, because different types of errors

Table 5. Linear mixed-effects models for error types in the masker neighborhood density conditions.

Effects	Target approximations			Masker approximations		
	Estimate	SE	p	Estimate	SE	p
Neighborhood density	−0.406	0.12	< .001	−0.533	0.21	< .05
Linear contrast age	0.061	0.15	.68	1.063	0.23	< .00001
Quadratic contrast age	−0.153	0.14	.29	−0.0001	0.23	.99
Neighborhood Density × Linear Contrast Age	−0.254	0.20	.21	0.472	0.34	.17
Neighborhood Density × Quadratic Contrast Age	0.335	0.19	.07	0.352	0.32	.27

Effects	Blend responses			Masker responses		
	Estimate	SE	p	Estimate	SE	p
Neighborhood density	0.264	0.15	.07	−0.159	0.18	.37
Linear contrast age	0.376	0.14	< .01	0.746	0.31	< .05
Quadratic contrast age	−0.260	0.13	< .05	−0.092	0.31	.76
Neighborhood Density × Linear Contrast Age	−0.336	0.25	.18	0.48	0.28	.08
Neighborhood Density × Quadratic Contrast Age	−0.172	0.23	.46	0.146	0.27	.59

may stem from different underlying processes. The segmental makeup of an error can provide insight into the influence of target and masker words during recognition. This study analyzed error patterns to test three main hypotheses. We predicted (a) that the influence of the masker on the phonological makeup of errors increases across the adult life span, (b) that the ease of recognizing target words and masker words modulates their influence on error responses, and (c) that the influence of lexical characteristics of masker words would vary across the adult life span.

Changes in the Influence of the Masker on Speech Perception Across the Adult Life Span

Middle-aged and older adults experience disproportionately more problems than younger adults in difficult listening situations, a finding that cannot be entirely explained by a lack of audibility of the speech message (e.g., Helfer & Freyman, 2008, 2014; Tun & Wingfield, 1999). The focus of the analyses in the current article was to help identify the underlying processes used in a listening situation that is especially problematic for older adults: when the masker consists of a single competing talker (e.g., Duquesnoy, 1983; Helfer & Freyman, 2014; Tun et al., 2002; Tun & Wingfield, 1999). Although part of older adults' challenges in this type of listening situation could be from age-related decline in the ability to use brief glimpses that occur within a fluctuating speech masker (e.g., George, Festen, & Houtgast, 2006; Takahashi & Bacon, 1992), it is unclear whether older adults' increased susceptibility to the content of the masker stream also plays a role. Previous studies have had conflicting outcomes: Whereas some investigations support this claim (Helfer & Freyman, 2014; Helfer & Jesse, 2015; Helfer et al., 2013; Lee & Humes, 2012), others found no difference between older and younger adults' probability of reporting words from an intelligible speech masker (Helfer et al., 2016; Humes et al., 2006; Meister et al., 2013). In the present article, we presented a finer-grained analysis of error patterns as a means of identifying subtle age-related differences in how to-be-attended and to-be-ignored speech are processed.

Our results showed that the probability of responding with an error that was phonologically related to the target word did not change with aging. We assume that these target approximation errors are related primarily to energetic masking: The listener was able to correctly determine which word was the target, but portions of the word were obscured by energetic masking. It appears that listeners of all ages were susceptible to this type of masking. However, the probability of combining parts of the masker and target into a blend response increased and peaked by midlife. Responses of this type could be linked to an inability to segregate the target from the masker. As predicted, we also found an increase in incorrect responses that were phonologically related to the masker words in our middle-aged and older groups. The older the age group, the more likely its members responded on a trial with an error that was a word from the masker or one that was phonologically related

to the masker. The latter of these responses (masker approximations) are likely due to a breakdown in segregation combined with susceptibility to energetic masking: That is, listeners selected the wrong word, and energetic masking led to portions of that word to become inaudible. This increase in the masker's influence appears to begin by middle age, in line with prior work showing that middle-aged listeners have more difficulty understanding speech in a competing speech background than younger listeners (Glyde, Cameron, Dillon, Hickson, & Seeto, 2013; Helfer, 2015; Helfer & Freyman, 2016; Helfer & Jesse, 2015; Helfer & Vargo, 2009). In this study, the increased influence of the masker across the adult life span was not only evident in phonological intrusions in the responses but also in the probability of reporting both masker words on a trial. This latter result is in line with prior work suggesting a decline in the ability to selectively attend to the target stream that begins by midlife, perhaps due to difficulties with temporal processing (Ruggles et al., 2012). An alternative explanation for the observed age-related increase in the masker's influence is that, as listeners age, their ability to inhibit irrelevant information declines (Hasher & Zacks, 1988). Regardless of the source of these errors, our results support the idea that aging brings about an increase in interference from to-be-ignored speech streams in a competing speech situation.

In summary, the results support our hypothesis that, as listeners age, the influence of the masker stream increases. As a consequence, listeners are more likely to experience phonological intrusions into their responses. In this study, aging also brought about a tendency to follow the wrong speech stream, suggesting a breakdown in selective attention. However, it is possible that responding with words from the masker reflects a response strategy rather than a breakdown in stream selection; that is, when participants were tasked with repeating back words they heard, they chose to report a word that was in the masker even if they knew that it was not a target word.

The Influence of Lexical Properties of Target and Masker Words on the Phonological Makeup of Errors Across the Adult Life Span

Both neighborhood density and lexical frequency determine the ease of recognition of a word (e.g., Dahan et al., 2001; Goldinger et al., 1989; Luce & Pisoni, 1998). The lexical properties of words in the target stream, but also of words in the masker, can modulate the recognition of target words (Boulenger et al., 2010; Helfer & Jesse, 2015). We thus predicted that the lexical properties of target and masker words would also affect the nature of the errors made by listeners. Furthermore, we predicted that these influences would change with age. Prior work on spoken word recognition in situations without competing speech suggests that the accurate recognition of a target word is impacted more by the target word's properties in older adults than in younger adults (e.g., Dirks et al., 2001; Revill & Spieler, 2012; Sommers & Danielson, 1999). On the other hand, this

increased susceptibility to lexical properties of target words with aging has only been partially observed in our prior analysis of recognition accuracy in a listening situation with competing speech (Helfer & Jesse, 2015). Older adults' ability to correctly recognize the target was less influenced by that word's lexical frequency as compared to middle-aged and younger adults tested at the same SNR. Older adults were more susceptible to the target words' neighborhood density than younger adults tested at the same SNR, but not when younger adults were tested at a more difficult SNR. Hence, older adults may not be more susceptible than younger adults to the lexical properties of target words in their ability to correctly recognize a target word in a listening situation with competing speech. In terms of the lexical properties of words in the masker, the results of our prior analysis (Helfer & Jesse, 2015) indicated that the influence of the masker's neighborhood density on correctly reporting the target words was only apparent for younger adults. Only older adults were susceptible to masker's lexical frequency in their probability of responding with words from the masker (although all other age groups showed only numerical trends). In this study, we sought to determine if this result persists when examining error patterns in a more comprehensive manner and whether this more detailed analysis of errors allows us to detect subtle influences in listeners spanning the adult age range.

Indeed, lexical properties of words in both streams modulated errors in all age groups: The ease of recognition of target words affected all error types. That is, when target words were easier to recognize as they were higher in lexical frequency or from sparse neighborhoods, all types of errors decreased. The lexical properties of target words thus have a global effect on errors, indicating that they affect overall accuracy. In contrast, the lexical properties of maskers only affected certain error types. When masker words had a higher rather than lower frequency, the error responses were less likely to be related to only the target words and, instead, masker errors became more probable. In this case, masker words may have been recognized more quickly. Participants could have thus selected the masker words because they became available before the target words. Alternatively, as more frequent (vs. less frequent) masker words would start to compete earlier with target words, they were more difficult to inhibit. Masker words were therefore more likely to be selected. The latter explanation would predict, however, that the masker frequency effect on masker responses would increase with aging, as inhibitory control declines. This was not observed. The results of our current analyses differ in that regard from those previously reported by Helfer and Jesse (2015), where the masker frequency effect on masker responses was only significant for older adults. Note, however, that in our previous article, we had analyzed the proportion of masker responses. Here, we instead analyzed the probability of responding with at least one word from the masker. That probability is influenced by the masker word's frequency in all age groups.

One other interesting result of the present analyses is the prevalence of blend errors. When faced with ambiguity

or difficulty in identifying the target and masker words, participants sometimes responded with a word that contained elements of both the target and the masker. These types of responses were made more often when the masker was from sparse rather than dense lexical neighborhoods. In that situation, error responses were less likely to be phonologically related to just target words or to just masker words and instead were more likely to show phonological influences of both. These blend errors suggest that, in adverse listening situations, individuals may meld parts of to-be-ignored and to-be-attended speech streams. This could be another reflection of listeners' difficulty segregating the target from the masker on a segmental level.

In contrast to our prediction, the influence of the ease of lexical processing of target and masker words on errors remained similar across the adult life span. The only hint of an age effect was that the influence of target word frequency on target approximations was larger at midlife than at younger or older age. Although the ease of recognition of target and masker words can influence how processing breaks down, it does so similarly for individuals spanning the adult age range. This result differs from the increased susceptibility to lexical characteristics of target words observed in listening situations without competing speech (e.g., Dirks et al., 2001; Revill & Spieler, 2012; Sommers & Danielson, 1999). This apparent disparity in results could be due to the difference in listening conditions, as task demands and the need for cognitive resources certainly differ when trying to process a target speech stream in isolation versus when faced with the challenge of recognizing a message in the presence of competing speech. Hence, our results do not support the idea that the influence of lexical characteristics on the types of errors produced in the presence of a single competing talker changes with aging. Although adults are more and more prone to be affected by the masker as they age, the influence of lexical properties on this susceptibility remains constant.

Conclusions

Lexical properties of words in both attended and unattended speech streams influence listeners' ability to correctly recognize a target word as well as the breakdown in word recognition. Although how easily the target word is recognized affects all error types, the effects of the ease of recognition of the masker are more specific. Higher lexical frequency in masker words reduces the phonological influence of target words and increases the probability of responding with the actual masker word. Moreover, lower levels of lexical competition for the masker increase the probability that listeners report an error that is phonologically related to both masker and target words. Across the adult life span, listeners' error responses are increasingly affected by the phonological composition of words in the masker. The increasing influence of the masker also is evident in the temporary failure in selectively attending to the target stream. Hence, the processes involved in age-related changes in error responses are likely both a reduction in

inhibitory ability and problems with selective attention. The increase in both cases is linear and is already evident in midlife. The effects of the ease of lexical processing of target and masker words on error types remain, however, largely the same across the adult life span.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grant R01 012057, awarded to Karen S. Helfer. We would like to thank Michael Rogers for his help with this project and Aline Sayer, Adrian Staub, and Andrew Cohen for valuable comments. Part of this work has been presented as “Lexical Influence on Error Patterns in Competing Speech Perception” at the Meeting of the Acoustical Society of America, Boston, MA, in June 2017, and as “Lexical Characteristics of Words in Competing Speech Streams Predict How Processing Breaks Down” at the Aging and Speech Communication Research Conference, Tampa, FL, in November 2017.

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Boulenger, V., Hoen, M., Ferragne, E., Pellegrino, F., & Meunier, F. (2010). Real-time lexical competitions during speech-in-speech comprehension. *Speech Communication*, 52(3), 246–253. <https://doi.org/10.1016/j.specom.2009.11.002>
- Brouwer, S., & Bradlow, A. R. (2016). The temporal dynamics of spoken word recognition in adverse listening conditions. *Journal of Psycholinguistic Research*, 45(5), 1151–1160. <https://doi.org/10.1007/s10936-015-9396-9>
- Calandruccio, L., Dhar, S., & Bradlow, A. R. (2010). Speech-on-speech masking with variable access to the linguistic content of the masker speech. *The Journal of the Acoustical Society of America*, 128(2), 860–869. <https://doi.org/10.1121/1.3458857>
- Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, 42(4), 317–367. <https://doi.org/10.1006/cogp.2001.0750>
- Dirks, D. D., Takayanagi, S., Moshfegh, A., Noffsinger, P. D., & Fausti, S. A. (2001). Examination of the neighborhood activation theory in normal and hearing-impaired listeners. *Ear and Hearing*, 22(1), 1–13. <https://doi.org/10.1097/00003446-200102000-00001>
- Duquesnoy, A. J. (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *The Journal of the Acoustical Society of America*, 74(3), 739–743. <https://doi.org/10.1121/1.389859>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Minimal state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198. [https://doi.org/10.1002/\(SICI\)1099-1166\(199805\)13:5<285::AID-GPS753>3.0.CO;2-V](https://doi.org/10.1002/(SICI)1099-1166(199805)13:5<285::AID-GPS753>3.0.CO;2-V)
- Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, 12(5–6), 613–656. <https://doi.org/10.1080/016909697386646>
- George, E. L. J., Festen, J. M., & Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 120(4), 2295–2311. <https://doi.org/10.1121/1.2266530>
- Glyde, H., Cameron, S., Dillon, H., Hickson, L., & Seeto, M. (2013). The effects of hearing impairment and aging on spatial processing. *Ear and Hearing*, 34(1), 15–28. <https://doi.org/10.1097/AUD.0b013e3182617f94>
- 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. [https://doi.org/10.1016/0749-596x\(89\)90009-0](https://doi.org/10.1016/0749-596x(89)90009-0)
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation*, 22, 193–225. [https://doi.org/10.1016/s0079-7421\(08\)60041-9](https://doi.org/10.1016/s0079-7421(08)60041-9)
- Helfer, K. S. (2015). Competing speech perception in middle age. *American Journal of Audiology*, 24(2), 80–83. https://doi.org/10.1044/2015_AJA-14-0056
- Helfer, K. S., & Freyman, R. L. (2008). Aging and speech-on-speech masking. *Ear and Hearing*, 29(1), 87–98.
- Helfer, K. S., & Freyman, R. L. (2014). Stimulus and listener factors affecting age-related changes in competing speech perception. *The Journal of the Acoustical Society of America*, 136(2), 748–759. <https://doi.org/10.1121/1.4887463>
- Helfer, K. S., & Freyman, R. L. (2016). Age equivalence in the benefit of repetition for speech understanding. *The Journal of the Acoustical Society of America*, 140(5), EL371–EL377. <https://doi.org/10.1121/1.4966586>
- Helfer, K. S., & Jesse, A. (2015). Lexical influences on competing speech perception in younger, middle-aged, and older adults. *The Journal of the Acoustical Society of America*, 138(1), 363–376. <https://doi.org/10.1121/1.4923155>
- Helfer, K. S., Mason, C. R., & Marino, C. (2013). Aging and the perception of temporally interleaved words. *Ear and Hearing*, 34(2), 160–167. <https://doi.org/10.1097/AUD.0b013e31826a8ea7>
- Helfer, K. S., Merchant, G. R., & Freyman, R. L. (2016). Aging and the effect of target-masker alignment. *The Journal of the Acoustical Society of America*, 140(5), 3844–3853. <https://doi.org/10.1121/1.4967297>
- Helfer, K. S., & Vargo, M. (2009). Speech recognition and temporal processing in middle-aged women. *Journal of the American Academy of Audiology*, 20(4), 264–271. <https://doi.org/10.3766/jaaa.20.4.6>
- Humes, L. E., Lee, J. H., & Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *The Journal of the Acoustical Society of America*, 120(5), 2926–2937. <https://doi.org/10.1121/1.2354070>
- Ihlefelf, A., & Shinn-Cunningham, B. (2008). Spatial release from energetic and informational masking in a divided speech identification task. *The Journal of the Acoustical Society of America*, 123(6), 4380–4392. <https://doi.org/10.1121/1.2904825>
- Lee, J. H., & Humes, L. E. (2012). Effect of fundamental-frequency and sentence-onset differences on speech-identification performance of young and older adults in a competing-talker background. *The Journal of the Acoustical Society of America*, 132(3), 1700–1717. <https://doi.org/10.1121/1.4740482>
- Luce, P. A. (1986). *Neighborhoods of words in the mental lexicon* (Doctoral dissertation). In Research on Speech Perception (Technical Report No. 6). Speech Research Laboratory, Department of Psychology, Indiana University.
- Luce, P. A., Goldinger, S. D., Auer, E. T., Jr., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics*, 62(3), 615–625. <https://doi.org/10.3758/bf03212113>

- Luce, P. A., & Pisoni, D. B.** (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1–36. <https://doi.org/10.1097/00003446-199802000-00001>
- Luce, P. A., Pisoni, D. B., & Goldinger, S. D.** (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 122–147). Cambridge, MA: MIT Press.
- Marslen-Wilson, W. D.** (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*(1–2), 71–102. [https://doi.org/10.1016/0010-0277\(87\)90005-9](https://doi.org/10.1016/0010-0277(87)90005-9)
- Marslen-Wilson, W. D.** (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148–172). Cambridge, MA: MIT Press.
- McClelland, J. L., & Elman, J. L.** (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1–86. [https://doi.org/10.1016/0010-0285\(86\)90015-0](https://doi.org/10.1016/0010-0285(86)90015-0)
- Meister, H., Schreitmüller, S., Grugel, L., Ortman, M., Beutner, D., Walger, M., & Meister, I. G.** (2013). Cognitive resources related to speech recognition with a competing talker in young and older listeners. *Neuroscience, 232*, 74–82. <https://doi.org/10.1016/j.neuroscience.2012.12.006>
- Norris, D.** (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*(3), 189–234. [https://doi.org/10.1016/0010-0277\(94\)90043-4](https://doi.org/10.1016/0010-0277(94)90043-4)
- Norris, D., & McQueen, J. M.** (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review, 115*(2), 357–395. <https://doi.org/10.1037/0033-295X.115.2.357>
- Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S.** (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology, 34*(3), 191–243. <https://doi.org/10.1006/cogp.1997.0671>
- R Core Team.** (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Revill, K. P., & Spieler, D. H.** (2012). The effect of lexical frequency on spoken word recognition in young and older listeners. *Psychology and Aging, 27*(1), 80–87. <https://doi.org/10.1037/a0024113>
- Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B.** (2012). Why middle-aged listeners have trouble hearing in everyday situations. *Current Biology, 22*, 1417–1422.
- Sommers, M. S.** (2000). *Washington University neighborhood database* [database]. Retrieved from <http://neighborhoodsearch.wustl.edu/neighborhood/Home.asp>
- 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), 458–472. <https://doi.org/10.1037/0882-7974.14.3.458>
- Takahashi, G. A., & Bacon, S. P.** (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *Journal of Speech and Hearing Research, 35*(6), 1410–1421. <https://doi.org/10.1044/jshr.3506.1410>
- 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), P551–P560. <https://doi.org/10.1093/geronb/gbq039>
- Tun, P. A., O’Kane, G., & Wingfield, A.** (2002). Distraction by competing speech in young and older adult listeners. *Psychology and Aging, 17*(3), 453–467. <https://doi.org/10.1037/0882-7974.17.3.453>
- Tun, P. A., & Wingfield, A.** (1999). One voice too many: Adult age differences in language processing with different types of distracting sounds. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 54*(5), P317–P327. <https://doi.org/10.1093/geronb/54B.5.P317>
- Van Engen, K. J., & Bradlow, A. R.** (2007). Sentence recognition in native- and foreign-language multi-talker background noise. *The Journal of the Acoustical Society of America, 121*(1), 519–526. <https://doi.org/10.1121/1.2400666>