

Cognitive Psychology

Effects of Age, Word Frequency, and Noise on the Time Course of Spoken Word Recognition

Kristin J. Van Engen^{1, a}, Avanti Dey², Nichole Runge¹, Brent Spehar³, Mitchell S. Sommers¹, Jonathan E. Peelle³

¹ Department of Psychological and Brain Sciences, Washington University in St. Louis, MO, US, ² Department of Psychological and Brain Sciences, Washington University in St. Louis, MO, US; Center for Vital Longevity, University of Texas at Dallas, TX, US, ³ Department of Otolaryngology, Washington University in St. Louis, MO, US

Keywords: noise, lexical frequency, aging, eyetracking, word recognition

<https://doi.org/10.1525/collabra.17247>

Collabra: Psychology

Vol. 6, Issue 1, 2020

This study assessed the effects of age, word frequency, and background noise on the time course of lexical activation during spoken word recognition. Participants (41 young adults and 39 older adults) performed a visual world word recognition task while we monitored their gaze position. On each trial, four phonologically unrelated pictures appeared on the screen. A target word was presented auditorily following a carrier phrase ("Click on _____"), at which point participants were instructed to use the mouse to click on the picture that corresponded to the target word. High- and low-frequency words were presented in quiet to half of the participants. The other half heard the words in a low level of noise in which the words were still readily identifiable. Results showed that, even in the absence of phonological competitors in the visual array, high-frequency words were fixated more quickly than low-frequency words by both listener groups. Young adults were generally faster to fixate on targets compared to older adults, but the pattern of interactions among noise, word frequency, and listener age showed that older adults' lexical activation largely matches that of young adults in a modest amount of noise.

Introduction

Contemporary models of spoken word recognition generally agree on a lexical competition framework in which similar-sounding words (phonological neighbors) compete with each other for recognition, so that recognition difficulty is related to the number of competitors a word has in the lexicon (Luce & Pisoni, 1998; Marslen-Wilson & Tyler, 1980; Norris & McQueen, 2008). Lexical representations are activated by incoming acoustic information, and listeners must select a response from the possible candidates while inhibiting incorrect options. A word like "cat", for example, has a large number of similar-sounding competitors ("cap", "calf", "hat"...), and would thus be more difficult to accurately perceive than a word like "orange", which has few competitors. Many studies infer the challenge caused by competitors by looking at identification errors, observing that words with many competitors are recognized less often in noise (e.g., (Goldinger et al., 1989; Luce & Pisoni, 1998; Sommers & Danielson, 1999). However, a key characteristic of lexical competition frameworks is that these processes of lexical activation, inhibition, and selection operate even when recognition is successful. While lexical activation may be partly automatic, selection and inhibition have been proposed to rely on additional cognitive resources (Sommers & Danielson, 1999). Thus, lexical competition may underlie at least a portion of the cognitive challenge associated with effortful listening (Peelle, 2018; Pichora-Fuller et al., 2016).

The problem of effortful listening is particularly relevant for older adults, who experience increased difficulty with spoken word recognition, especially in the context of back-

ground noise. This difficulty is likely due to a combination of auditory and cognitive factors (Humes et al., 2013): in addition to high rates of hearing loss among older adults, which can limit auditory access to speech signals, there is evidence to suggest that they are particularly affected by the cognitive challenge associated with lexical competition. Sommers & Danielson (1999), for example, showed that older adults had disproportionate difficulty recognizing words with many phonological neighbors (i.e., from dense phonological neighborhoods) relative to young adults when they were tested at noise levels that equated the groups' recognition of words with few neighbors. In the current study, we examine the effect of another lexical factor—word frequency—on young and older adult listeners in both quiet and noisy conditions. High-frequency words are typically recognized more quickly and reliably than low-frequency words. As such, most models of speech recognition assume that frequency affects the baseline activation levels of lexical candidates (e.g., Marslen-Wilson, 1987) or the strength of connections between sublexical and lexical representations (MacKay, 1982, 1987).

One productive approach to measuring lexical activation in the absence of recognition errors has been the visual world paradigm (Allopenna et al., 1998; Cooper, 1974). In a typical visual world experiment, an array of pictures is presented on a screen in front of a participant, who hears a word and is asked to indicate what they heard. The direction of the participant's gaze is tracked and used to index lexical activation. Conveniently, then, eye tracking can be used to measure the activation of both target words and distractors in the visual array. For example, for the target word

“beaker”, competitors like “beetle” and “speaker” also receive some looks from listeners (Alloppenna et al., 1998).

Using a visual world paradigm, Ben-David et al. (2009) compared the effects of competitors that either shared onsets (*candle-candy*) or endings (*candle-sandal*) on young and older adults in quiet and in noise. The authors found that spoken word recognition processes were generally similar across the age groups. They did find one age difference: older adults were slightly more slowed down by rhyme competitors (e.g., *candle-sandal*) than were younger adults. Since there was no such age effect in the resolution of onset competition (e.g., *candle-candy*), the authors argue that older adults benefit from the additional time to “catch up” with younger listeners when the contrast comes late in the word.

When it comes to word frequency effects and aging, the visual word processing literature indicates that word frequency tends to have a stronger influence on older adult readers compared to younger readers (Balota et al., 2004; Spieler & Balota, 2000). Revill & Spieler (2012) used the visual world paradigm to determine whether this pattern also characterized spoken word recognition. In their study, visual arrays included high- and low-frequency targets as well as high- and low-frequency competitors. Their results showed that older adults were more likely than young adults to fixate on high-frequency competitors. However, the same study showed only a marginal effect of frequency on target recognition. Importantly, degrading the signal for young adults in Revill & Spieler (2012) did not increase their fixations to high-frequency competitors, suggesting that the source of that effect for older listeners is not hearing loss, but rather changes in the cognitive processes associated with word recognition. This result contrasts with studies that have shown that presenting acoustically-degraded speech to young adults can reduce or eliminate differences between younger and older adults, suggesting that peripheral distortion is at the heart of age differences in speech perception (Ben-David et al., 2009; Pichora-Fuller et al., 2007).

Rather than focusing on competition between targets and displayed competitors, our current study focuses on how lexical frequency affects the time course of word recognition, even when there are no competitors in the visual display. The effects of word frequency on young adults were first examined using eye tracking by Dahan et al. (2001). In their first experiment, phonologically-related words were presented together in visual arrays, and the ones that were higher in frequency were looked at by listeners more quickly than the lower-frequency competitors. In their second experiment, each target word (which was either high or low frequency) was presented with three phonologically unrelated words. Listeners looked to high frequency words more quickly than low frequency words, even in the absence of phonologically similar competitors. Similarly, Magnuson et al. (2003) showed a main effect of frequency in an eye tracking study using an artificial lexicon, and Magnuson et al. (2007) found that fixation proportions to high frequency words were greater than low frequency words, even without related competitors in the display. In this case, the general advantage for high frequency words did not depend on time (i.e., listeners looked to high frequency words more than

low frequency words across the entire trial). In the current study, we aimed to replicate the general finding that high-frequency words are looked to earlier than low-frequency words and extend it to investigate the effects of aging and noisy listening environments on the temporal dynamics of spoken word recognition.

We consider two broad sets of processes that contribute to correct word identification. First, auditory processes deal with the accumulation of sensory evidence for a particular word. At the beginning of a word, not enough information has been processed to correctly identify it; when the word is complete, the maximum amount of sensory evidence is available. However, differentiating acoustic characteristics and contextual information usually allow listeners to tell what word has been presented before the end of the word (Grosjean, 1980; Tyler & Wessels, 1983; Wingfield et al., 1991). We expect that degrading the acoustic signal—for example, through the addition of background noise—will generally slow this process. Complementing auditory processes are cognitive factors that are important for both inhibiting responses that were initially activated but are no longer consistent with the acoustic input, and for selecting the correct word. Thus, two listeners with access to identical acoustic input may differ in spoken word recognition due to individual differences in their ability to select the appropriate word from among the possible competitors. Because of age-related changes in both hearing and cognition (e.g., general slowing), we expect older adults to show slower word recognition than young adults.

In the current study we examined spoken word recognition by young and older adults in the absence of phonological competitors among the visually-presented foils in order to focus on the speed of target activation for these two age groups. In this paradigm, a greater reliance on frequency in word recognition for older adults would predict a significant interaction between age and frequency, such that word frequency has a larger effect on the dynamics of word recognition for older than young adults.

Materials and Methods

Stimuli are available at <https://osf.io/5kuct/>. All materials and methods were approved by the Institutional Review Board at Washington University in St. Louis.

Materials

Two hundred words were used for the experiment: 25 low-frequency targets (Log Freq HAL¹ range of 5.1–6.8); 25 high-frequency targets (Log Freq HAL range of 10.0–11.9); and 150 mid-frequency distractors (Log Freq HAL 6.9–9.3). All words were closed monosyllables that referred to imageable nouns and were matched for phonological neighborhood density. For each word, a color image on a white background was found online (200 × 200 pixels). Three distractor words were pseudo-randomly grouped with each of the critical words ensuring that none of the distractors were phonological neighbors of the targets or were obviously related to the target semantically (as judged by the authors). Distractors sharing the same phonological onset as the critical word were also avoided. Fifty experimental displays

¹ The Hyperspace Analogue to Language (HAL) frequency norms based on the HAL corpus (Lund & Burgess, 1996, <https://paperpile.com/c/Cz25Ht/dB3d>), which consists of approximately 131 million words gathered across 3,000 Usenet newsgroups during February 1995. The log-transformed HAL frequency norms were used here.

were created out of these groups, with each of the four pictures in a different quadrant of the computer screen. The location of the target in each trial was randomized once and that location was used for all participants. The order of the trials was also randomized once, with this same order used for all participants. Consistency across participants was prioritized in order to facilitate analyses of individual differences.

Each display occurred with the spoken instructions “Click on the _____”. Recordings were made by an American male from the Midwest.² A single 1000 ms recording was used for the carrier phrase, and recordings of each target word were appended to it. The pictures appeared at the onset of the carrier phrase. Half of the participants heard the stimuli in quiet, while the other half heard them in steady speech-shaped noise (created using the long-term average spectrum of the target word stimuli) at a signal-to-noise ratio (SNR) of +3 dB.

Participants

Participants were 41 young adults aged 18–25 years (25 female, $M = 21.2$, $SD = 1.8$) and 39 older adults aged 65–84 years (24 female, $M = 71.7$, $SD = 5.1$). Three additional young adults and 3 additional older adults participated, but their data were excluded because of problems with eye tracking (e.g., the eye tracker could not locate their eye) or because they fell asleep (one older participant). We selected a sample size that was larger than those used in similar eye tracking studies (Reville & Spieler, 2012 had 16 per age group; Dahan et al. (2001) had 18 in their group). Young adult participants were recruited from the undergraduate psychology pool at Washington University in St. Louis and received course credit or \$10/hour for their participation. Older adults were recruited from the St. Louis community and were paid for their participation. All participants were community-dwelling, native English speakers, had self-reported normal or corrected-to-normal vision, did not use hearing aids, and were not color blind. All older adults scored at least 25 on the Mini Mental State Examination ($M = 29.0$, $SD = 1.6$) and had an average of 15.8 years of education ($SD = 2.4$). Young adults had a mean of 14.9 years of education ($SD = 1.4$) and were not administered the Mini Mental State Examination.

All participants were tested on vocabulary knowledge and hearing acuity. Vocabulary knowledge was assessed using the vocabulary subtest of the Wechsler Adult Intelligence Scale (Wechsler, 2008). To determine hearing acuity, pure-tone air-conduction thresholds were determined at 250, 500, 1000, 2000, 4000, and 8000 Hz. A pure-tone average (PTA) was calculated for each listener in each ear by averaging the thresholds at 500, 1000, and 2000 Hz. Group data are provided in Figure 1.

Unpaired *t*-tests showed statistically significant differences between the groups for hearing acuity in both ears (Left: $t = 8.93$, $df = 47.84$, $p < .001$; Right: $t = 8.64$, $df = 64.82$, $p < .001$). Their difference in years of education was marginal ($t = 1.93$, $df = 60.77$, $p = 0.06$), and there was no significant difference in vocabulary ($t = 0.98$, $df = 75.62$, $p = 0.33$).

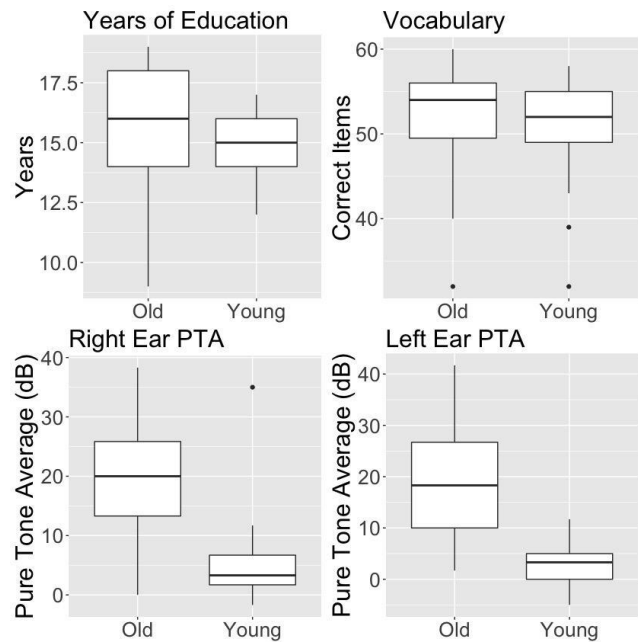


Figure 1. Education, vocabulary, and hearing information for young and older adult participants.

Procedure

Participants were tested individually in a sound-attenuated booth. They were instructed that a fixation cross would appear in the center of the computer screen at the beginning of every trial. When ready, they should click on the cross. Upon clicking the cross, an experimental array with a picture in each of the four corners would appear on the screen and the phrase “Click on the [TARGET]” would be heard through the speakers at a comfortable level. Using a mouse, the participant was instructed to move the cursor to the appropriate picture and click. No instructions were given regarding speed of response. A fixation cross would then appear, which they clicked to begin the next trial. Eye movements were tracked with a Tobii X120 eye tracker controlled by LabView 6.2 at a rate of 60 samples per second. Participants sat 0.5 meters from the screen, and a nine-point calibration procedure was conducted before testing began. Auditory stimuli were presented through a calibrated Madsen Auricle audiometer using two loudspeakers each approximately 1 meter from the listener and oriented ± 45 degrees from the participants’ forward-looking position when facing the monitor.

Results

Data processing and statistical modelling

Data and analysis scripts are available from <https://osf.io/5kuct/>. Looks to the target were analyzed for the 1-second time window from 300 ms to 1300 ms after target word onset, and only trials in which the target was correctly recognized were included (accuracy was $> 98\%$ in all conditions). Each frame in the eye tracker output was cod-

² Additional recordings of words from high- versus low-density phonological neighborhoods are available on OSF as well. They were not used for the present experiment.

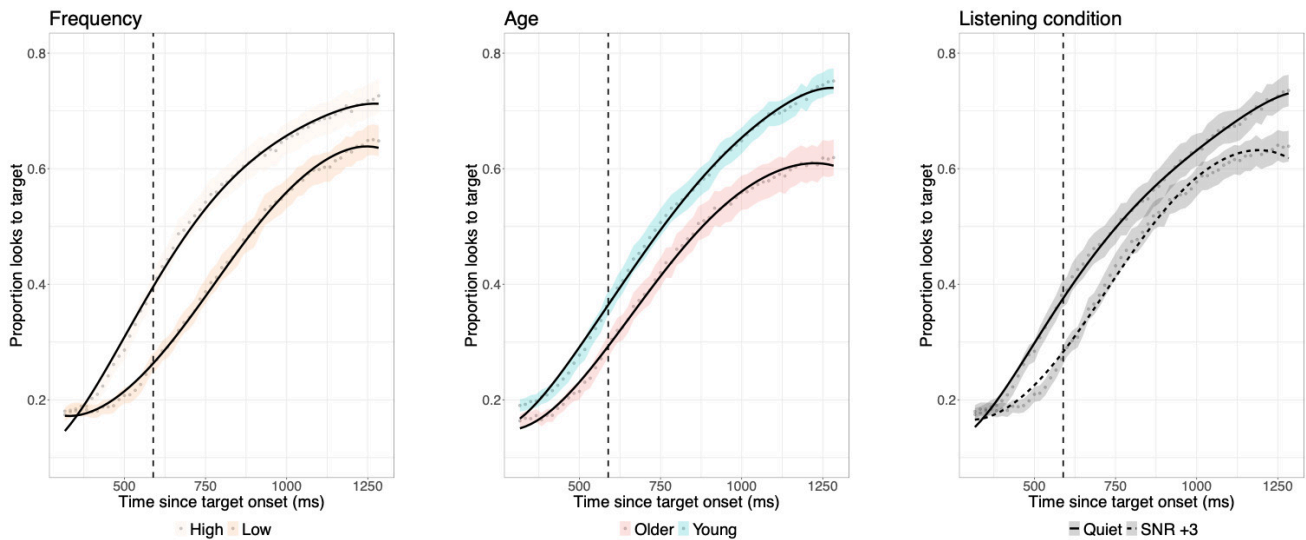


Figure 2. Fixed effects of word frequency, age, and noise (listening condition). Lines represent statistical model fits; dots represent raw averages; ribbons indicate standard error. Dashed vertical lines represent average target word offset. One second of data is presented, beginning at 300 ms after target onset.

ed as “1” if the eye was directed at the corner containing the target and “0” if it was not (i.e., frames where the eye was directed elsewhere or where the individual was blinking would be coded as 0).

We used logistic growth curve analysis (GCA) to model the by-participant target fixation data using the *lme4* package in R version 3.6.2. GCA is similar to polynomial regression, but controls for collinearity problems by orthogonalizing the polynomial time terms (Mirman, 2014). We modelled the time course with a third-order (cubic) orthogonal polynomial, which allowed us to model the sigmoidal shape of the raw data (i.e., two inflection points in the curves). Fixed effects were included for age (young vs. older), lexical frequency (high vs. low), and noise (quiet vs. noise), along with the interactions among these three factors. All three factors were sum coded (i.e., -1, 1). The model also included participant and participant-by-frequency random effects to capture both overall individual differences and differences in the effect of the frequency manipulation on each subject. Inclusion of all of the time terms in the random effects led to a singular model, so the structure was simplified minimally by removing the cubic time term from the subject random effects. (This term was involved in the two highest correlations among the random effects in the overfit model.)

Statistical significance was determined using p-values based on asymptotic Wald tests (the default in the *glmer* function from the *lme4* package in R). The full model output is included in the Appendix. Although all of the abovementioned factors were included in the model and in our considerations of statistical significance, we have plotted subsets of the effects to more clearly illustrate our results.

Main effects: age, noise, word frequency

Figure 2 shows the effects of age, noise, and word frequency. Visual inspection of the data suggests there were more fixations to the targets overall for high- vs. low-frequency words, for young adults vs. older adults, and for quiet vs. noisy stimuli. The overall effects of each of these factors were tested by comparing a statistical model that included random effects only to a model that also included

the fixed effect of interest. Age ($\chi^2=3.77$, $df=1$, $p=.05$) and frequency ($\chi^2=66.10$, $df=1$, $p<.001$) were significant predictors of overall looks to the target, but SNR was not ($\chi^2=.28$, $df=1$, $p=.60$).

Time course effects: age, noise, word frequency

The results of the full growth curve analysis (see Appendix) indicate that age significantly affected the linear time term ($\beta = -.59$, $SE = .26$, $z = -2.29$, $p = .02$), with young adults' fixations to the target increasing more rapidly than older adults' (see also Reville and Spieler, 2012). Noise significantly affected the cubic time term only ($\beta = .39$, $SE = .09$, $z = 4.56$, $p < .001$), as can be seen in the more curved shape of the model for noisy presentations. Frequency interacted with all three time terms (linear: $\beta = .36$, $SE = .15$, $z = 2.39$, $p = .02$; quadratic: $\beta = -.60$, $SE = .11$, $z = -5.46$, $p < .001$; cubic: $\beta = .40$, $SE = .09$, $z = 4.70$, $p < .001$).

Interactions

There were also interactions among these factors, shown in Figure 3. Age interacted significantly with listening condition on the quadratic time term ($\beta = -.47$, $SE = .15$, $z = -3.18$, $p = .001$) and with frequency on the cubic term ($\beta = -.17$, $SE = .09$, $z = -1.98$, $p < .05$). That is, the effects of noise and frequency differed for young and older adults.

Inspection of the data shows that the interaction between age and listening condition arises because there is a larger effect of noise on the young adults: the model of the young adults' looks to the target words continues to increase in quiet but flatten in noise while the model of the older adults' fixations flattened in both listening conditions. The interaction between age and frequency on the cubic term similarly arises because the young listeners' modelled fixations to high-frequency targets continue to increase while their looks to low-frequency targets flatten like the older adults'. In both cases, then, older adults' looking behavior across conditions looks more similar to young adults' behavior in challenging conditions (low frequency words, noisy environment).

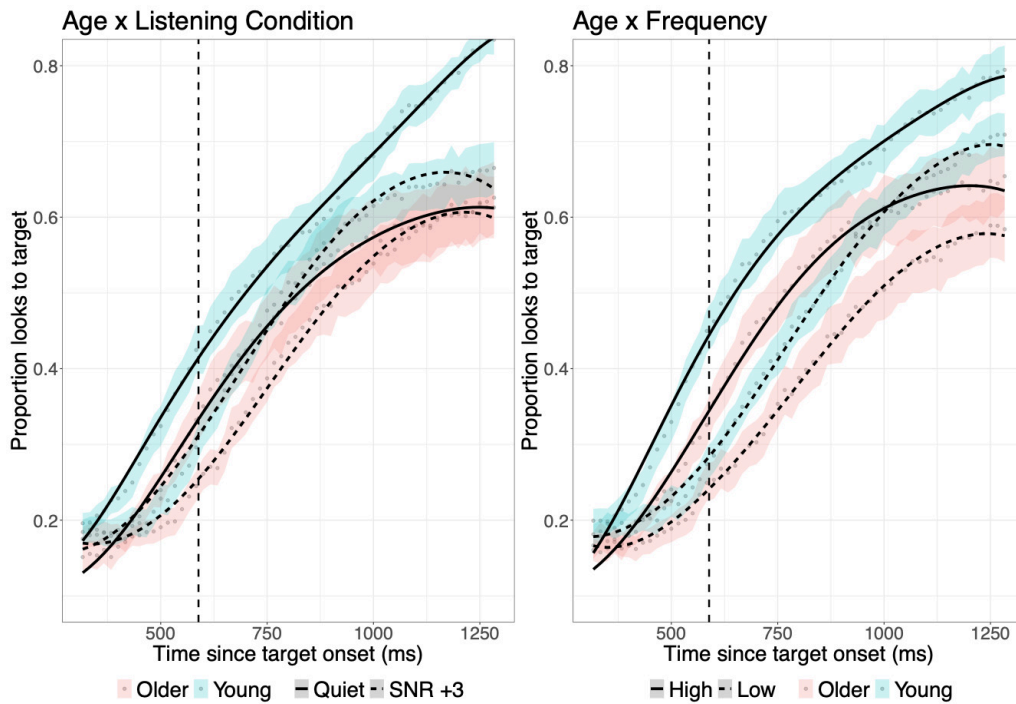


Figure 3. Significant two-way interactions between age and listening condition and between age and word frequency. Lines represent model fits; dots represent raw averages; ribbons indicate standard error.

Word frequency and noise also interacted with one another on the quadratic time term ($\beta = .34$, $SE = .11$, $z = 3.11$, $p < .01$). Visual inspection of this interaction (Figure 4) shows that the model fit for high-frequency words in quiet is shaped quite differently from the other conditions, such that looks to the target were still increasing in the analysis time window for that condition only. Because of this, the noise effect for high-frequency words appears stronger than for low-frequency words. It is also worth noting here that high frequency words in noise were recognized more quickly than low frequency words in quiet (i.e., the high-frequency data is all “above” the low-frequency data, even in noise.)

Finally, there was a three-way interaction among age, noise, and frequency on the linear time term ($\beta = -.30$, $SE = .15$, $z = -2.04$, $p = .04$). This interaction likely arises because although only age and word frequency affect the linear time term in general, younger and older adults differ more from one another in quiet than in noise.

Exploratory analyses of individual differences: hearing, education, and vocabulary

Hearing. Hearing acuity was not included in the general analysis because of its correlation with age (Cruickshanks et al., 1998; Homans et al., 2017). A follow-up analysis restricted to the older adults was conducted with better-ear PTA as a fixed factor to assess whether hearing acuity would predict the time course of lexical activation. Two models were tested: one that included PTA as a fixed factor but did not include its interactions with the other fixed factors; the other also included interactions with SNR and frequency. Despite the variability in hearing acuity among the older adults, there was no significant effect of better-ear PTA on the temporal dynamics of word recognition (i.e., there was no improvement to the model fit when PTA was added either on its own or with the interactions with other fixed factors).

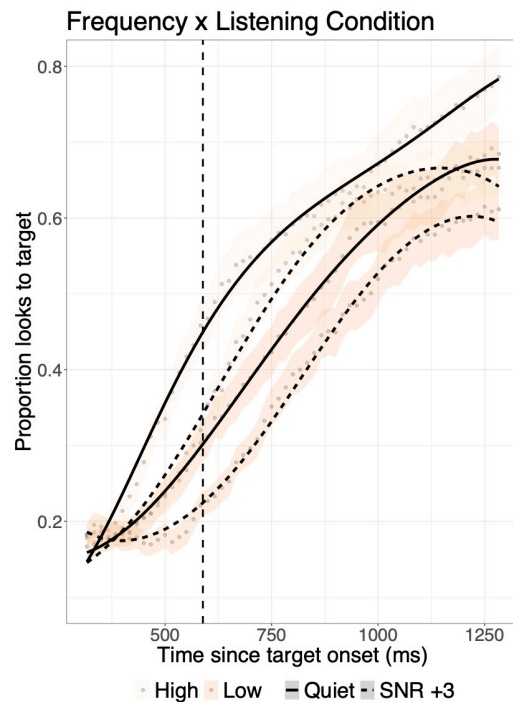


Figure 4. Significant interaction between word frequency and noise. Lines represent model fits; dots represent raw averages; ribbons indicate standard error.

Education. A parallel analysis on the older adult data was run on years of education, given that the older adults were marginally more educated than the younger adults. Adding education (without interactions between it and SNR or frequency) did improve the model fit in this case. However, neither the interaction between frequency and edu-

education nor the interaction between SNR and education improved the model. Thus it does not appear that years of education is modulating the frequency effect in older adults. Furthermore, closer inspection of the older adult data revealed that one participant had both the least education (> 2 standard deviations below the mean) and the lowest passing MMSE score in the group. With this person excluded from the analysis of the older adult data, education no longer significantly improved the fit of the statistical model.

Vocabulary. Vocabulary measures (WAIS scores) were included in this study primarily to ensure that our younger and older groups were matched, and indeed, there was no difference between the groups on this measure. An exploratory analysis was conducted, however, to investigate the potential role of vocabulary size in the time course of word recognition. As such, WAIS scores were entered into our main statistical model, along with all interactions among WAIS and our other fixed factors (age group, listening condition, and frequency). This model indicated that WAIS was a significant predictor, both overall and on the first two time terms. It also interacted with age group overall and on the first two time terms and there was a three-way interaction among vocabulary, age group, and listening condition overall and on the first two time terms. (Note that there was no interaction with frequency.)

To begin to understand this pattern of results, we conducted separate analyses of the two age groups that collapsed over frequency. For younger adults, vocabulary scores were significant overall ($\beta = -.31, z = -1.96, p = .05$) and there was a significant interaction between vocabulary scores and listening condition on the quadratic time term ($\beta = .47, z = 2.26, p = .02$). (Note that the estimate is negative, indicating that larger vocabulary is actually associated with fewer looks to the target.) For the older adults, vocabulary scores were a stronger predictor of looks to the target across the time course ($\beta = 1.35, SE = .25, z = 5.45, p < .001$), with higher vocabulary scores being associated with more looks to the target. They were also significant for all time terms (linear: $\beta = 2.48, SE = .59, z = 4.21, p < .001$; quadratic: $\beta = -.95, SE = .32, z = -2.98, p < .01$; cubic: $\beta = -.40, SE = .09, z = -4.50, p < .001$) and they interacted with listening condition overall and on all three time terms (overall: $\beta = 1.08, SE = .25, z = 2.60, p < .001$; linear: $\beta = 1.53, SE = .59, z = 2.60, p < .01$; quadratic: $\beta = -.73, SE = .32, z = -2.30, p = .02$; cubic: $\beta = -.25, SE = .09, z = -2.84, p < .01$). For older adults in particular, therefore, it appears that vocabulary size is an important predictor of the time course of word recognition.

Picture-word ratings

The words for this experiment were selected for their frequency characteristics, similar neighborhood densities, and basic phonological structure (closed monosyllables). However, it is important to note that the pictures we used were not normed for their prototypicality as referents of the words. To address this concern, we collected rating data online from 38 individuals (19 younger adults; 19 older adults). Participants were asked to rate the word-picture pairs on a 1-5 scale, with 1 being “word does not describe the object in the photo at all” and 5 being “describes the object in the photo very well”. The 50 target items were presented along with 114 fillers designed to range in terms of their prototypicality. The images for these fillers were taken from the eye tracking study as well. The high-frequency target items received an average score of 4.9 (range: 4.46-5.0) and the low-frequency items received an average score of 4.4 (range: 2.44-4.84; 3 of the 25 items received an average

rating below 4: *cot*, *hearth*, and *mitt*). It is unsurprising that low-frequency words received slightly lower match ratings, given that people may typically use higher-frequency labels for various items. For example, most people would use the word *fireplace* for the image that was presented for the target *hearth*. Rating data are available on OSF.

Discussion

It has long been observed that common words are recognized more rapidly and accurately than less common words (Goldinger et al., 1989; Howes, 1957; Marslen-Wilson, 1987). The current study replicated this lexical frequency effect in both young and older adult listeners using eye tracking with a visual array that did not include phonological competitors. Like Dahan et al. (2001) and others, the current data show a very early influence of lexical frequency on the word-recognition process.

While we cannot rule out the possibility that our observed frequency effects are influenced by the slightly poorer match between the low-frequency words and their images, there are several reasons to think this may not be a significant problem. First, given the closed-set nature of the task (only four images on the screen at a time as possible referents of a given item), there is never any doubt as to which image is being referred to by a particular stimulus. Furthermore, the participants could see the images before the onset of the target word, so they would already have scanned the array by the time they heard the target. Second, the individual listeners' data is collapsed over items for each condition, reducing the influence of those few items that were less well matched. Finally, it is quite possible that items whose names are low in frequency are simply less common items, such that they may capture *more* visual attention because of their relative novelty. Looking at the data (Figure 2, Figure 3) it seems possible that listeners looked at the low-frequency items slightly *more* immediately preceding target-word onset. If so, then low-frequency status might actually privilege pictures in terms of very early looks (an effect that would, if anything, reduce the effect of slower recognition for low-frequency words).

While the frequency effect was present in both age groups, we also observed differences between the age groups. First of all, younger adults generally looked to images depicting target words more quickly than older adults. While Revill & Spieler (2012) found that allowing age to affect the linear coefficient improved the fit of their growth curve model for target fixations and Ben-David et al. (2009) found that older adults were slower than younger adults when they had to distinguish target words from rhyming alternatives, the current study is, to our knowledge, the first to show age-related slowing in the time course of word recognition using visual arrays that do not also include phonological competitors. This result thus bolsters the general conclusion that older adults are slower to recognize spoken words and is consistent with general slowing accounts of cognitive aging (Lima et al., 1991; Madden et al., 1993; Salthouse, 1985, 1996).

More important, perhaps, is the fact that we did not find evidence for the proposed greater reliance of older adults on word frequency during target word recognition. While there was a significant interaction between age and word frequency, it occurred on the cubic time term only and was driven by a greater difference between the high- and low-frequency model fits for *young* adults as compared to older adults. Thus, although Revill & Spieler (2012) found stronger competition from high-frequency distractors in

older adults, there is still little evidence supporting the hypothesis that older adults are more affected by the frequency of target words in auditory word recognition.

Interestingly, we also found that young adults in noise (+3 dB SNR) showed similar time courses of word recognition as older adults in quiet. To visualize this, Figure 5 shows the raw means and model fits for young adults in noise and older adults in quiet.

On one hand, this might suggest that age-related changes in spoken word recognition are primarily impacted by age-related changes in auditory processing (given that changing the acoustic demands can produce “older adult”-like performance in young adults). However, it is important to remember that understanding speech in noise also increases cognitive demand. Thus, young adults listening to speech in noise face increases in both acoustic and cognitive challenge compared to listening in quiet, which results in a slowing of spoken word recognition similar to that seen in normal aging.

Another pattern worth mentioning in these results is that the high frequency words in *noise* were still recognized more quickly than low-frequency words in *quiet* (see Figure 4), indicating that the frequency effect is quite robust (i.e., acoustic degradation of the high-frequency words did not slow them to the level of low-frequency words). We purposely selected a noise level at which listeners would still correctly identify target words for this study, but further manipulation of SNR is needed to better delineate the relative effects of noise and word frequency on the time course of word recognition.

In summary, we have shown that young and older adults’ spoken word recognition appears to be similarly affected by word frequency. Although we observe age differences when presenting materials in the same level of noise to both groups of listeners, adding noise to the young adults results in comparable patterns of lexical activation to the older adults in quiet. These findings are consistent with similar processes supporting spoken word recognition in young and older adults that are sensitive to both auditory and cognitive aspects of speech recognition.

Contributions

Contributed to conception and design: KVE, AD, NR, BS, MS, JP

Contributed to acquisition of data: AD, NR, BS

Contributed to analysis and interpretation of data: KVE, JP

Drafted and/or revised the article: KVE, JP

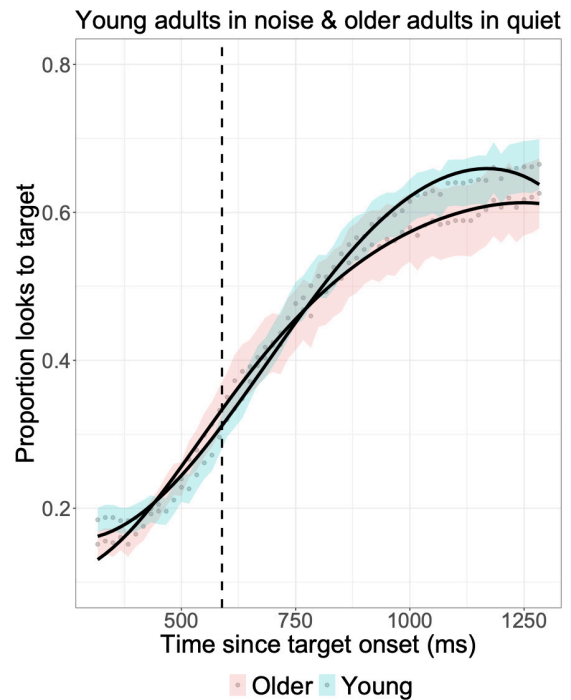


Figure 5. Young adults in a +3 dB SNR and older adults in quiet.

Approved the submitted version for publication: KVE, AD, BS, MS, JP

Acknowledgments

We thank Kirk Ballew and Maggie Zink for assistance with data collection.

Funding information

Research reported here was supported by grant R01DC014281 from the US National Institutes of Health (Jonathan Peelle, PI; Mitch Sommers and Kristin Van Engen, Co-Investigators).

Competing interests

None of the authors have any competing interests.

REFERENCES

- Alloppenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38(4), 419–439. <https://doi.org/10.1006/jmla.1997.2558>
- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, 133(2), 283–316. <https://doi.org/10.1037/0096-3445.133.2.283>
- Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E., & Schneider, B. A. (2009). Controlling for age related hearing loss can eliminate aging differences in lexical competition: Evidence from eye-tracking as an online measurement of age and noise effects on listening. *Canadian Acoustics*, 37(3), 150–151.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 6(1), 84–107. [https://doi.org/10.1016/0010-0285\(74\)90005-x](https://doi.org/10.1016/0010-0285(74)90005-x)
- Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E., Klein, R., Mares-Perlman, J. A., & Nondahl, D. M. (1998). Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. *The Epidemiology of Hearing Loss Study. American Journal of Epidemiology*, 148(9), 879–886.
- 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>
- 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)
- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, 28(4), 267–283. <https://doi.org/10.3758/bf03204386>
- Homans, N. C., Metselaar, R. M., Dingemanse, J. G., van der Schroeff, M. P., Brocaar, M. P., Wieringa, M. H., Baatenburg de Jong, R. J., Hofman, A., & Goedegebure, A. (2017). Prevalence of age-related hearing loss, including sex differences, in older adults in a large cohort study. *The Laryngoscope*, 127(3), 725–730. <https://doi.org/10.1002/lary.26150>
- Howes, D. (1957). On the Relation between the Intelligibility and Frequency of Occurrence of English Words. *The Journal of the Acoustical Society of America*, 29(2), 296–305. <https://doi.org/10.1121/1.1908862>
- Humes, L. E., Busey, T. A., Craig, J., & Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Attention, Perception, & Psychophysics*, 75(3), 508–524. <https://doi.org/10.3758/s13414-012-0406-9>
- Lima, S. D., Hale, S., & Myerson, J. (1991). How general is general slowing? Evidence from the lexical domain. *Psychology and Aging*, 6(3), 416–425. <https://doi.org/10.1037/0882-7974.6.3.416>
- 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>
- Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, 28(2), 203–208. <https://doi.org/10.3758/bf03204766>
- MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, 89(5), 483–506. <https://doi.org/10.1037/0033-295x.89.5.483>
- MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. Springer-Verlag.
- Madden, D. J., Pierce, T. W., & Allen, P. A. (1993). Age-related slowing and the time course of semantic priming in visual word identification. *Psychology and Aging*, 8(4), 490–507. <https://doi.org/10.1037/0882-7974.8.4.490>
- 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. <https://doi.org/10.1080/03640210709336987>

- Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General*, *132*(2), 202–227. <https://doi.org/10.1037/0096-3445.132.2.202>
- 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., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1–71. [https://doi.org/10.1016/0010-0277\(80\)90015-3](https://doi.org/10.1016/0010-0277(80)90015-3)
- Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R* (1st ed.). Chapman and Hall/CRC.
- 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>
- Peelle, J. E. (2018). Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear and Hearing*, *39*(2), 204–214. <https://doi.org/10.1097/aud.0000000000000494>
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, *37*(Suppl 1), 5S-27S. <http://doi.org/10.1097/aud.0000000000000312>
- Pichora-Fuller, M. K., Schneider, B. A., MacDonald, E., Pass, H. E., & Brown, S. (2007). Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. *Hearing Research*, *223*(1–2), 114–121. <https://doi.org/10.1016/j.heares.2006.10.009>
- 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>
- Salthouse, T. A. (1985). Speed of behavior and its implications for cognition. In *Handbook of the Psychology of Aging* (2nd ed., Vol. 2, pp. 400–426).
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. <https://doi.org/10.1037/0033-295x.103.3.403>
- 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>
- Spieler, D. H., & Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychology and Aging*, *15*(2), 225–231. <https://doi.org/10.1037/0882-7974.15.2.225>
- Tyler, L. K., & Wessels, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception & Psychophysics*, *34*(5), 409–420. <https://doi.org/10.3758/bf03203056>
- Wechsler, D. (2008). *Wechsler adult intelligence scale-Fourth Edition (WAIS-IV)*. NCS Pearson, 22, 498.
- Wingfield, A., Aberdeen, J. S., & Stine, E. A. L. (1991). Word onset gating and linguistic context in spoken word recognition by young and elderly adults. *Journal of Gerontology*, *46*(3), P127–P129. <https://doi.org/10.1093/geronj/46.3.p127>

SUPPLEMENTARY MATERIALS

Appendix: R output for growth curve analysis

Download: https://collabra.scholasticahq.com/article/17247-effects-of-age-word-frequency-and-noise-on-the-time-course-of-spoken-word-recognition/attachment/47366.txt?auth_token=xceevtp8SizOoFdhPz73
