Efficiency of spoken word recognition slows across the adult lifespan

Sarah E. Colby[1,2] and Bob McMurray[1,2,3,4]

[1]Department of Psychological and Brain Sciences
[2]Department of Otolaryngology – Head and Neck Surgery
[3]Department of Communication Sciences and Disorders
[4]Department of Linguistics
University of Iowa
Iowa City, IA 52242

Keywords: Word recognition, aging, lexical competition
Preprint: https://psyarxiv.com/gcj76/ (CC-By Attribution 4.0 International)

Corresponding Author
Sarah Colby, Ph.D.
Psychological and Brain Sciences Building, G60
340 Iowa Ave., Iowa City, IA 52242
Email: sarah-colby@uiowa.edu
Phone: 319-353-6462

Acknowledgements

This work was supported by National Institutes of Health Grant P50 000242 to B. Gantz and B. McMurray and DC 0008089 to B. McMurray. The authors would like to thank Kristin Rooff and Marissa Huffman for their help with data collection; and Francis X. Smith for assistance in developing the version of the VWP used here.

Abstract

Spoken word recognition is a critical hub during language processing, linking hearing and perception to meaning and syntax. Words must be recognized quickly and efficiently as speech unfolds to be successfully integrated into conversation. This makes word recognition a computationally challenging process even for young, normal hearing adults. Older adults often experience declines in hearing and cognition, which could be linked by age-related declines in the cognitive processes specific to word recognition. However, it is unclear whether changes in word recognition across the lifespan can be accounted for by hearing or domain-general cognition. Participants (N = 107) responded to spoken words in a Visual World Paradigm task while their eyes were tracked to assess the real-time dynamics of word recognition. We examined several indices of word recognition from early adolescence through older adulthood (ages 11 – 78). The timing and proportion of eye fixations to target and competitor images reveals that spoken word recognition became more efficient through age 25 and began to slow in middle age, accompanied by declines in the ability to resolve competition (e.g., suppressing *sandwich* to recognize *sandal*). There was a unique effect of age even after accounting for differences in inhibitory control, processing speed, and hearing thresholds. This suggests a limited age range where listeners are peak performers.

Introduction

Typical (and atypical) aging is associated with two major changes that impact quality of life: hearing loss and cognitive decline. Both impact social functioning, which in turn predicts physical, mental and cognitive health (DuPertuis et al., 2001). A critical link between these is speech recognition. When older adults find it challenging to recognize speech, communication becomes difficult or tiring and they may withdraw from social situations. This isolation leads to psychosocial issues and potentially to cognitive decline in a downward spiral (Lin et al., 2011).

The ability to understand spoken language is likely a critical mediator between hearing and social engagement. Spoken language comprehension lies at the intersection of auditory processing and cognition and could mediate the relationship between auditory and cognitive declines. Older adults with poorer hearing may be able to rely on top-down cognitive strategies to compensate for peripheral auditory declines, while older adults with weaker cognitive processing may struggle to keep up in conversation. Supporting this, older adults often report difficulty understanding speech even when they can hear it (Humes et al., 2020; Pichora-Fuller, 2003), suggesting more complex changes to auditory cognition and language processing than simply declines to hearing detection thresholds (such as changes to processes like temporal processing, auditory streaming, and higher-level cognitive factors involved in language processing). Conversely, stronger language skills could offer resilience to the effects of hearing loss, allowing people to benefit from a more socially enriched environment despite some hearing loss. Language processing is highly plastic and can adapt to novel circumstances (Clayards et al., 2008; Kapnoula & McMurray, 2016) even into older adulthood (Colby et al., 2018). This raises the possibility of interventions to improve the efficiency of spoken language processing, which in turn could improve psychosocial outcomes (Humes et al., 2020).

There are several known changes to language processing that accompany aging. Older adults are slower to process complex sentences (Payne et al., 2014; Waters & Caplan, 2001) and have more difficulty comprehending longer discourse (Schneider et al., 2002). When processing speech, older adults also rely more on top-down information, showing larger effects of higher-level lexical knowledge (e.g., whether an item is a word or a nonword, Mattys & Scharenborg, 2014) and sentential context (Pichora-Fuller, 2008). They also show greater reliance on working memory for complex sentences (Payne et al., 2014), suggesting an ability to recruit domain-general resources to offset peripheral declines. However, as these domain-general resources themselves may also decline with age, it is not clear that these represent successful strategies.

Along with these changes in performance, older adults report more fatigue when comprehending language, particularly in noise, and exert more cognitive effort towards the goal of language processing (Ayasse et al., 2017; Kuchinsky et al., 2013; Pichora-Fuller, 2003; Tun et al., 2009). The increased effort and fatigue are commonly thought to be the result of recruiting additional cognitive resources to offset peripheral auditory declines. Additional resources may also be recruited to compensate for any changes to the underlying mechanisms of speech perception itself (Phillips, 2016; Pichora-Fuller et al., 2016).

The overall picture appears to be one in which language performance declines with age. However, the overall trajectory of these age-related changes remains unclear. Studies often compare older adults as a discrete group to younger adults, leaving middle-aged adults and the continuous effect of age over the lifespan unexplored. Moreover, existing work often focuses on factors like performance and effort – important outcomes in their own right – while leaving open the specific changes to the underlying mechanisms of language processing. Thus, the present study investigated continuous age-related differences across the lifespan in the mechanisms of real-time processing for one important aspect of language: spoken word recognition.

Word recognition across the lifespan

Vocabulary knowledge is one aspect of language that remains a strength in older adults (Kavé & Halamish, 2015), and vocabulary is widely seen as being preserved throughout the lifespan. However, vocabulary represents crystallized knowledge and does not reflect how language is used and processed by an individual (fluid ability). In fact, despite crystallized abilities remaining a strength in older adulthood, there is still a relationship between fluid and crystallized abilities, with individuals who have the largest declines in fluid abilities showing the smallest gains (or even losses) to crystallized abilities (Tucker-Drob et al., 2022). Thus, there is a need to investigate the cognitive mechanisms underpinning word recognition directly. To this end, the present study examines spoken word recognition across the adult lifespan, from early adolescence to older adulthood, and some of the possible sources of variation that might affect word recognition in older adults (hearing, processing speed, and inhibitory control).

Word recognition is cognitively challenging, tapping language-specific processes that are well understood in younger adults. Research with young, normal hearing (NH) listeners suggests lexical access – the process of recognizing a word and activating its meaning – starts by immediately activating an array of candidates as soon as any amount of input is heard. For example, after hearing *ro-* at the onset of *rocket*, listeners activate a range of words like *rocket*, *rocker*, *rock* and *robin*. As the input unfolds, a competition process plays out: candidates that best match the input remain active, and others are ruled out, until one remains (*rocket*). This is not entirely driven by bottom up match to the input: competition is observed for words that mismatch at different times (the bone in trombone; Luce & Cluff, 1998) or even words whose sounds are out of order (cat after hearing tack; Dufour & Grainger, 2019; Toscano et al., 2013). As lexical items compete, more active items inhibit less active competitors (Dahan et al., 2001; Luce & Pisoni, 1998)—once *rocket* becomes more active than *rocker,* it actively suppresses *rocker*. This ultimately speeds the process, and helps competition resolve more completely. Thus, lexical competition is a cognitive process that balances speed, efficiency, and flexibility to achieve robust speech recognition (McMurray et al., 2022). Importantly, similar competition mechanisms have been invoked at every level of language comprehension, from individual phonemes to discourse (Altmann, 1998; Dell & O'Seaghdha, 1992; Lotto & Holt, 2011; Weber &

Scharenborg, 2012). Thus, word recognition can serve as a model system for understanding how aging affects a key mechanism that underlies much of language cognition more broadly.

Older adults have more difficulty managing competition during word recognition; they are slower to recognize words with many competitors (Sommers & Danielson, 1999) and have difficulty suppressing competitors (Dey et al., 2017). Without sentence context, older adults require more information to recognize words (Lash et al., 2013; Wingfield et al., 1991). Most of this work, however, relies on accuracy, response time, and self-report measures that do not capture online processing, and none has examined the full lifespan nor controlled for domain-general cognitive abilities.

There are well-established methods that can trace the time course of word recognition on the order of milliseconds in a way that offers a close match to computational models (McMurray et al., 2010; Mirman et al., 2008). This allows us to identify age-related changes in not just how well listeners recognize words, but to characterize how the fundamental process changes with aging. In the Visual World Paradigm (VWP), participants hear a word and match it to one of four pictures. For example, participants might hear *rocket* with a display containing pictures of the target (*rocket*), an onset competitor (*rocker*, a cohort), a rhyme (*pocket*), and an unrelated word (*bubble*). Eye movements are recorded while they do this task. Participants must find the target to respond; this requires them to launch fixations while lexical access is unfolding. Fixations to each object over time capture how much that item is considered. For example, in Figure 1A, after hearing *rocket*, participants fixate both the *rocket* and the *rocker*, but by around 500 msec they suppress *rocker* (though they may briefly fixate *pocket*). By around one second, they have fully committed to the target.

The VWP has been used to assess individual differences, development, disorders and challenging listening conditions (Brouwer & Bradlow, 2016; Farris-Trimble et al., 2014; McMurray et al., 2010; Rigler et al., 2015). These illustrate how word recognition is not unidimensional (McMurray et al., 2022, 2023): typical development (through adolescence), for

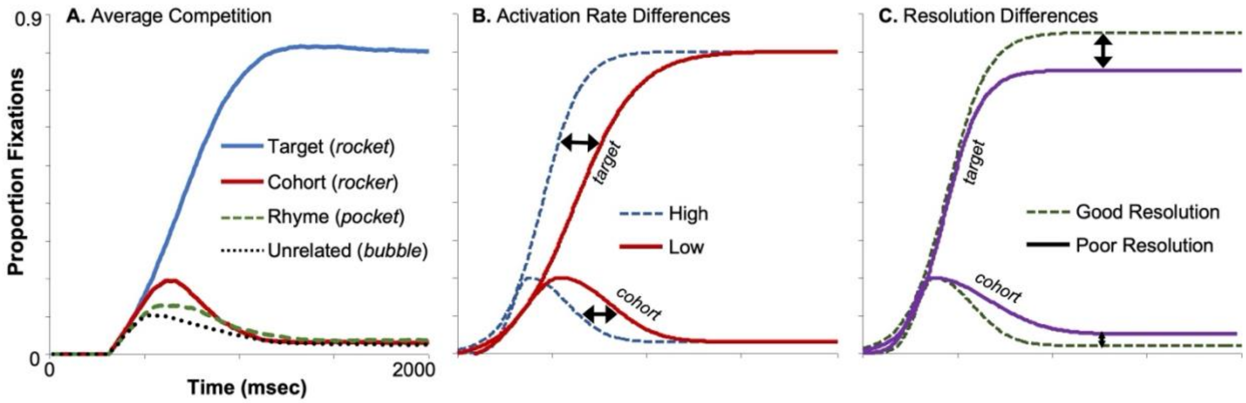

Figure 1. Average time course of fixations to different image types. A) Proportion fixations to target, cohort, rhyme, and unrelated images from the spoken word VWP experiment; B) Differences in the activation rate appear across multiple components of the curves and are associated with typical development; C) Differences in resolution affect the asymptotes and have been linked to disrupted language or challenging listening conditions.

example, is linked to changes in the speed by which targets are activated and competitors are suppressed (Figure 1B; Rigler et al., 2015), while language disorders and challenging listening has its effects on the asymptotes as the ultimate degree of competitor suppression (Figure 1C; Brouwer & Bradlow, 2016; McMurray et al., 2010). This creates the opportunity to characterize aging in terms of both the efficiency and disruption of the underlying processes of spoken word recognition (among other dimensions).

Several studies have used the VWP to characterize age-related changes in word recognition. Older adults are slower to fixate targets (Van Engen et al., 2020), experience more competition from high-frequency words (Revill & Spieler, 2012), and have more difficulty distinguishing targets from rhymes in noise (Ben-David et al., 2011). These effects are generally seen even after excluding trials where the participant clicks the incorrect image, suggesting differences in the process of word recognition, controlling for accuracy. This work supports the idea that they struggle to manage competition. However, these studies are limited in several regards. First, they only assess one type of competitor (either rhyme or cohort competitors; Ben-David et al., 2011; Revill & Spieler, 2012) or in the case of Van Engen et al. (2020), display no related competitors and only picture the target with unrelated distractors. In the latter case, no claims can be made about the nature of lexical competition, only about its speed. In addition, existing work has not examined the full lifespan, treating age in terms of two discrete groups with little to no sampling from middle age. Thus, it cannot characterize the complete lifespan.

Importantly, most of these studies have not accounted for age-related changes in decision-making, speed of processing, and/or visual search that could also impact performance in the VWP. A few studies have attempted to account for these differences using analyses *within* the VWP. For example, Ben-David et al. (2011) used the difference between looks to the target and competitor to compensate for overall differences in the amount of looks made by older and younger adults. However, this does not control for other processes like differences in visual search, decision making, or overall speed of processing.

There are non-linguistic versions of the VWP that are closely matched to the word recognition task and which can be used to estimate the summed domain-general processes that may contribute to performance in the VWP (Farris-Trimble & McMurray, 2013). These tasks capture all of the decision making, visual search[1], and oculomotor demands of the VWP but in a completely non-linguistic task, and they yield eye-movement indices that are analogous to the VWP. They can thus pinpoint the contribution of aging specifically on the dynamics of word recognition by accounting for virtually all of the non-linguistic aspects of cognition relevant to VWP performance. Together, the theoretical and methodological paradigms of spoken word recognition offer an unparalleled view of an aging language process.

One important counter-hypothesis is that the differences in language processing in aging may derive entirely from sensory or perceptual differences. Indeed, several lines of research support this. For example, Schneider et al. (2005) investigated the role of time compressed speech. They found that when speech is sped up in a way that does not degrade the signal,

---

[1] Many variants of the VWP, including the one used here, use a pre-scan period – showing the pictures prior to the auditory stimulus to minimize demands on visual search (c.f., Apfelbaum et al., 2021). Nonetheless, this does not eliminate more general factors like decision speed, or more specific ones like oculomotor control that could vary with age that can be captured by the non-linguistic VWP.

older and younger adults recognize sentences similarly. However, if speech is sped up in a way that affects the clarity of the signal, older adults show a larger decrement (Schneider et al., 2005). Two studies of speech in noise also suggest that there may be few differences. Ben-David et al. (2011) found that when older adults were presented with an easier signal-to-noise ratio, there was no difference in the relative looks to targets and competitors between older and younger adults (though differences emerged at lower SNRs). Similarly, younger adults at a relatively easier +3 dB SNR show a similar pattern of looks to older adults in quiet (Van Engen et al., 2020).

Together these results suggest that there may be few differences in language processing and word recognition if perceptual issues are minimized and domain-general cognitive variables are controlled for. However, there are several reasons to reinvestigate this claim. First, Schneider et al. (2005) relied on accuracy measures, whereas a real-time measure may be more sensitive. Second, while the Van Engen and Ben-David studies used the VWP, Van Engen et al. (2020) did not include any competitors on the screen (potentially making the measure less sensitive), and Ben-David et al. (2011) only had 32 critical trials and a relatively small sample, raising questions of reliability.

In this light, the present study restricts our sample to those with relatively preserved hearing, uses the pure-tone average as a covariate, and tests individual words spoken in quiet. Thus, our approach is particularly conservative—if age-related differences derive from purely perceptual concerns, we should see few effects. At the same time, our version of the VWP had a large number of critical trials and underwent extensive psychometric evaluation. Thus, our paradigm was optimally suited to detect such a difference.

Similarly, we considered the possibility that domain-general inhibitory control—which is known to decline with age—may specifically play a role in the resolution of lexical competition. This is particularly relevant for an investigation of aging, as older adults are known to have more difficulty suppressing competitors once they have been activated (Campbell & Hasher, 2018;

Dey et al., 2017). Here it is important to contrast two forms of inhibition. In classic models of word recognition, lexical candidates are thought to directly inhibit one another as they compete for activation via lateral inhibition (Dahan et al., 2001; McClelland & Elman, 1986). This is likely to be distinct from more general forms of inhibitory control (e.g., as measured with Stroop or Flanker tasks) and two studies have not found correlations between lateral inhibition in word recognition and domain-general measures of control (Blomquist & McMurray, 2023; Kapnoula & McMurray, 2021). One study that directly addressed whether general inhibitory control is involved in word recognition found that individual inhibitory ability predicted competitor resolution in a mouse-tracking study in which the target objects were Mandarin characters (Zhao et al., 2022). This suggests that inhibitory control may play a role in the late stages of lexical competition when competitors are being suppressed. However, here the relevant item to be suppressed was at the output stage (the two characters were visually similar) not in the phonological competitors to spoken words. At the same time, Dey and Sommers (2015) found a relationship between inhibitory control and accuracy recognizing audiovisual speech, but only in older adults recognizing words with high competition (and not easier words with fewer competitors), suggesting that older, but not younger, adults may recruit domain-general inhibition when speech perception becomes challenging. However, it is unclear if general forms of inhibitory control are involved in suppressing phonological competitors more broadly (and this of course may change with age).

The Present Study

The present study adopted a comprehensive approach to the lifespan development of spoken word recognition. We tested a large sample of individuals with ages spanning 11-78, evenly sampled across the lifespan with a standard spoken word recognition version of the VWP. We present two different types of phonological competitors along with the target to get a more complete picture of the nature of lexical competition.

Critically, we deployed a visual analogue of the VWP (Farris-Trimble & McMurray, 2013) in which participants matched a centrally-presented shape to an array of shapes, one of which served as a competitor by matching the target in color (but not shape). Fixations in this task can be analyzed similarly to the standard VWP to yield a timecourse of target and competitor looking that is analogous to what is obtained from the spoken VWP. This task captures most, if not all, of the contributions of non-linguistic factors to the VWP. For example, it has similar visual search demands, reflects similar speed of processing demands, and it requires object recognition and saccade targeting. It also has similar demands on more abstract levels of cognition—maintaining instructions in working memory and sustained attention. Unlike standardized assessments of cognition or intelligence, the visual VWP is an identical task with the language modality removed, making it a better control task for factoring out differences in domain-general cognition as it precisely targets the types of cognition that are used in the VWP and yields indices that are of the same numerical form as the VWP (e.g., target slope, competitor peak). In addition to this task, we also estimated inhibitory control using a spatial Stroop task.

We address four research questions. First, we investigated the lifespan development of spoken word recognition: when does it peak and what is the timecourse of any decline? At the older end of the age range, we expected to see declines in fluid processing ability, despite the relative maintenance of crystallized vocabulary knowledge. At the earlier part of the lifespan, we extend prior work suggesting development up to ages 16-18 (Apfelbaum et al., 2022; Rigler et al., 2015) to ask how long into young adulthood developmental changes are visible.

Second, we investigated the nature of any age-related changes. If age primarily affects efficiency, we should see differences in activation rate (Figure 1B). Based on prior work this should be the primary change in the younger (developmental) portion of the sample, if development is ongoing (Rigler et al., 2015). It is unclear what should be observed with aging as prior work has not evaluated VWP with this level of specificity. In addition, if cognitive decline reflects instability in language processing, we should observe changes in resolution (Figure 1C) as is often seen in conditions like Developmental Language Disorder (McMurray et al., 2010) or for words acquired in a second language (Sarrett et al., 2021). We did not anticipate this would be observed in the adolescent range as prior work has shown only changes in activation rate.

Third, we ask if these changes hold over and above differences in the non-linguistic processes involved in the Visual World Paradigm. For example, if age-related changes primarily reflect domain-general speed of processing, then any age-related differences should be eliminated once we control for activation rate from the visual VWP. To the extent that age-related changes are unique to word recognition, we should see a continued effect of age, even controlling for the visual VWP. Note that this is a purposely conservative approach. Performance in the visual VWP likely reflects a mix of general processes and factors. Some of these are not relevant to word recognition (e.g., eye-movement control), while others may reflect domain-general factors that are highly relevant (speed of processing). Our hierarchical regression approach assigned any shared variance to the visual VWP (even though it may capture some of these processes).

Finally, we investigated the contributions of two standard factors: inhibitory control (as reflected in a Stroop task) and hearing thresholds. With respect to hearing, all our participants had what we operationally defined as age-typical hearing. This was assessed by pure-tone audiometry but with a slightly relaxed criterion for inclusion to accommodate some high-frequency hearing loss in older adults (see methods for details). Subtle differences in normal audiometry have been linked to things likes speech in noise performance (Holmes & Griffiths, 2019). Thus, we ask if hearing predicts word recognition and if age-related differences in word recognition persist after accounting for hearing.

Methods

Participants

Participants were recruited from the University of Iowa and surrounding communities. 111 participants between the ages of 11 – 78 participated in the study. Participants were evenly distributed across this age range, with at least 10 participants per decade of age. All participants were monolingual speakers of English and had no history of neurological or cognitive impairment. All participants had normal or corrected-to-normal vision, and no history of degenerative vision impairments (i.e., retinal degeneration, cataracts). There were no participants who used hearing-assistive devices (e.g., hearing aids, cochlear implants).

We adopted a slightly relaxed criterion for inclusion based on hearing in order to accommodate sub-clinical high frequency hearing loss that is typical in older adults. Here our goal was to broadly sample from the population and include variation in hearing that is typical of aging (which can then be assessed statistically) without testing people who would normally use a hearing aid or cochlear implant (as these can distort the auditory input). First, each subject's pure-tone detection thresholds were estimated at seven frequencies (250, 500, 1000, 2000, 4000, 6000, and 8000 Hz) using standard audiometry. Pure-tone average (PTA) was calculated as the average hearing threshold at the lower six frequencies. We did not include 8 kHz in our PTA calculation to allow for some variability in high-frequency hearing loss, which is typical of age-related hearing loss (see Supplemental Figure S1 for average audiograms by age group). Next, we excluded any participants who did not have a modified (six-tone) PTA of less than 30 dB HL in at least one ear (n = 4). Note that 20 dB HL is the ASHA and WHO standard for "normal hearing" (World Health Organization, 2021). While this criterion only affected the older end of our sample, we believed it necessary to have a stringent hearing requirement as we were primarily interested in how age affects language processing aside from declines to peripheral hearing thresholds. We operationally define this as "age typical hearing", and it yields a similar exclusion rate to the use of standards like the ISO standards (International Organization for Standardization, 2017). This left 107 individuals available for analyses (39 male, 68 female, M_{age} = 47.8 years, SD_{age} = 19.5, range = 11.2 – 78.1 years).

Due to missing data on subsets of tasks, the full multiple linear regressions contain data from 89 individuals, while other regressions included larger samples as noted. All recruitment and experimental protocols were approved by the University of Iowa's Institutional Review Board.

Procedure

All tasks were completed in one visit to the lab that lasted approximately 1.5 hours. Participants were seated in front of a 19" computer screen in a sound-attenuated booth. Auditory stimuli were presented over two loudspeakers placed approximately 1 meter in front of the participant at 30° and 330° azimuth. The session started with the informed consent process and audiogram. The experimental tasks were always completed in the following fixed order: 1) the spoken word VWP; 2) the spatial Stroop task; 3) the visual VWP. For both VWP tasks, eye movements were recorded using an Eyelink 1000 desktop-mounted eye tracker with chinrest (SR Research; Ontario, Canada) sampling at 500 Hz. Before each VWP task, the eye-tracker was calibrated using a 9-point calibration and during the tasks, drift corrects were performed every 30 trials.

Spoken Word VWP

Items. Stimuli sets were comprised of a target, onset (cohort) competitor, rhyme competitor, and unrelated item (e.g., *rocket, rocker, pocket, bubble*). There were 30 monosyllabic sets and 30 bisyllabic sets. These sets were developed over the course of a series of pilot studies intended to build a canonical VWP task. These pilot studies started with 120 sets which were developed and tested with 68 NH young adults. We then selected the 60 items with the most prototypical pattern of competition. The final 60 item sets were then tested for test-retest reliability in 29 young adults who completed the spoken word VWP task twice with a week delay. Test-retest correlations between our indices of interest were moderate to strong (Target activation rate: r = 0.75; Competitor resolution: r = 0.62; Peak Cohort Activation: r = 0.54), and generally exceeded prior estimates (Farris-Trimble & McMurray, 2013).

Stimuli. Auditory stimuli were recorded by a female monolingual speaker of English in a sound-attenuated room sampled at 44.1 kHz. Auditory tokens were edited to reduce noise and remove clicks. They were then amplitude normalized to 70 dB SPL. Visual stimuli were images from a commercial clipart database that were selected by a small focus group of students and edited to have a cohesive style using a standard lab protocol (McMurray et al., 2010). Images were all scaled to 300 x 300 pixels.

Design and Procedure. Each of the four items from a set was used as the auditory target once. In addition, one additional item from each set was randomly selected to serve as the target word on an additional trial to discourage participants from predicting the upcoming target word once the items in the display are visible (i.e., they cannot assume that once they have already heard *rocket, rocker* must be the target). This led to a total of 300 trials (60 sets x 5 targets/set). Image placement was pseudo-randomized across trials and participants, such that each image type was equally likely to appear in any quadrant.

On every trial, participants saw a blue circle in the middle of the computer screen with the four images corresponding to an item set in each of the corners. After 500 msec, the circle turned red, and they clicked on it to play the auditory stimulus. Participants were instructed to click on the image that best represents the auditory target. Their eye movements were recorded while they completed this task.

Data Processing. Eye movements (fixations, saccades, and blinks) were processed using EyeLink Analysis 4.12 (McMurray, 2019). Saccades and the subsequent fixations were combined into a single unit, a look, which started at the onset of the saccade and ended at the end of the following fixation. Looks were assigned to one of four regions of interest, which were the regions in which the images were displayed extended by 100 pixels. Looks were then identified as directed to one of the four image types (target, cohort, rhyme, unrelated) or to nothing. Any fixations launched before the onset of the target word (accounting for a 200 msec oculomotor delay) were ignored. Only trials where the correct target image was selected were included in further analyses as the logic of the analyses sought to identify differences in processing (given accurate word recognition), not differences in accuracy.

Accuracy was high for all ages. For participants younger than 32.4 years, mean accuracy was 99.3%. For participants between the ages of 32.4-53.7, it was 99.4%. For participants between the ages of 53.7-66.1, it was 99.5%, and for the oldest participants (≥66.1), accuracy was 99.1%[2]. All statistical analyses treat age as a continuous variable.

For each participant, the average fixations to each item type at each sampling point was calculated (every four msec). Nonlinear curves were then fit to each participant's data (Farris-Trimble & McMurray, 2013) using a constrained gradient descent that minimized the least squared error between the function and the data (McMurray, 2020). Target looks were fit with four-parameter logistic curves which include: the upper and lower asymptotes, the crossover (the point in time where the function switches) and the slope (the derivative at the crossover). Cohort, rhyme, and unrelated competitor curves were fit with an asymmetric Gaussian with six parameters: the initial and final asymptotes, the location and height of the peak, and the onset and offset slopes. We used the parameters from these functions to derive theoretically-meaningful indices that capture characteristics of each participant's looks.

Activation rate captures the speed of target activation and was indexed by *target timing*. Slope and crossover values were log-transformed and z-scored for standardization. Crossover was multiplied by -1 and then the two values were averaged into a composite score of timing (where higher = faster). *Competitor resolution* was the difference between the upper asymptote of target looks and the average of the later baselines of competitor and unrelated looks. This

---

[2] Note that these age groups reflect equal sized groups and were created solely for descriptive purposes to identify any gross age differences in accuracy and for visualization purposes (e.g., Figure 2).

reflects the final portion of the timecourse during which competitor and target fixations have reached their asymptotes. This was calculated separately for both cohort and rhyme competitors (i.e., average of cohort and unrelated baselines, average of rhyme and unrelated baselines). *Peak competitor activation* is the maximum of fixations at the inflection point of the cohort (or rhyme) looks.

Visual VWP

The visual VWP was a modified version of the task introduced in Farris-Trimble and McMurray (2013). It was designed to be analogous to the spoken word VWP, though with a condensed design as 1) we anticipated less variability among items; 2) there was not an a priori reason to examine differences in competitor types (e.g., cohorts vs. rhymes) since the stimuli did not unfold over time which allows us to use only a single competitor; and 3) we needed a shorter task to fit into the session.

Stimuli. Stimuli were 16 uncommon shapes (e.g., chevron, hourglass, trefoil) in 8 colors that were chosen to be color-blind friendly. Unlike the prior versions of this tasks, shapes were chosen that would not be easily named. Final images were 300 x 300 pixels.

Design and Procedure. Sixteen sets were constructed such that each comprised four shapes in two contrastive colors (e.g., lavender chevron, lavender teardrop, yellow trefoil, yellow cross). This allowed every trial to include a target and a competitor. For example, when the lavender chevron was the target the lavender teardrop was the competitor, and the two yellow objects were unrelated; however, when the yellow trefoil was the target, the yellow cross was the competitor, and the lavender items were unrelated.

Each shape and color appeared in four different sets. The same shape did not always appear as the same color, nor were the same two colors always paired in a set. Each item was used as the target three times for a total of 192 trials. Image placement was pseudo-randomized across trials and participants, such that the image type (target, competitor, unrelated) was equally likely to appear in any one quadrant.

Trials proceeded in a similar manner to the spoken word VWP. After clicking on the red circle, instead of hearing an auditory target, the target shape appeared for 100 msec in the center of the screen. The participant was instructed to click on the shape in the display that identically matched the target.

Data Processing. Eye movements were processed in the same way as for the spoken word VWP task described above. The same curve fitting procedure was also used for target and competitor looks.

Other tasks

Pure-tone audiometry. Hearing thresholds were established using a standard audiometric procedure. Starting at 25 dB HL, participants were instructed to raise their hand when they could hear a tone played to one ear over headphones. If the participant could hear the tone, amplitude was decreased by 10 dB. If they could not hear the tone, amplitude was increased by 5 dB. The lowest amplitude that participants consistently responded to was recorded as the threshold for that frequency. Both ears were tested at .25, .5, 1, 2, 4, 6, and 8 kHz. PTA was calculated as the average at .25 – 6 kHz (excluding 8 kHz). PTA from the better ear (i.e., the lower PTA of the two ears) was used as a predictor in analyses.

Spatial Stroop. Participants responded to the direction of an arrow on a computer screen as quickly and accurately as possible using the left and right arrow keys on a keyboard. On each trial, a fixation cross appeared for 200 msec before the arrow appeared. The arrows could appear on either the left or right side of the screen to create congruency or incongruency between the direction and the side of presentation. On congruent trials, the arrow pointed in the same direction as it was presented on the screen (e.g., the arrow pointed left and was presented on the left half of the screen). On incongruent trials, the arrow pointed in the opposite direction as its presentation side (e.g., pointed left and presented on the right side). The incongruent information slows response time, as participants must inhibit the congruent response. The arrow remained on the screen until the participant responded. There was an inter-trial interval of 1 second. There were 32 incongruent trials and 64 congruent trials (96 trials).

Accuracy was high in this task (congruent trials M = 99.0%, incongruent trials M = 93.9%). We calculated an effect of congruency for each participant by fitting a linear mixed effects model and extracting the random slope of trial type (congruent vs. incongruent). First, we excluded incorrect trials, trials where the previous answer was incorrect (to eliminate the effect of post-error slowing), and any trials where the response time was slower than 2000 msec or faster than 200 msec. We then ran a mixed effects model with trial type, prior trial type, z-scored response time on the previous trial, and an interaction between trial type and prior trial type predicting response time[3]. Random effects of trial type and prior trial response time were included by participant. We used each individual's random effect of trial type as our metric of Stroop congruency. This score represents an individual's variability around the fixed effect of congruency, and thus represents how large of a congruency effect an individual showed relative to the group. That is, if a participant has a high, positive random effect, they have a larger effect of congruency on their response time.

Statistical Analyses

Our first analyses addressed the first two research questions. These used separate multiple linear regressions predicting target timing (activation rate), competitor resolution, and peak competitor activation. Our initial models included age as linear and quadratic factors. The

---

[3] These latter three effects were included to account for processes like the congruency sequence effect or the auto-correlation of RTs across trials, that could add noise to the RT on any given trial but were not of interest.

quadratic was added to account for any gains in performance during the developmental window. Age was z-scored prior to the quadratic transformation.

Next, we addressed our third and fourth questions with hierarchical regressions that included the corresponding index from the visual analogue of the VWP (e.g., visual target timing was included for spoken word target timing), each participant's PTA from their better ear, and their random effect of Stroop congruency as first level predictors. Age and age^2 were then added in the second level. In all models, all predictors were z-scored to obtain standardized regression coefficients, which allows for easier comparisons of effect size.

Because the analysis of the rhyme competitors was similar to that of the cohort competitors, it is presented in full in Online Supplement 1. All processed data, analysis scripts, and supplementary analyses are available at https://osf.io/zthbw/.

Results

The time course of fixations to each competitor is presented in Figure 1A averaged across all participants. The typical pattern of competition between lexical candidates is apparent: fixations to the target and cohort competitors begin around 300 msec. This is as early as could be expected given that it takes approximately 200 msec to launch an eye movement, and there was 100 msec of silence before each word. Shortly after that, looks to the target continue to increase, while looks to the cohort and rhyme competitors peak early and are suppressed as more information supporting the target word unfolds.

Figure 2 presents the time course of looks to the target (Figure 2A), cohort (Figure 2B), and rhyme (Figure 2C) split by age group. Note that cohorts and rhymes are plotted after subtracting looks to the unrelated image to isolate competition from general fixations. Younger adults were faster to activate targets (slope of the target fixations) and had higher peak competitor fixations (Figure 2B & 2C). There were also differences at the asymptotes of the function at the end of the time course (higher target asymptote and lower baseline competitor looks), suggesting that younger adults are better able to resolve competition.

The Nature of Age-Related Changes in Word Recognition

To address our first two research questions, we characterized the time course of fixations in terms of three key indices motivated by prior work: *activation rate*, *competition resolution*, and *peak activation* (McMurray et al., 2010, 2022; Rigler et al., 2015). Figure 3 presents these indices as a function of age. These were analyzed in a series of multiple regressions predicting each index from age, as well as the quadratic effect of age. There was a linear and quadratic effect of age on *activation rate* (age: $\beta = -0.44$, $t(104) = -2.72$, $p = .007$;

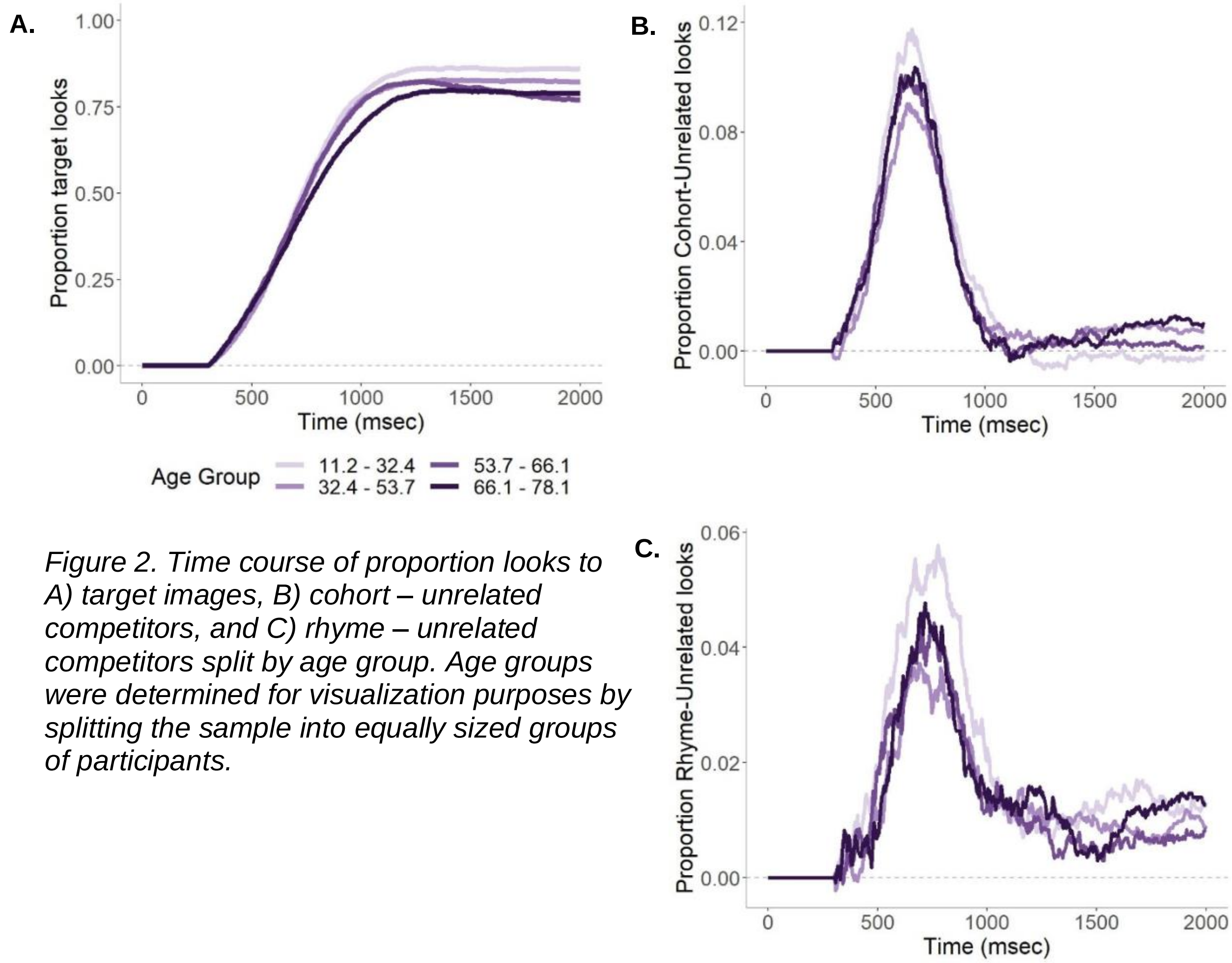

Figure 2. Time course of proportion looks to A) target images, B) cohort – unrelated competitors, and C) rhyme – unrelated competitors split by age group. Age groups were determined for visualization purposes by splitting the sample into equally sized groups of participants.

age^2: β = -0.43, t(104) = -2.64, p = .009), suggesting that lexical activation becomes more efficient (speeds up) from childhood into middle age, when it begins to slow down. For *cohort competitor resolution*, there was a linear effect of age (β = -0.06, t(104) = -2.07, p = .04; Age^2: β = -0.02, t(104) = -0.98, p = .32). As age increases, competitors are not as fully suppressed during word recognition. There was not a significant quadratic effect, suggesting that this index did not exhibit development in the early range. There was no effect of age on *peak cohort activation* (Age: β = -0.004, t(104) = -0.32, p = .75; Age^2: β = 0.02, t(104) = 1.30, p = .19). The analysis of the rhyme competitors largely mirrored that of the cohort competitors and is thus presented in full in an online supplement.

Visual-Cognitive Processing, Hearing, and Inhibitory Control

These initial analyses document robust age effects on real-time spoken word processing. However, these effects could be the result of different underlying changes. For instance, poorer hearing could slow older adults' processing as the incoming signal is more degraded compared to a younger adult. Similarly, domain-general processes like processing speed or inhibitory control are also likely declining with age, and thus could be driving differences in word recognition. To account for this and to address our last two research

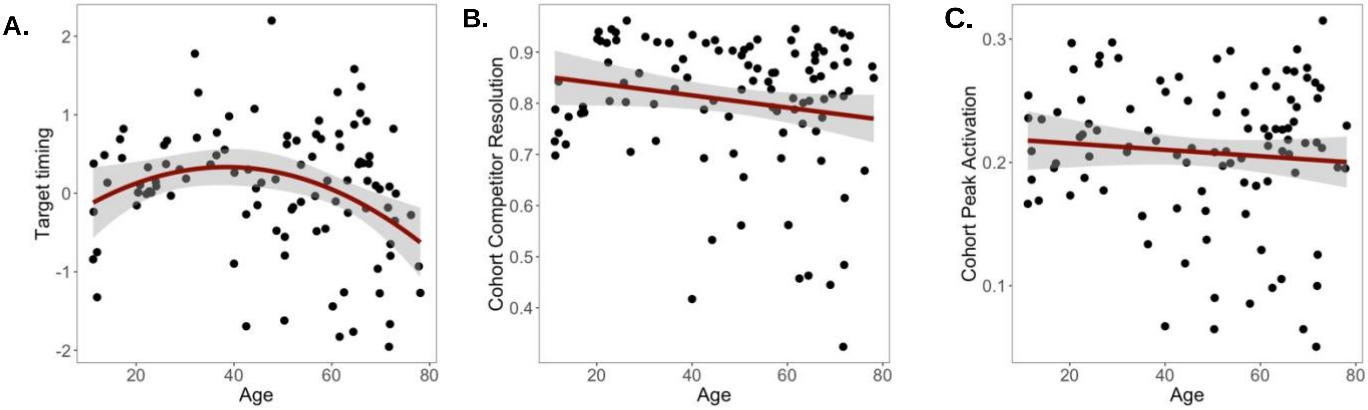

Figure 3. Indices of spoken word recognition presented by age. A) Composite target timing, B) Cohort competitor resolution, and C) Peak cohort activation by age.

questions, we next ran a series of hierarchical regressions that were based on the prior analysis but included factors capturing visual-cognitive processes, including domain-general visual-cognition (processing speed, visual search, etc. indexed by the visual VWP), domain-general inhibitory control (spatial Stroop), and hearing ability (PTA). We included these cognitive and sensory factors in the first level models and then added age and age^2 in the second level models to estimate the unique effect of age after accounting for these other variables. Separate regressions were run for activation rate, resolution, and peak competitor activation.

Table 1 presents the results of a hierarchical regression predicting activation rate. In the first level model, hearing and inhibitory control were not significant. However, the activation rate index from the visual task significantly predicted activation rate in the spoken task ($\beta = 0.44$, $t(83) = 4.24$, $p < .001$): participants who were faster to fixate on targets images in the visual-only VWP were also faster to fixate in the auditory VWP. Nonetheless, in the second-level model, even after accounting for this, we still found a significant quadratic effect of age ($\beta = -0.52$, $t(83) = -2.80$, $p = .006$), and age and age^2 accounted for 6% of the variance over and above the first-level model without these factors ($F(2, 83) = 4.43$, $p = .01$). Figure 4A shows a visualization of this effect, plotting target timing after it has been residualized by the first-level model (i.e., variance in target timing once other factors have been removed). It shows the same (or even more robust) quadratic effect observed in Figure 3A. This supports the unique contribution of age to changes in the efficiency of target word recognition, as the effect of age is not fully explained by additional factors.

We next ran the same analysis with the index of cohort competitor resolution (Table 2). Like the prior analysis, in the first level model, domain-general inhibitory control and hearing were not significant, but we again found a significant effect of resolution in the visual VWP task ($\beta = 0.10$, $t(83) = 3.57$, $p < .001$), suggesting that individuals who can better resolve competition in the visual VWP also do so in the spoken word VWP. Nonetheless, even controlling for this,

Table 1. Summary of hierarchical regressions predicting spoken word activation rate by 1) visual target timing, PTA, and Stroop congruency and 2) those factors along with age.

		Estimate	**Std. Error**	**t(83)**	**p**
Model 1 (Cognitive + sensory factors)	Visual target timing	0.44	0.10	4.24	< 0.001
	Better Ear PTA	-0.14	0.19	-0.76	0.45
	Stroop congruency	-0.0007	0.002	-0.44	0.67
Model 2 (Model 1 + age)	Age	-0.11	0.27	-0.41	0.68
	Age^2	-0.52	0.19	-2.80	0.006
	Better Ear PTA	-0.08	0.25	-0.32	0.75
	Visual target timing	0.81	0.17	4.72	< 0.001
	Stroop congruency	-0.2	0.18	-1.08	0.28
Model Comparison			**ΔR^2**	**F(2,83)**	**p**
	Model 1 ~ Model 2		0.058	4.43	0.015

Table 2. Summary of hierarchical regressions predicting spoken word competitor resolution by 1) visual competitor resolution, PTA, and Stroop congruency and 2) those factors along with age.

		Estimate	**Std. Error**	**t(83)**	**p**
Model 1	Visual competitor resolution	0.10	0.03	3.57	< 0.001
	Better Ear PTA	-0.007	0.03	-0.23	0.82
	Stroop congruency	-0.02	0.03	-0.68	0.49
Model 2	Age	-0.11	0.05	-2.29	0.02
	Age^2	-0.06	0.03	-1.92	0.06
	Better Ear PTA	0.06	0.04	1.44	0.15
	Visual competitor resolution	0.12	0.03	4.04	< 0.001
	Stroop congruency	-0.01	0.03	-0.45	0.65
Model comparison			**ΔR^2**	**F(2,83)**	**p**
	Model 1 ~ Model 2		0.041	3.07	0.051

Table 3. Summary of a linear regression predicting peak spoken word competitor activation by 1) visual competitor peak, PTA, and Stroop congruency and 2) those factors with age.

		Estimate	**Std. Error**	**t(83)**	**p**
Model 1	Peak visual competition	0.04	0.01	2.96	0.004
	Better Ear PTA	-0.0009	0.01	-0.07	0.95
	Stroop congruency	-0.004	0.01	-0.28	0.78
Model 2	Age	-0.02	0.02	-1.01	0.32
	Age^2	-0.003	0.01	-0.21	0.83
	Better Ear PTA	0.01	0.02	0.64	0.52
	Peak visual competition	0.04	0.01	2.69	0.008
	Stroop congruency	0.0003	0.01	0.02	0.98
Model comparison			**ΔR^2**	**F(2,83)**	**p**
	Model 1 ~ Model 2		-0.010	0.54	0.58

the second level model found a small, but significant linear effect of age (linear: $\beta = -0.11$, t(83) = -2.29, p = .02; quadratic: $\beta = -0.14$, t(83) = -1.92, p = .06), uniquely accounting for 5% of the variance (F(2, 83) = 3.06, p = .05). As age increases, individuals do not resolve competition as

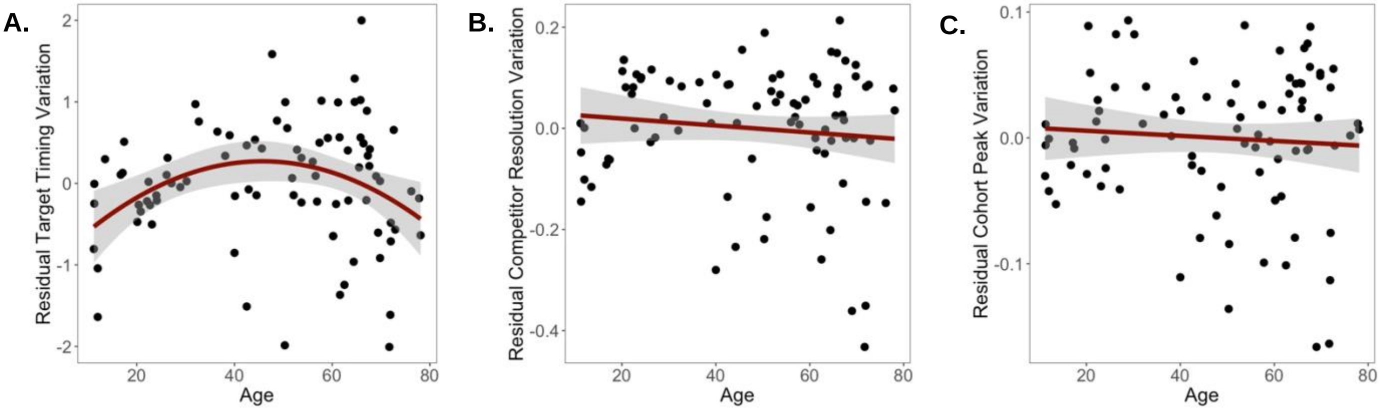

Figure 4. Residual variation from multiple linear regressions with PTA, visual cognition, and inhibitory control as predictors. A) Residual target timing variation, B) Cohort competitor variation, and C) Cohort peak variation by age. Notably, this is residual, unexplained variance from models that did not contain age or age^2 as predictors.

completely. While this is a small effect, there is also a high degree of variability in resolution in middle- and older-aged adults (Figure 3B). This suggests there may be some other factor driving this increased variability in older adulthood.

Finally, we ran the same regression predicting peak cohort activation (Table 3). Peak competitor activation in the visual analogue was the only significant predictor (β = 0.04, t(83) = 2.96, p = .004). In the second level model, age and age^2 did not have an effect on peak activation of lexical competitors.

Discussion

Across the adult lifespan, spoken word recognition increased in efficiency until middle age, when it then began to slow down. This held even when accounting for differences in hearing acuity, visual cognition, processing speed, and inhibitory control. These non-linguistic factors thus do not explain the robust changes in the process of spoken word recognition across the lifespan. However, not all dimensions of lexical competition declined. There was no unique age-related effects on peak competitor activation, suggesting there are specific aspects of the word recognition process that are affected by age over and above what can be explained by changes to hearing, processing speed, and inhibitory control.

While our study focused on aging, it also offers clear evidence that language-specific cognitive processes related to the speed of target activation continue to increase in efficiency (speed up) well into adulthood, from 11 to about 25 years of age. Previous work has found developmental changes to spoken word recognition into adolescence (Rigler et al., 2015); our study extends this developmental window into early adulthood (up through the mid 20's). Consistent with previous work (McMurray et al., 2022; Rigler et al., 2015), we find that developmental changes largely affect activation rate and not competitor resolution.

Conversely, we found surprisingly early age-related declines to both activation rate and competitor resolution. For activation rate, speed of target activation began to slow down in middle age (around the mid 40's). The nature of this change is similar to Van Engen et al. (2020), who found that older adults were slower to fixate on targets compared to a group of younger adults. However, our evidence extends this by providing a clear picture of when word recognition begins to slow down across the lifespan. It is possible that this decline in efficiency reflects not only a slower general processing speed as a result of age, but an increase in the amount of time it takes to eliminate competitors. That is, given that the lexicon continues to grow throughout the lifespan (Kavé & Halamish, 2015), it is feasible to assume that the competition process underlying word recognition will take longer to sift through potential lexical candidates (Ramscar et al., 2013; Wulff et al., 2019).

In addition, starting around middle age, we also observed that increasing age leads to poorer resolution overall (though this was not accompanied by any changes early in life). Previous work has identified poor competitor resolution as a hallmark of Developmental Language Disorder (McMurray et al., 2010) and it has also been recently observed in second language learners (Sarrett et al., 2021), suggesting it is a marker of less stable lexical competition. Here we see it with typical aging, potentially providing a mechanism for prior results suggesting that older adults struggle with lexical competition (Dey et al., 2017; Sommers & Danielson, 1999). Alternatively, poor resolution has also been observed with moderately degraded speech and in good performing cochlear implant users (Farris-Trimble et al., 2014). In these contexts, it has been proposed as a way to maintain flexibility – if a word is initially incorrectly perceived, it may be easier to reactivate words if competition has not fully committed. In aging, this could serve a similar purpose. Older adults are more susceptible to interference from competitors (Dey et al., 2017; Gazzaley et al., 2005), but maintaining activity for competitors could also allow for a greater reliance on contextual information (Goy et al., 2013; Hasher & Campbell, 2020). Future work should ask whether what appears to be poorer resolution is a sign of poorer language skills or is compensatory by relating the degree of competition resolution to real-world outcomes. In addition to differences in the mean competitor resolution, we also observed large increases in the variability of competitor resolution in older adults. This could reflect a mix of causes (e.g., less stable lexical representations *and* more flexibility) across listeners. Future work should investigate whether cognitive outcomes relate to this index, as perhaps these competition dynamics can serve as early indicators of future cognitive health.

We found little evidence that inhibitory control plays a role in spoken word recognition across the lifespan—traditional measures of control (the spatial Stroop) showed no correlation. While the absence of a relationship has been suggested by others (Blomquist & McMurray, 2023; Kapnoula & McMurray, 2021), the present study provides evidence that this is true across the lifespan. That is, across the lifespan, whatever inhibitory processes are used to suppress competitors (if any[4]), they are not the kind of processes tapped by general inhibitory control tasks like Stroop. Rather, they are more likely to be internal to the lexicon (e.g., Dahan et al.,

---

[4] Several classic models of spoken word recognition (Hannagan et al., 2013; Marslen-Wilson, 1987; McClelland & Elman, 1986; Norris, 1994) do not use inhibitory processes to suppress competitors; competitors drop out when they are no longer consistent with the input.

2001; Luce & Pisoni, 1998). More specific to aging, even as older adults begin to show declines to spoken word recognition, there is no evidence that they recruit domain-general resources to compensate. It is possible that older adults only begin to rely on domain-general resources when there is notable challenge in the signal, as there is evidence for a relationship between inhibitory control and word recognition in difficult signal-to-noise ratios (Sommers & Danielson, 1999). We note however, that while this excludes inhibitory control as a factor in word recognition, this does not exclude a role for other aspects of cognitive control such as working memory or planning (c.f. Nitsan et al., 2019).

In contrast, we did observe a relationship between our visual analogue of the VWP. This task captures a large range of lower-level individual differences such as visual search or eye-movement control that are not relevant for language. Our intention was to use it to control for these factors (which may change with age). At the same time, it also taps some higher-level aspects of cognition (e.g., decision making, inhibitory control, speed of processing) that are likely to be highly relevant for word recognition. Thus, our hierarchical approach—in assigning all of this shared variance to the visual VWP—offers an extremely conservative approach to isolating age-related effects on word recognition. This likely explains why the unique variance for age had only a moderate effect size. It could also explain why inhibitory control on its own did not play a role. Listeners who were better able to suppress visual competitors also suppressed lexical competitors more fully. The similarity in tasks may have left little room for Stroop to explain any remaining variance. Nonetheless, the nature of the contribution of any shared components remains an important avenue for future work. Importantly, even after accounting for this variation, we saw clear age-related changes to both the activation rate of targets and competitor resolution. This may reflect the fact that in mechanistic models of word recognition (e.g., TISK [Hannagan et al., 2013]; TRACE [McClelland & Elman, 1986]), domain-general inhibitory control is not the only locus (or even the primary locus) of competitor suppression – competitors are assumed to be suppressed by lateral inhibitory connections internal to the lexicon.

It is possible that poorer auditory encoding is driving age-related declines to spoken word recognition. While we found no influence of hearing thresholds and age-related differences in word recognition were significant even after accounting for differences in PTA, this does not comprehensively rule out an auditory contribution for several reasons. First, and most obviously, at more significant degrees of hearing loss word recognition will clearly be impacted. While real-time lexical processing has been investigated in severe hearing loss (Farris-Trimble et al., 2014; McMurray et al., 2017) and in quiet speech with normal hearing listeners (Hendrickson et al., 2020), real-time lexical processing has not been extensively examined in more typical situations of moderate age-related hearing loss. The interaction of hearing loss and aging for the real-time dynamics of language processing is an important open question.

Second, it is possible that other aspects of auditory processing are linked to the changes we observe. For instance, the encoding fidelity of spectral or temporal cues is known to change with age (Bidelman et al., 2014; Humes et al., 2009), even at audible levels. Older adults who have typical hearing thresholds can struggle with speech-related tasks (Bharadwaj et al., 2015) and there is wide variability in performance on a variety of speech tasks in individuals with normal hearing thresholds (Holmes & Griffiths, 2019). More work is needed to examine the influence of both peripheral auditory function and early auditory cognition (e.g., feature extraction, grouping, and streaming) on spoken word recognition.

Finally, it is important to point out that our sample is not entirely representative of older adults in the United States. Our inclusion criteria deliberately excluded people with mild hearing loss which is highly prevalent in this population. Our cut-off likely excludes individuals past the 75th percentile of hearing ability (Morrell et al., 1996). That is, our study examined the case of people who had largely preserved hearing. As we describe, this may underestimate the contributions of hearing. However, at the same time it leads to a very conservative approach to estimating the effects of aging. Thus, the true effects of age on word recognition are likely to be much bigger in the general population than what we estimate here.

The present work highlights the relatively short developmental window where individuals are at their most efficient when processing spoken words. The implications for this are twofold. First, it calls into question what psycholinguists consider ‘normal’ or ‘ideal’ in terms of language processing (McMurray et al., 2023). As word recognition clearly continues to develop into early adulthood and begins to slow down again in middle age, this leaves perhaps a 15-year window where individuals are ‘peak’ performers. This emphasizes the importance of embracing the diversity of function that exists in other populations that are outside of the oft-studied young, normal hearing participant group. Second, it is likely that word recognition ability impacts a broad variety of skills. Even if the only language deficit is poorer word recognition, this would still slow down sentence recognition and thus cascade to other aspects of language processing. However, given that competition processes are broadly invoked across multiple levels of language (word recognition, sentence processing, and speech production; Dell, 1986; MacDonald et al., 1994), this could be a marker of global competition deficits throughout the language system that affect many levels of language. In turn, difficulty with language negatively impacts social outcomes, with individuals withdrawing from social situations when conversing becomes difficult. That is, even a mild hearing loss, when compounded with less efficient language processing, may make a functional impact that is much larger. Thus, these indices of spoken word recognition have the potential to serve as important indicators of social and cognitive well-being, particularly in older adulthood.

References

Altmann, G. T. M. (1998). Ambiguity in Sentence Processing. *Trends in Cognitive Science*, *2*(4), 146–152.

Apfelbaum, K. S., Goodwin, C., Blomquist, C., & McMurray, B. (2022). The development of lexical competition in written- and spoken-word recognition. *Quarterly Journal of Experimental Psychology*, 174702182210904. https://doi.org/10.1177/17470218221090483

Apfelbaum, K. S., Klein-Packard, J., & McMurray, B. (2021). The pictures who shall not be named: Empirical support for benefits of preview in the Visual World Paradigm. *Journal of Memory and Language*, *121*, 104279. https://doi.org/10.1016/j.jml.2021.104279

Ayasse, N. D., Lash, A., & Wingfield, A. (2017). Effort Not Speed Characterizes Comprehension of Spoken Sentences by Older Adults with Mild Hearing Impairment. *Frontiers in Aging Neuroscience*, *8*(329), 1–12. https://doi.org/10.3389/fnagi.2016.00329

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., & Schneider, B. A. (2011). Effects of Aging and Noise on Real-Time Spoken Word Recognition: Evidence From Eye Movements. *Journal of Speech, Language, and Hearing Research*, *54*(1), 243–262. https://doi.org/10.1044/1092-4388(2010/09-0233)

Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., & Shinn-Cunningham, B. G. (2015). Individual Differences Reveal Correlates of Hidden Hearing Deficits. *Journal of Neuroscience*, *35*(5), 2161–2172. https://doi.org/10.1523/JNEUROSCI.3915-14.2015

Bidelman, G. M., Villafuerte, J. W., Moreno, S., & Alain, C. (2014). Age-related changes in the subcortical–cortical encoding and categorical perception of speech. *Neurobiology of Aging*, *35*(11), 2526–2540. https://doi.org/10.1016/j.neurobiolaging.2014.05.006

Blomquist, C., & McMurray, B. (2023). The Development of Lexical Inhibition in Spoken Word Recognition. *Developmental Psychology*, *59*(1), 186–206.

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

Campbell, K. L., & Hasher, L. (2018). Hyper-binding only apparent under fully implicit test conditions. *Psychology and Aging*, *33*(1), 176–181. https://doi.org/10.1037/pag0000216

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, *108*(3), 804–809. https://doi.org/10.1016/j.cognition.2008.04.004

Colby, S. E., Clayards, M., & Baum, S. R. (2018). The role of lexical status and individual differences for perceptual learning in younger and older adults. *Journal of Speech Language and Hearing Research*, *61*, 1855–1874.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Lang Cogn Process*, *16*(5–6), 507–534. https://doi.org/10.1080/01690960143000074

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*(1–3), 287–314. https://doi.org/10.1016/0010-0277(92)90046-K

Dey, A., & Sommers, M. S. (2015). Age-related differences in inhibitory control predict audiovisual speech perception. *Psychology and Aging*, *30*(3), 634–646. https://doi.org/10.1037/pag0000033

Dey, A., Sommers, M. S., & Hasher, L. (2017). An age-related deficit in resolving interference: Evidence from speech perception. *Psychology and Aging*, *32*(6), 572–587. https://doi.org/10.1037/pag0000189

Dufour, S., & Grainger, J. (2019). Phoneme-Order Encoding During Spoken Word Recognition: A Priming Investigation. *Cognitive Science*, *43*(10). https://doi.org/10.1111/cogs.12785

DuPertuis, L. L., Aldwin, C. M., & Bossé, R. (2001). Does the source of support matter for different health outcomes?: Findings from the Normative Aging Study. *Journal of Aging and Health*, *13*(4), 494–510.

Farris-Trimble, A., & McMurray, B. (2013). Test–Retest Reliability of Eye Tracking in the Visual World Paradigm for the Study of Real-Time Spoken Word Recognition. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1328–1345. https://doi.org/10.1044/1092-4388(2012/12-0145)

Farris-Trimble, A., McMurray, B., Cigrand, N., & Tomblin, J. B. (2014). The process of spoken word recognition in the face of signal degradation. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 308–327. https://doi.org/10.1021/nn300902w.Release

Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, *8*(10), 1298–1300. https://doi.org/10.1038/nn1543

Goy, H., Pelletier, M., Coletta, M., & Pichora-Fuller, M. K. (2013). The Effects of Semantic Context and the Type and Amount of Acoustic Distortion on Lexical Decision by Younger and Older Adults. *Journal of Speech, Language, and Hearing Research*, *56*(6), 1715–1732. https://doi.org/10.1044/1092-4388(2013/12-0053)

Hannagan, T., Magnuson, J. S., & Grainger, J. (2013). Spoken word recognition without a TRACE. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00563

Hasher, L., & Campbell, K. L. (2020). Inhibitory Theory: Assumptions, Findings, and Relevance to Interventions. In A. K. Thomas & A. Gutchess (Eds.), *The Cambridge Handbook of Cognitive Aging* (1st ed., pp. 147–160). Cambridge University Press. https://doi.org/10.1017/9781108552684.010

Hendrickson, K., Spinelli, J., & Walker, E. (2020). Cognitive processes underlying spoken word recognition during soft speech. *Cognition*, *198*, 104196. https://doi.org/10.1016/j.cognition.2020.104196

Holmes, E., & Griffiths, T. D. (2019). ‘Normal’ hearing thresholds and fundamental auditory grouping processes predict difficulties with speech-in-noise perception. *Scientific Reports*, *9*(1), 16771. https://doi.org/10.1038/s41598-019-53353-5

Humes, L. E., Busey, T. A., Craig, J. C., & Kewley-Port, D. (2009). The effects of age on sensory thresholds and temporal gap detection in hearing, vision, and touch. *Attention, Perception, & Psychophysics*, *71*(4), 860–871. https://doi.org/10.3758/APP.71.4.860

Humes, L. E., Pichora-Fuller, M. K., & Hickson, L. (2020). Functional Consequences of Impaired Hearing in Older Adults and Implications for Intervention. In K. S. Helfer, E. L. Bartlett, A. N. Popper, & R. R. Fay (Eds.), *Aging and Hearing* (Vol. 72, pp. 257–291). Springer International Publishing. https://doi.org/10.1007/978-3-030-49367-7_11

International Organization for Standardization. (2017). *Acoustics—Statistical distribution of hearing thresholds related to age and gender (Standard 7029)*. https://www.iso.org/standard/42916.html

Kapnoula, E. C., & McMurray, B. (2016). Training alters the resolution of lexical interference: Evidence for plasticity of competition and inhibition. *Journal of Experimental Psychology: General*, *145*(1), 8–30. https://doi.org/10.1037/xge0000123

Kapnoula, E. C., & McMurray, B. (2021). Idiosyncratic use of bottom-up and top-down information leads to differences in speech perception flexibility: Converging evidence from ERPs and eye-tracking. *Brain and Language*, *223*, 105031. https://doi.org/10.1016/j.bandl.2021.105031

Kavé, G., & Halamish, V. (2015). Doubly blessed: Older adults know more vocabulary and know better what they know. *Psychology and Aging*, *30*(1), 68–73. https://doi.org/10.1037/a0038669

Kuchinsky, S. E., Ahlstrom, J. B., Vaden, K. I., Cute, S. L., Humes, L. E., Dubno, J. R., & Eckert, M. A. (2013). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss: Pupil size in older adults. *Psychophysiology*, *50*(1), 23–34. https://doi.org/10.1111/j.1469-8986.2012.01477.x

Lash, A., Rogers, C. S., Zoller, A., & Wingfield, A. (2013). Expectation and entropy in spoken word recognition: Effects of age and hearing acuity. *Exp Aging Res*, *39*(3), 235–253. https://doi.org/10.1080/0361073X.2013.779175

Lin, F. R., Metter, E. J., O’Brien, R. J., Resnick, S. M., Zonderman, A. B., & Ferrucci, L. (2011). Hearing loss and incident dementia. *Archives of Neurology*, *68*(2), 214–220. https://doi.org/10.1001/archneurol.2010.362

Lotto, A., & Holt, L. (2011). Psychology of auditory perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 479–489. https://doi.org/10.1002/wcs.123

Luce, P. A., & Cluff, M. S. (1998). Delayed commitment in spoken word recognition: Evidence from cross-modal priming. *Perception and Psychophysics*, *60*(3), 484–490. https://doi.org/10.3758/BF03206868

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

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical Nature of Syntactic Ambiguity Resolution. *Psychological Review*, *101*(4), 676–703.

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

Mattys, S. L., & Scharenborg, O. (2014). Phoneme categorization and discrimination in younger and older adults: A comparative analysis of perceptual, lexical, and attentional factors. *Psychology and Aging*, *29*(1), 150–162. https://doi.org/10.1037/a0035387

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

McMurray, B. (2019). *Eyelink Analysis* (4.12) [Computer software]. https://osf.io/c35tg/

McMurray, B. (2020). *Nonlinear Curvefitting for Psycholinguistic (and other) Data* (Version 30) [Computer software]. https://osf.io/4atgv/

McMurray, B., Apfelbaum, K. S., & Tomblin, J. B. (2022). *The slow development of real-time processing: Spoken word recognition as a crucible for new thinking about language acquisition and disorders*. https://psyarxiv.com/uebfc/

McMurray, B., Baxelbaum, K. S., Colby, S., & Bruce Tomblin, J. (2023). Understanding language processing in variable populations on their own terms: Towards a functionalist psycholinguistics of individual differences, development, and disorders. *Applied Psycholinguistics*, 1–28. https://doi.org/10.1017/S0142716423000255

McMurray, B., Farris-Trimble, A., & Rigler, H. (2017). Waiting for lexical access: Cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition*, *169*, 147–164. https://doi.org/10.1016/j.cognition.2017.08.013

McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology*, *60*(1), 1–39. https://doi.org/10.1016/j.cogpsych.2009.06.003

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*(4), 475–494. https://doi.org/10.1016/j.jml.2007.11.006

Morrell, C. H., Gordon-Salant, S., Pearson, J. D., Brant, L. J., & Fozard, J. L. (1996). Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. *The Journal of the Acoustical Society of America*, *100*(4), 1949–1967. https://doi.org/10.1121/1.417906

Nitsan, G., Wingfield, A., Lavie, L., & Ben-David, B. M. (2019). Differences in Working Memory Capacity Affect Online Spoken Word Recognition: Evidence From Eye Movements. *Trends in Hearing*, *23*, 1–12. https://doi.org/10.1177/2331216519839624

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234. https://doi.org/10.1016/0010-0277(94)90043-4

Payne, B. R., Grison, S., Gao, X., Christianson, K., Morrow, D. G., & Stine-Morrow, E. A. L. (2014). Aging and individual differences in binding during sentence understanding: Evidence from temporary and global syntactic attachment ambiguities. *Cognition*, *130*(2), 157–173. https://doi.org/10.1016/j.cognition.2013.10.005

Phillips, N. A. (2016). The Implications of Cognitive Aging for Listening and the Framework for Understanding Effortful Listening (FUEL). *Ear & Hearing*, *37*(1), 44S-51S. https://doi.org/10.1097/AUD.0000000000000309

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

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology*, *47 Suppl 2*, S72-82. https://doi.org/10.1080/14992020802307404

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 & Hearing*, *37*(1), 5S-27S. https://doi.org/10.1097/AUD.0000000000000312

Ramscar, M., Hendrix, P., Love, B., & Baayen, R. H. (2013). Learning is not decline: The mental lexicon as a window into cognition across the lifespan. *The Mental Lexicon*, *8*(3), 450–481. https://doi.org/10.1075/ml.8.3.08ram

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.The

Rigler, H., Farris-Trimble, A., Greiner, L., Walker, J., Tomblin, J. B., & McMurray, B. (2015). The slow developmental time course of real-time spoken word recognition. *Developmental Psychology*, *51*(12), 1690–1703. https://doi.org/10.1037/dev0000044

Sarrett, M. E., Shea, C., & McMurray, B. (2021). *Within- and between-language competition in adult second language learners: Implications for language proficiency*. https://psyarxiv.com/bwv7c/

Schneider, B. A., Daneman, M., & Murphy, D. R. (2005). Speech Comprehension Difficulties in Older Adults: Cognitive Slowing or Age-Related Changes in Hearing? *Psychology and Aging*, *20*(2), 261–271. https://doi.org/10.1037/0882-7974.20.2.261

Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K. (2002). Listening in aging adults: From discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *56*(3), 139–152. https://doi.org/10.1037/h0087392

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.

SR Research. (n.d.). *Eyelink 1000* [Computer software]. sr-research.com

Toscano, J. C., Anderson, N. D., & McMurray, B. (2013). Reconsidering the role of temporal order in spoken word recognition. *Psychonomic Bulletin and Review*, *20*(5), 981–987. https://doi.org/10.3758/s13423-013-0417-0

Tucker-Drob, E. M., de la Fuente, J., Köhncke, Y., Brandmaier, A. M., Nyberg, L., & Lindenberger, U. (2022). A strong dependency between changes in fluid and crystallized abilities in human cognitive aging. *Science Advances*, *8*(5), eabj2422. https://doi.org/10.1126/sciadv.abj2422

Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, *24*(3), 761–766. https://doi.org/10.1037/a0014802

Van Engen, K. J., Dey, A., Runge, N., Spehar, B., Sommers, M. S., & Peelle, J. E. (2020). Effects of age, word frequency, and noise on the time course of spoken word recognition. *Collabra: Psychology*, *6*(1), 1–10. https://doi.org/10.1525/collabra.17247

Waters, G. S., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging*, *16*(1), 128–144. https://doi.org/10.1037/0882-7974.16.1.128

Weber, A., & Scharenborg, O. (2012). Models of spoken-word recognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 387–401. https://doi.org/10.1002/wcs.1178

Wingfield, A., Aberdeen, J. S., & Stine, E. A. (1991). Word onset gating and linguistic context in spoken word recognition by young and elderly adults. *Journal of Gerontology*, *46*(3), P127-129.

World Health Organization (Ed.). (2021). *World Report on Hearing*. World Health Organization.

Wulff, D. U., De Deyne, S., Jones, M. N., & Mata, R. (2019). New Perspectives on the Aging Lexicon. *Trends in Cognitive Sciences*, *23*(8), 686–698. https://doi.org/10.1016/j.tics.2019.05.003

Zhao, L., Yuan, S., Guo, Y., Wang, S., Chen, C., & Zhang, S. (2022). Inhibitory control is associated with the activation of output-driven competitors in a spoken word recognition task. *The Journal of General Psychology*, *149*(1), 1–28. https://doi.org/10.1080/00221309.2020.1771675

Efficiency of Spoken Word Recognition Slows Across the Adult Lifespan

ONLINE SUPPLEMENT

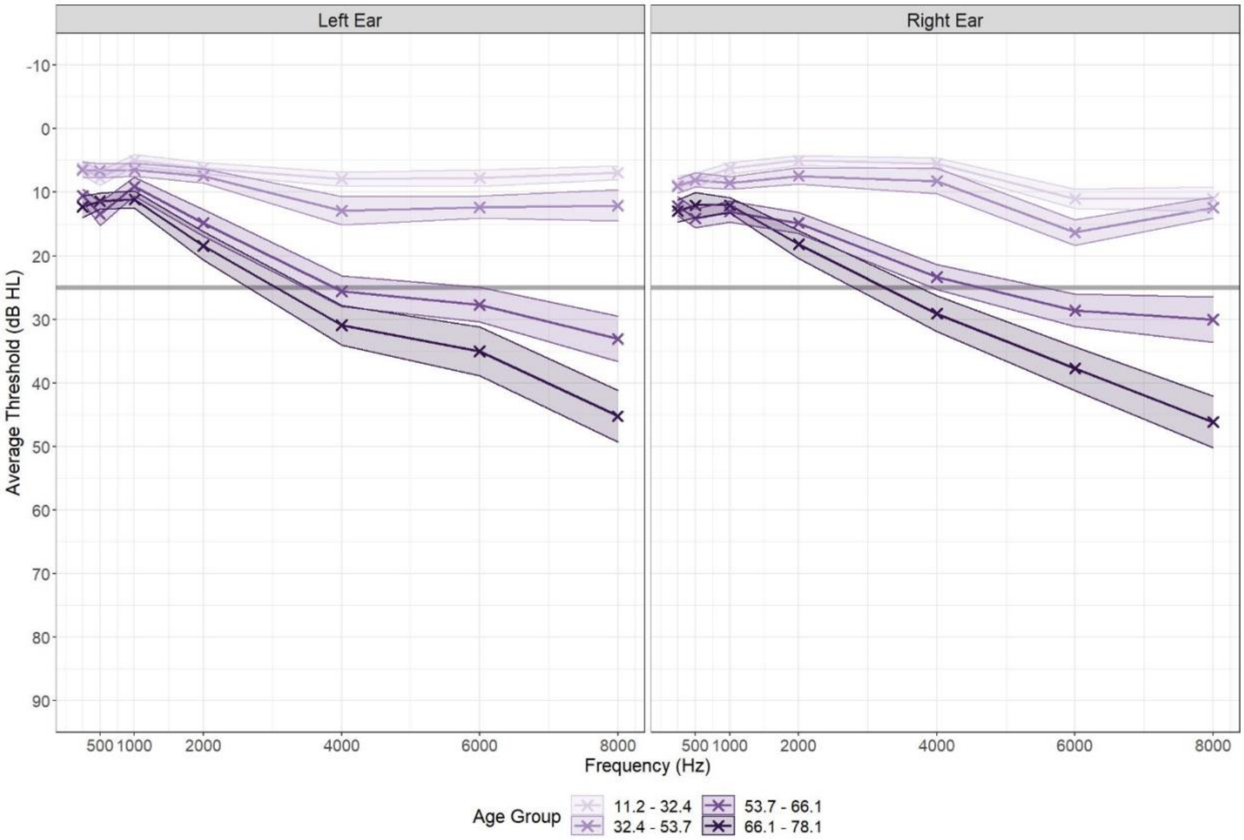

Figure S1. Average pure-tone thresholds (in dB HL) from each ear for each age group. Ribbon width represents standard error. Age group was determined for visualization by splitting the sample into groups with an equal number of participants. Horizontal line at 25 dB HL indicates the standard cut-off for 'normal' hearing.

Supplemental Analysis of Rhyme Competitors.

The analysis of looks to rhyme competitors strongly parallels that of cohort competitors and is presented for completeness here. Indices were calculated as described in the Data Processing section in the main text.

Figure S2 presents rhyme competitor resolution and peak activation as a function of age. These were analyzed in two multiple linear regressions predicting each index from age and the quadratic effect of age. There was a linear effect of age on ***rhyme competitor resolution*** (Age: β = -0.05, t(104) = -1.92, p = .05; Age^2: β = -0.03, t(104) = -1.07, p = 0.28), suggesting that as age increases, resolution becomes poorer. This is consistent with the analysis of the cohort competitors. For ***peak rhyme activation***, there was a significant quadratic effect of age (Age: β = -0.0008, t(104) = -0.10, p = .91; Age^2: β = 0.03, t(104) = 3.97, p < .001). This is in contrast to our analysis of cohort competitors, where there was no significant effect of age. However, we must account for changes to domain-general cognition and hearing to better explain these effects.

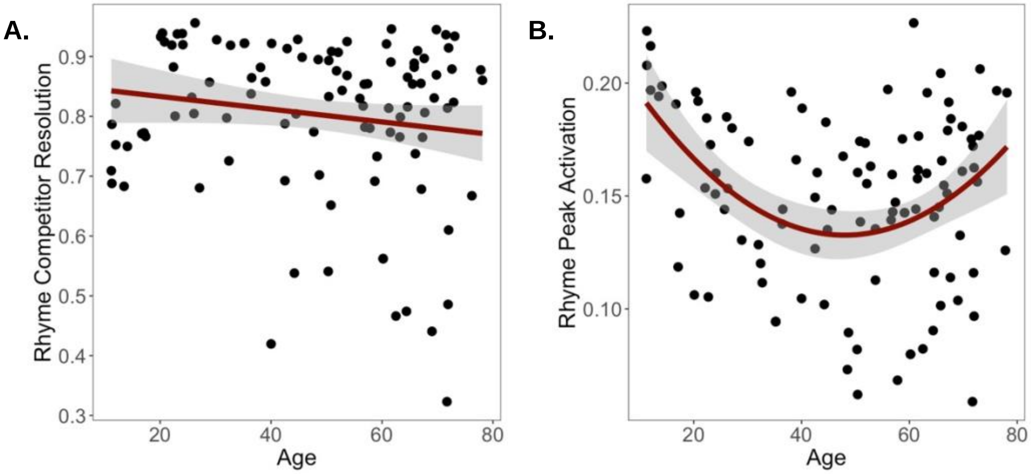

Figure S2. A) Rhyme competitor resolution and B) Peak rhyme activation by age.

Thus, we next conducted a hierarchical regression that predicted ***rhyme resolution*** from PTA, visual cognition, and inhibitory control at the first level, and added age at the second level (Table S1). Hearing and inhibitory control were not significant. Competitor resolution from the visual-only analogue of the VWP significantly predicted rhyme resolution (β = 0.10, t(83) = 3.58, p < .001), suggesting that individuals who are better able to resolve competition do so in both modalities. When controlling for this, we still found a significant effect of age on rhyme resolution (Age: β = -0.10, t(83) = -2.19, p = .03; Age^2: β = -0.06, t(83) = -1.96, p = .05). The addition of Age and Age^2 to a model containing the other factors was significant (F(2) = 2.94, p = .05). This is consistent with our analysis of cohort competitors, suggesting that as age increases, individuals become worse at resolving competition, regardless of competitor type.

Table S1. Summary of hierarchical regressions predicting rhyme competitor resolution by 1) visual target timing, PTA, and Stroop congruency and 2) those factors along with age.

		Estimate	**Std. Error**	**t(83)**	**p**
Model 1 (Cognitive + sensory factors)	Visual competitor resolution	0.10	0.03	3.58	< 0.001
	Better Ear PTA	-0.003	0.03	-0.10	0.92
	Stroop congruency	-0.02	0.03	-0.66	0.51
Model 2 (Model 1 + age)	Age	-0.10	0.05	-2.19	0.03
	Age^2	-0.07	0.03	-1.96	0.05
	Better Ear PTA	0.06	0.04	1.46	0.15
	Visual competitor resolution	0.12	0.03	4.06	< 0.001
	Stroop congruency	-0.02	0.03	-0.48	0.64
Model Comparison			**ΔR^2**	**F(2,83)**	**p**
	Model 1 ~ Model 2		0.04	2.94	0.05

Table S2 presents the results of a hierarchical regression predicting ***peak rhyme activation*** by age, PTA, visual cognition, and inhibitory control. Only peak competitor activation from the visual VWP task significantly predicted peak rhyme activation ($\beta = 0.04$, $t(83) = 5.16$, $p < .001$). Despite an earlier indication of an effect of age, this is no longer significant once accounting for other factors. This is also consistent with our analysis of cohort competitors.

Table S2. Summary of a linear regression predicting peak rhyme competitor activation by 1) visual competitor peak, PTA, and Stroop congruency and 2) those factors with age.

		Estimate	**Std. Error**	**t(83)**	**p**
Model 1	Peak visual competition	0.04	0.007	5.16	< 0.001
	Better Ear PTA	-0.005	0.008	-0.62	0.53
	Stroop congruency	-0.003	0.008	-0.38	0.70
Model 2	Age	-0.004	0.01	-0.32	0.75
	Age^2	0.02	0.008	1.72	0.08
	Better Ear PTA	-0.003	0.01	-0.26	0.79
	Peak visual competition	0.03	0.008	4.33	< 0.001
	Stroop congruency	0.002	0.008	0.24	0.81
Model comparison			**ΔR^2**	**F(2,83)**	**p**
	Model 1 ~ Model 2		0.02	2.19	0.12