



Published in final edited form as:

Neuroimage. 2021 January 15; 225: 117511. doi:10.1016/j.neuroimage.2020.117511.

Neural sensitivity to phonological characteristics is stable across the lifespan

Michele T. Diaz^{a,b,*}, Hossein Karimi^c, Sara B.W. Troutman^a, Victoria H. Gertel^a, Abigail L. Cosgrove^a, Haoyun Zhang^b

^aDepartment of Psychology, The Pennsylvania State University, USA

^bSocial, Life, and Engineering Sciences Imaging Center, The Pennsylvania State University, USA

^cDepartment of Psychology, Mississippi State University, USA

Abstract

Aging is often associated with declines in language production. For example, compared to younger adults, older adults experience more tip-of-the-tongue (TOT) states, show decreased speed and accuracy in naming objects, and have more pauses and fillers in speech, all of which indicate age-related increases in retrieval difficulty. While prior work has suggested that retrieval difficulty may be phonologically based, it is unclear whether there are age-related differences in the organization of phonological information per se or whether age-related difficulties may arise from accessing that information. Here we used fMRI to investigate the neural and behavioral basis of phonological neighborhood density (PND) effects on picture naming across the lifespan (N=91, ages 20-75). Consistent with prior work, behavioral results revealed that higher PND led to faster picture naming times and higher accuracies overall, and that older adults were less accurate in their responses. Consistent with the behavioral analyses, fMRI analyses showed that increasing PND was associated with decreased activation in auditory and motor language regions, including bilateral superior temporal gyri and bilateral precentral gyri. Interestingly, although there were age-related increases in functional activation to picture naming, there were no age-related modulations of neural sensitivity to PND. Overall, these results suggest that having a large cohort of phonological neighbors facilitates language production, and although aging is associated with increases in language production difficulty, sensitivity to phonological features during language production is stable across the lifespan.

1. Introduction

Although most language abilities are remarkably stable over the lifespan, older adults often have increased difficulty with language retrieval and production. Indeed, older adults often

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

*Corresponding author at: Department of Psychology, The Pennsylvania State University, 356 Moore Building, University Park, PA 16801, USA.. mtd143@psu.edu (M.T. Diaz).

Author credit statement

MTD was involved in the Conceptualization, Funding acquisition, Project administration, Data analysis, and Writing - original draft; HK, SBWT, VHG, ALC, and HZ were involved in Data collection, Data analysis, and Writing - review & editing.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2020.117511.

cite word retrieval as a frequent and frustrating challenge (Ossher et al., 2013; Sadat et al., 2014, Vitevitch and Stamer, 2006). In laboratory studies, older adults are often slower to name pictures (Mortensen et al., 2006), and have more omissions in speech. However, while these issues have been widely investigated behaviorally, less is known about the neural bases of these effects. Moreover, many prior studies of language production and aging have examined only younger and older adults, limiting our understanding of how language production differs across the lifespan. To examine these issues, individuals aged 20 – 75 years old completed an fMRI picture naming study to investigate neural and behavioral sensitivity to phonological aspects of production by manipulating a well-established phonological variable — phonological neighborhood density (PND).

PND is the number of words that differ from a target word by the addition, subtraction, or substitution of a single sound (Landauer and Streeter, 1973). For example, the word “phone” has many neighbors including “cone”, “tone”, “loan”, “own”, “foam”, etc., whereas “month” has none. Phonological neighborhood density has been shown to influence both perception and production, where larger neighborhoods often produce interference during perception and facilitation during production (for a review see, Baus et al., 2008; Goldrick et al., 2010; Harley and Brown, 1998; Perez, 2007; Stemberger, 2004; Tsai, 2018; Vitevitch, 2002; Vitevitch and Luce, 2016) but see also (Newman and German, 2005; Sadat et al., 2014; Tabak et al., 2010; Vitevitch and Stamer, 2006). Focusing on language production, items with larger phonological neighborhoods are generally named faster (Vitevitch, 2002) and more accurately (Stemberger, 2004), and others have shown that words with dense phonological neighborhoods are named more accurately in tongue twisters (Vitevitch, 2002), and elicit fewer malapropisms (e.g., mitten for muffin, Vitevitch, 1997).¹ These facilitatory effects are thought to arise from form-related activation of phonologically-similar neighbors which then feeds back to the intended target. Phonological neighborhoods have also been found to influence articulatory aspects of production, in which words with dense phonological neighborhoods have shorter vowel durations, more expanded vowel space (Gahl et al., 2012; Munson and Solomon, 2004; Wright, 2004), longer voice onset times (Fox et al., 2015; Fricke et al., 2016; Zhang et al., 2019), and more pronounced coarticulatory effects (Scarborough, 2013; Scarborough and Zellou, 2013). Collectively, these findings suggest that phonological neighbors influence the speed, accuracy, and form of language production.

However, many of these studies have been conducted in younger adults. Although less is known about middle-aged adults, the few studies that have examined language production suggest that middle-aged individuals have the highest picture naming accuracy (Kavé et al., 2010; Newman and German, 2005). In contrast, older adults often have increased difficulty with language production. Several empirical studies have suggested that at least some part of these age-related increases in retrieval difficulty may be related to phonological aspects of the stimuli. For example, older adults have more tip-of-the-tongue (TOT) experiences (Burke et al., 2004; Burke et al., 1991; Maylor, 1990; Rastle and Burke, 1996) which reflect

¹However, not all studies of language production find facilitatory effects of PND, e.g., (Newman and German, 2005; Sadat et al., 2014). Moreover, effects of PND may interact with other factors, such as age of acquisition (Karimi & Diaz, In Press). However, in the present study we focus on basic neural effects of PND, since these have not been previously examined.

phonological, but not semantic retrieval failure. These age-related phonological deficits have been taken as evidence that phonological connections may be particularly vulnerable to decline (Burke et al., 1991).

In contrast to a phonological deficit account, others have suggested that cognitive decline among older adults may be influenced by inhibitory deficits (Hasher et al., 1991; Hasher and Zacks, 1988). If older adults inhibit less efficiently, then this could lead to activation of the intended target as well as other less relevant candidates, producing increased selection demands. Although phonological and inhibitory deficits are not mutually exclusive, these theoretical perspectives predict different effects of phonological neighborhoods in older adults. If phonological links weaken across the lifespan, the facilitory effect of phonological neighbors should also weaken. This may leave words with smaller phonological neighborhoods most vulnerable to age-related declines because such words have fewer neighbors from which to benefit. Alternatively, if inhibitory deficits increase with age, then we may expect a more pronounced influence of phonological neighbors on language production because if more phonological neighbors are active this might boost the activation of the intended target.

We know of only one study that has examined age-related differences in the effects of phonological neighborhoods (Vitevitch and Sommers, 2003). Vitevitch and Sommers found that older adults experienced more TOTs compared with younger adults, but only for items that had both low phonological neighborhood density and low neighborhood frequency, suggesting that older adults had increased retrieval difficulty when there were particularly weak neighborhood effects. In a separate picture naming experiment, they found that high neighborhood frequency led to faster naming speed and higher accuracy rates for older adults. These findings support the idea that older adults may be disproportionately impaired when processing items with relatively weak or few phonological neighbors, and that strong phonological neighborhoods facilitate retrieval.² These results also suggest that both the size and the strength of the phonological neighborhood may influence retrieval.

All of the studies discussed thus far have focused on behavior; however, less is known about the neural bases of phonological neighborhood effects on language production. Peramunage and colleagues examined how younger adults processed minimal pairs (words varying by only a single phonemic element such as voicing, e.g., cape - gape, Peramunage et al., 2011). They found that words with close phonological neighbors elicited less activation than words without close phonological neighbors in left hemisphere regions that have been implicated in phonological processing (left superior temporal gyrus and left supramarginal gyrus) as well as language selection (left inferior frontal gyrus) and motor planning (left precentral gyrus). Although they did not look at phonological neighborhood density effects per se, these findings suggest that greater phonological overlap elicited less activation in younger adults. Others have used the picture-word interference design to examine how explicitly presenting a phonologically-related word affects retrieval. Consistent with what Peramunage and colleagues found, these picture-word interference studies find that written and auditorily

²It may also be the case that the influence of phonological neighbors varies as a function of that item's relative activation in relation to its neighborhood (e.g., as reflected in frequency in Vitevitch and Sommers, 2003).

presented phonologically-related words also engage bilateral superior temporal gyri, left supramarginal gyrus, and left inferior frontal gyrus (Abel et al., 2009; de Zubicaray and McMahon, 2009; de Zubicaray et al., 2002; Diaz et al., 2014). Collectively, these findings highlight the importance of bilateral superior temporal gyri, and left hemisphere language regions (precentral, supramarginal, and inferior frontal gyri) in phonological processing.

Neuroimaging studies that have examined phonological processing in older adults suggest that they may be impaired in phonological aspects of production. For example, Shafto and colleagues showed that older adults had less brain activation in the left insula compared to younger adults during TOT states, and linked these lower levels of activation to poorer naming performance (Shafto et al., 2010). Interestingly, they found no age differences during successful naming, suggesting that age differences in the neural bases of language production may be strongest specifically during phonological retrieval failures, as TOTs do not involve failed semantic access. Consistent with these results, research from our lab has demonstrated that older adults are less accurate and less efficient when making phonological decisions and have weaker brain-behavior relations (Diaz et al., 2014; Diaz et al., 2018). Moreover, older adults may engage key language regions such as the insula less, when retrieval demands are high, as when naming low frequency items (Gertel et al., 2020).

Although age-related increases in language production impairments have been documented, older adults show stability in other aspects of language, such as similar patterns of semantic priming (Burke et al., 1987; Madden et al., 1993), and similar sensitivity to frequency manipulations (Gertel et al., 2020; Gollan et al., 2008; LaGrone and Spieler, 2006; Newman and German, 2005). While none of these previous investigations have specifically examined phonological processes, these findings raise the possibility that the underlying organization of language may be largely intact. Indeed several studies within the domain of language suggest that age-related increases in functional activation may be due to external task demands. For example, in language production tasks, older adults have more difficulty with increasing task demands compared to younger adults (Zhang et al., 2019), and age-related increases in fMRI activation during language comprehension have been attributed to the addition of an explicit task (Davis et al., 2014). Thus, an open question is whether age-related deficits in language production reflect differences in the organization of basic linguistic abilities or differences in executive aspects such as selection or adjusting to task demands.

In the present study we investigated language production using an fMRI picture naming paradigm. Specifically, we were interested in assessing neural and behavioral sensitivity to PND across the lifespan using a minimally demanding task. If language production deficits are related to a purely phonological deficit, then we would expect older adults to show less behavioral and neural sensitivity to PND, compared to younger adults. Alternatively, if language production deficits are related to decreased inhibition we might expect age-related increases in sensitivity to PND. However, if the underlying organization of the phonological system is maintained across the lifespan, we would expect similar sensitivities to PND across ages. We adopted a simple picture naming paradigm, to limit the potential influence of task demands, and focused our question of interest on an implicit manipulation of the underlying PNDs of the items. Moreover, we expand upon many of the previous studies of

aging and language by examining age as a continuous variable to assess age-related differences across the lifespan.

2. Methods

2.1. Participants

Ninety-three adults participated in the experiment, one was removed due to poor performance and a second participant was removed due to the possibility of depression, leaving 91 complete data sets. Participants were aged 20-75 (mean age = 47.40 years, $sd = 17.45$ years, 54 female). Participants' ages were distributed across the lifespan in order to assess age-related differences in language production, see Fig. 1 for a histogram of the age distribution. All participants were community-dwelling, right-handed, native American-English speakers who had minimal exposure to other languages. All participants had normal or corrected-to-normal vision, and reported no history of neurological, psychological, or major medical conditions (Christensen et al., 1992). Prior to the MRI session, each participant completed a battery of psychometric and neuropsychological tests to assess basic cognitive functions such as speed, working memory, executive function, and language. These tasks included the Mini-Mental State Exam to screen for mild cognitive impairment or dementia (MMSE, Folstein et al., 1975); the geriatric depression scale (GDS) to screen for depression (Yesavage et al., 1982); WAIS-III vocabulary, digit-symbol, and digit span subtests (Wechsler et al., 1997); phonemic (F, A, S) and categorical (animals) verbal fluency (Patterson, 2011); the author recognition test to assess reading habits (Acheson et al., 2008); the California Verbal Learning Test to assess immediate and delayed memory (Woods et al., 2006); simple and choice reaction time tests to assess speed³; a reading span task (Conway et al., 2005; Loboda, 2012), a computerized version of the Stroop task (Stroop, 1935), and a story elicitation task as a more naturalistic measure of language production. Age was significantly positively correlated with RT measures of speed (simple speed, choice speed, digit symbol) and inhibition (Stroop) and negatively correlated with measures of working memory and recall (backward digit span, verbal working memory, immediate and delayed recall). Age was also positively correlated with reading habits, with older adults reporting greater familiarity with authors, which has been interpreted as increased reading experience (Acheson et al., 2008). Demographic characteristics and assessment scores are reported in Table 1. All participants provided written, informed consent, and all procedures were approved by the Institutional Review Board at The Pennsylvania State University.

2.2. Stimuli and procedure

Participants performed a picture naming task in the MRI scanner, see Fig. 2. On each trial, one of 191 color photographs ($396 \text{ pixels} \times 396 \text{ pixels}$, duration = 1500 ms) was presented on a white background and participants were instructed to overtly name the photograph as quickly as possible while still responding accurately. Participants were also asked to be as specific as possible in their answers (e.g., robin instead of bird). To control for brain activation to basic visual and motoric features, we included a control condition that

³Speed tests consisted of computerized, lab-developed, reaction time tasks involving detection of a square (simple) and pressing a left or right button in response to arrows pointing left or right (choice). The story elicitation task involved telling the story of the picture book "Frog, where are you?" by Mercer Mayer.

consisted of 50 photographs that had been diffeomorphically transformed to yield unrecognizable objects that maintained the basic visual properties of untransformed photographs (See Fig. 2 and supplemental materials for examples of transformed images, Stojanoski and Cusack, 2014). Participants were instructed to respond to these abstract items with the word 'picture'. Items were presented in a randomized order with a variable inter-stimulus interval (range = 1.5 – 15 s, mean = 3.40 s) that was determined using the optseq2 program, as jittered presentations have been shown to optimize the hemodynamic response (Dale, 1999) and prevent participants from anticipating the onset of events. Participants completed 4, 6-minute runs. During the task, overt verbal responses were recorded and filtered using an MR-compatible, dual-channel, fiber-optic microphone system (Optoacoustics Ltd., Or-Yehuda, Israel). Prior to scanning, participants practiced overt picture naming while minimizing head movement in a mock scanner.

Photographs were taken from two normed databases (Brodeur et al., 2014; Moreno-Martínez and Montoro, 2012), as well as publicly available images from the internet. Images depicted common concrete objects from a variety of categories such as animals, clothing, food, and household items. In developing the final experimental stimuli, all images were normed by a separate group of native English-speaking individuals (N = 28) to ensure that they could be identified accurately and consistently. Linguistic characteristics of the final stimuli were obtained from the English Lexicon Project (ELP, Balota et al., 2007, see Supplemental Materials for detailed word characteristics). The names of photographs had an average word length = 5.88 (SD = 2.01, range 3 - 11), log Hyperspace Analog to Language (HAL) word frequency = 8.00 (SD = 1.54, range = 4.48 - 12.16), number of phonemes = 4.81 (SD = 1.78, range = 2 - 9), number of syllables = 1.77 (SD = 0.83, range = 1 - 4), visual complexity = 4.00 (SD = 1.68, range = 0.43 - 7.5), and H-index = 0.25 (SD = 0.40, range = 0 - 1.55). H-index is a measure of name agreement that accounts for both the proportion and variability in acceptable names given by participants for a particular image (Snodgrass and Vanderwart, 1980). Higher name agreement corresponds with lower H-index values. Phonological Neighborhood Densities (PND) were varied across the items to allow us to investigate the role of these factors on picture naming. PND was restricted to the number of phonological neighbors that overlapped with the first syllable, as Fricke, Baese-Berk, & Goldrick have demonstrated that phonological overlap in the first syllable is most influential on processing (mean PND = 7.72, SD = 10.00, range = 0 – 37, Fricke et al., 2016).

2.3. Acquisition of MRI data

MRI scanning was completed on a 3T Siemens Prisma Fit MRI scanner with a 64-channel head coil. Sagittal T1 weighted localizer images were collected and used to define a volume for data collection, higher-order shimming, and alignment to the anterior and posterior commissures (AC-PC). T1 weighted anatomical images were collected using a magnetization-prepared rapid acquisition gradient echo (MP-RAGE) sequence (repetition time [TR] = 2300 ms; echo time [TE] = 2.28 ms; Inversion Time [TI] = 900 ms; flip angle = 8°; echo spacing = 7 ms; acceleration factor = 2; field of view [FOV] = 256 mm²; voxel size = 1 × 1 × 1 mm; 160 contiguous slices; duration ~ 5 min).

Functional images were collected using an echo-planar imaging (EPI) sequence (TR = 2000 ms; TE = 25 ms; flip angle = 90°; echo spacing = 0.49 ms; FOV = 240 mm²; voxel size = 3 × 3 × 4 mm; 33 contiguous axial slices, parallel to the AC–PC; sequential-descending acquisition; phase encoding = anterior to posterior; fat saturation = on; run duration ~ 6 minutes; 4 functional runs). Two additional volumes were acquired and deleted at the beginning of each functional run to reach steady state equilibrium. A field map sequence was also collected using a double-echo spoiled gradient echo sequence (TR = 446 ms; TE = 4.92 ms; flip angle = 60°; FOV = 240 mm²; voxel size = 3 × 3 × 4 mm; 33 contiguous slices; phase encoding = anterior to posterior, fat saturation = off; duration = 1:12 minutes) that generated 2 magnitude images and 1 phase image. Resting state and diffusion tensor images were also collected; however, those data were analyzed separately.

2.4. Behavioral data analyses

Responses were coded for accuracy and naming time based on the recordings from the MRI session. Responses were marked as correct if the participant provided the exact target name (e.g., *chicken* for *chicken*) or the plural form of the target name (e.g., *chickens* for *chicken*). Because our analyses focused on phonological neighborhood characteristics, exact overlap in onsets was required for the manipulation to be valid, and all other responses were classified as errors. Errors were further subdivided into acceptable alternatives (e.g., *jet* for *plane*), dysfluencies (e.g., *helicopter*, *I mean plane* or *uh, ... plane*), incorrect responses (e.g., *helicopter* for *plane*), omissions, and technical errors⁴. Incorrect responses accounted for 21.20% of the data. This high percentage is largely a result of our strict accuracy rubric, as acceptable alternatives accounted for 79.9% of errors. The distribution of errors was analyzed with a chi-square goodness of fit test to assess distributional differences. In this and other by-group follow up analyses, we divided the participants into 3 roughly equal groups based on their distance from the mean age (younger age range = 20-40, mean age = 27.42, N = 33; middle-aged age range = 41-57, mean age = 49.38, N = 26; Older age range = 58-75, mean age = 66.38, N = 32).

Accuracy was analyzed using generalized logistic mixed-effect modeling, employing the *glmer* function in the *lme4* package (Bates et al., 2014) in R (R Core Team, 2017). This approach has the advantage of accounting for individual data points, allowing intercepts and slopes to be random across participants and items. As recommended by Barr, Levy, Scheepers, and Tily (Barr et al., 2013), for all regression models we always included the maximal random effects structure. By-participant random slopes for frequency⁵ and PND, random intercepts of participant and item, and by-item random slopes for the effect of Age were included in the final accuracy and naming time models.⁶

Naming times were calculated using customized PRAAT scripts (Boersma and Weenink, 2019). The PRAAT scripts identified response onsets by searching the recordings for pitch

⁴There were two types of technical errors—ones in which the responses were recorded but the naming times were missing, and ones in which both the recorded response and naming times are missing.

⁵We included effects of frequency as a covariate to better isolate effects due to PND (frequency – PND correlation in our data set: $r^2 = 0.42$, $p < .001$).

⁶Note that each participant has only one value for age, and each item has only one value for frequency and PND. Thus, by-participant random slopes for age and by-item random slopes for frequency and PND cannot be calculated and were not included.

deviations within the filtered auditory signal. These onsets were then manually verified by using both the audio and visual speech stream. The naming times were calculated as the difference between the photograph onsets generated from the E-Prime output and the response onsets. Only trials with correct responses, responses longer than 200 ms, and responses with naming times within 3 SDs were included in further analyses⁷. Excluded outliers accounted for 3.13% of total data points. Naming times were subjected to a square root transformation to better approximate a normal distribution and analyzed using mixed-effect regression modeling, employing the *lmer* function in the *lme4* package (Bates et al., 2014) in R (R Core Team, 2017). PND, Age, and the interaction between PND and age were included as independent variables and naming time was included as the continuous dependent variable. As with the accuracy model, frequency was included as a covariate in the naming time model. The R script and the full behavioral data are available via the Open Science Framework: <https://osf.io/7bc8w>

2.5. fMRI data analyses

The fBIRN QA tool was used to assess data quality (Glover et al., 2012, https://www.nitrc.org/projects/bxh_xcede_tools/), measuring the number of potentially clipped voxels, mean signal fluctuation to noise ratio (SFNR), and per-slice variation. Additionally, the anatomical and functional images were visually inspected for artifacts and signal drop-out. Non-brain tissue of the anatomical images was removed using Optimized Brain Extraction for Pathological Brains (optiBET: Lutkenhoff et al., 2014). We used FSL (version 5.0.9), with FEAT (fMRI expert analysis tool) version 6.0 (Smith et al., 2004; Woolrich et al., 2004), to carry out preprocessing and statistical analyses. Preprocessing steps included motion correction (FSL MCFLIRT), B0 unwarping, slice timing correction, spatial smoothing (FWHM = 5 mm), high-pass filtering (50 s), coregistration, and normalization. During normalization participants functional data was first aligned to his or her own anatomical data using boundary-based registration (Greve and Fischl, 2009) and then to the MNI 152 standard image (12 degrees of freedom). We conducted first level analyses on correct trials within each participant's individual runs, convolving each trial onset with a double-gamma hemodynamic response function⁸, and including the standard motion parameters as nuisance covariates. Participants moved less than ½ a voxel, as recommended for inclusion in data analyses (average motion = .26 mm, SD = .12 mm, range = .06 - .68 mm, Huettel et al., 2014). However, age and motion were significantly positively correlated ($r = .27, p < .01$). For the contrasts of interest (described below), only trials with correct responses were included. Errors were always included in the fMRI models, but as a separate regressor. Because there were few errors, we did not conduct further analyses on these trials⁹. Run-level analyses were combined across runs within participants and then across participants in group-level analyses using FMRIB's local analysis of mixed effects (FLAME 1+2, Beckmann et al., 2003; Woolrich et al., 2004). All significant activations were

⁷Naming times within 3 SDs should capture > 99% of a normal distribution. Naming times < 200ms are typically due to microphone problems or inadvertent noises.

⁸We used FSL's double gamma HRF which is a combination of two gamma functions, a standard positive function at a normal lag and a small delayed inverted gamma function that models a late undershoot in the BOLD response.

⁹We ran a second fMRI model that included effects of naming time and frequency, and the PND results were consistent across both models.

determined using a two-step process in which Z (Gaussianised T/F) statistical images were initially thresholded at the voxel level ($p < .001$). Clusters of identified voxels were then corrected for multiple comparisons ($p < .05$, corrected) based on Gaussian random field theory (Worsley, 2001) in which each cluster's estimated significance level was compared with the cluster probability threshold, and then only clusters whose estimated significance exceeded the threshold were included in the results (Hayasaka and Nichols, 2003). Additionally, results from comparisons between conditions were masked to ensure that only differences based on significant positive hemodynamic responses were included in the analyses. All reported brain regions were identified using the Harvard-Oxford Structural Atlas (Desikan et al., 2006). Coordinates are reported in MNI space, and results are overlaid on a representative brain in MNI space.

For our functional MRI analyses, we first examined the main effect of picture naming, looking at regions in which naming pictures elicited greater activation than the control condition (i.e., viewing scrambled images and saying "picture"). We then conducted a parametric analysis in which we looked for regions in which variability in PND elicited increases or decreases in functional activation. Then, we examined both analyses to determine if there were any influences of Age on the functional activation patterns to naming pictures, or sensitivity to PND. Lastly, we were interested in brain-behavior relations specifically whether PND-related and age-related differences in fMRI activation related to task performance or individual differences. To assess this, we correlated behavioral performance (i.e., RT) as well as language-related neurocognitive measures (verbal fluency, verbal working memory, vocabulary, and reading habits) with brain activation. This approach allowed us to examine the relationship between overall behavioral performance and functional activation across individuals.

3. Results

3.1. Behavioral results

We conducted a logistic mixed effects regression analysis on Accuracy, examining the effect of PND, Age, and the interaction between PND and Age, while covarying effects of frequency. This revealed a significant main effect of PND on accuracy in which increasing PND was associated with higher accuracies, $Z = 4.51$, $p < .001$, eta squared = .22. There was also a significant main effect of Age, in which increasing Age was associated with lower accuracies, $t = -2.92$, $p < .005$, eta squared = -.08. Moreover, there was a marginally significant quadratic effect of age on accuracy ($p = .08$), with the strongest effect of age on accuracy among the oldest adults. See Fig. 3. There was no significant interaction between Age and PND, suggesting that the effect of PND on accuracy is stable across the lifespan. We also examined the response data for any effect of error type, considering 4 types of errors: acceptable alternatives (e.g., baby deer for fawn), dysfluencies, incorrect responses (e.g., book for menu), and omissions. A chi-square analysis showed that the three groups differed from one another with middle-aged adults making the fewest errors both overall and within each category ($\chi^2 = 51.669$, $df = 6$, $p < .001$). Younger adults made significantly fewer errors than older adults ($\chi^2 = 16.008$, $df = 3$, $p < .01$), both overall and within each error category, except for dysfluencies.

We conducted a similar mixed effects regression analysis on naming time. Again, there was a significant main effect of PND in which larger PND values were associated with faster naming times, $t = -3.22$, $p < .005$, eta squared = $-.11$, Fig. 3. Although there was no significant main effect of Age on naming time, there was a marginally significant interaction between Age and PND, $t = -1.88$, $p = 0.06$. To better understand this interaction, we looked at the effect of PND across 3 age groups: younger adults, middle aged adults, and older adults. This analysis showed that there was a facilitatory effect of PND in all three groups, and that this effect was strongest among older adults. Follow-up analyses showed that this effect was driven by older adults responding similarly to younger and middle-aged adults on items with higher PND, but more slowly than the other two groups on items with lower PND. Consistent with our accuracy results, the naming time analyses suggest that the effect of PND on naming time is relatively stable across the lifespan but that older adults may be more slowed by low PND items compared to younger and middle-aged adults.

4. fMRI results

First, we looked at the main effect of picture naming (pictures > scrambled objects) collapsed across participants. Picture naming elicited significant patterns of activation in ventral visual, temporal, and frontal regions (see Fig. 4a, Table 2). Notably, these activations included left superior frontal gyrus, left medial prefrontal cortex, left inferior frontal gyrus, bilateral insula, bilateral orbital frontal cortex, anterior cingulate, bilateral precentral gyri, bilateral hippocampi, bilateral parahippocampal gyri, bilateral lateral occipital cortex, which extended into bilateral inferior temporal gyri, and right cerebellum.

Significant effects of Age on picture naming (Fig. 4b, Table 2) were found in several regions across the brain in which increases in age were associated with increases in functional activation. These regions included bilateral frontal pole, which extended into middle and superior frontal gyri, bilateral precentral gyri, bilateral superior and middle temporal gyri, bilateral thalamus, bilateral hippocampus, and bilateral lateral occipital cortices. In short, consistent with prior research, we observed age-related increases in functional activation. There were no regions in which age-related decreases in activation were observed.

Our main interests were in the effects of PND on brain activation and whether these effects differed as a function of age. Analysis of functional activation indicated that PND was negatively related to functional activation. That is, as PND increased, this was associated with decreases in fMRI activation, consistent with a facilitatory effect of larger phonological neighborhoods. Reductions in activation are shown in Fig. 5 and Table 3, and were found in left frontal pole, right orbital frontal gyrus, bilateral superior temporal gyri, which extended into Heschl's gyrus, bilateral precentral gyri, and bilateral lateral occipital cortex, which extended into bilateral fusiform and lingual gyri. There were no regions in which increases in PND were related to increases in fMRI activation. To examine the influence of age on PND, we looked for an interaction between PND and Age; however, there were no significant regions in which sensitivity to PND differed across ages.¹⁰ Consistent with our

¹⁰To further investigate the possibility of age-related differences, we categorized age across 3 groups: younger adults, middle-aged adults, and older adults, and conducted an ANOVA to look for an interaction of Age Group by PND. Consistent with our parametric analysis, the overall interaction was not significant.

behavioral results, this suggests a relative stability in the neural effect of PND across the lifespan.

Given the significant age-related increases in functional activation to picture naming and the significant decreases in functional activation to PND, we were interested in whether these patterns related to behavior. There was no relationship between these effects and picture naming RT, nor several language-related behavioral assessments (e.g., reading habits, verbal fluency, verbal working memory, or vocabulary). Although null results should be interpreted with caution, these suggest that sensitivity to phonological neighborhoods is relatively robust and less sensitive to individual differences in these behavioral measures.

5. Discussion

In an across-the-lifespan study we used a picture naming paradigm with objects whose names varied in Phonological Neighborhood Density (PND). Consistent with the previous literature, objects whose names had larger phonological neighborhoods were named faster and more accurately. Moreover, fMRI results showed that objects with larger PNDs elicited less activation, which is often interpreted as facilitation, primarily in bilateral superior temporal gyri, and bilateral lateral occipital cortex, regions that support phonological (Heim et al., 2003; Vaden et al., 2010) and visual object processing (Gerlach et al., 2002; Koutstaal et al., 2001) respectively. Additionally, decreases in activation were observed in bilateral precentral gyri, a region involved in motor control (Sakai et al., 2000). Consistent with the extant literature, behavioral age-related increases in difficulty were observed for naming: age was negatively correlated with accuracy. Moreover, age was also positively correlated with fMRI activation. This increased need for neural resources is often interpreted as reflecting increased difficulty (Reuter-Lorenz and Cappell, 2008). These age-related increases in activation were found in both language-relevant regions such as bilateral superior and middle temporal gyri and bilateral precentral gyri, as well as in regions associated with domain general executive function (bilateral middle and superior frontal gyri), memory (bilateral hippocampus), and vision (bilateral lateral occipital cortices). Critically, these age-related increases in errors and functional activation were related only to picture naming in general, as there was no significant Age x PND interaction in the accuracy or fMRI results. Overall, these results suggest that neural sensitivity to phonological neighborhood structure is intact across the lifespan, although retrieval difficulty, as measured through increased errors and functional activation, increases with age.

While no neuroimaging studies to date have explicitly examined the neural bases of PND on language production, several studies in younger adults have examined how the brain supports phonological aspects of language production (Abel et al., 2009; de Zubicaray and McMahan, 2009; de Zubicaray et al., 2002; Diaz et al., 2014; Peramunage et al., 2011). The most commonly incorporated task is the picture-word interference paradigm, in which a phonologically-related word is shown during a picture naming task. Closer to the present study's design, Peramunage and colleagues (2011) examined the influence of implicit phonological competitors (minimal pair words) on language production in younger adults. Whether the phonologically-related word was explicitly or implicitly present, these studies showed that phonological distractors engage bilateral superior temporal gyri, left supra

marginal gyrus, and left inferior frontal gyrus (Abel et al., 2009; de Zubicaray and McMahon, 2009; de Zubicaray et al., 2002; Diaz et al., 2014, Peramunage et al., 2011). Consistent with these results, we found less fMRI activation for words with larger PND values in regions that have been shown to be sensitive to phonology, such as bilateral superior temporal gyri, and in regions that have a role in articulatory and motor function, such as bilateral precentral gyri. Of note, the neural sensitivity to PND did not vary across the lifespan, suggesting that neural sensitivity to basic phonological features remains intact throughout the lifespan. We also examined the relationship between PND and other behavioral and cognitive factors, such as RT, verbal working memory, vocabulary, and reading habits. However, there were no significant relationships between the neural measures of PND sensitivity and behavioral measures. Although this suggests that the basic phonological organization is robust to other cognitive factors, this interpretation should be taken with caution as these are null results.

One caveat to the stability of our fMRI and accuracy results across the lifespan is that we did observe a marginally significant interaction between Age and PND in naming time. Although older adults were not slower to respond than younger adults to high PND items, they were slower to respond to low PND items. This suggests that although the overall phonological organization remains intact across the lifespan, older adults may have some vulnerability in processing items with weaker representations. This is consistent with what others have observed for items with low frequency and items with both low PND and low neighborhood frequency values (Gertel et al., 2020; Vitevitch and Sommers, 2003). However, this Age x PND interaction did not translate to age-related differences in neural patterns of activation, suggesting that the naming time differences were not sufficient to incur additional fMRI processing demands. It may be the case that neural age-related differences did not emerge because we focused on successful naming. That is, only pictures that were precisely named with our intended label were included in our fMRI analyses. Other work that has found age-related neural differences has focused on phonological retrieval failures (i.e., TOT incidents, Shafto et al., 2007; Shafto et al., 2010). Indeed, Shafto and colleagues found few age-related differences when they examined successful naming. Consistent with this idea, older adults named words with large PNDs similarly to younger adults, suggesting that age-related naming impairments may be selective to words with less robust representations (i.e., low frequencies, small PNDs).

While we did not observe strong age-related differences in sensitivity to phonology, there were age-related differences in picture naming, generally. First, although participants performed well in general, increasing age was associated with less accurate picture naming. Although it was not the primary goal of the study, picture naming performance was correlated with several cognitive measures suggesting that individuals who were better able to name pictures had higher cognitive abilities.¹¹ Picture naming was also associated with age-related increases in functional activation in both language regions, such as bilateral

¹¹Naming times were negatively correlated with immediate ($r = -.22, t = -2.43, p = .04$) and delayed recall ($r = -.20, t = -2.29, p = .06$). Picture naming accuracy was negatively correlated with digit symbol RT ($r = -.33, t = -2.83, p < .01$) and positively correlated with MMSE scores ($r = .37, t = 4.09, p < .001$), immediate ($r = .38, t = 4.10, p < .001$) and delayed recall ($r = .46, t = 5.49, p < .001$), and vocabulary ($r = .40, t = 4.37, p < .001$).

precentral gyri and bilateral superior and middle temporal gyri, as well as in dorsal superior frontal regions that have been shown to be sensitive to task difficulty (e.g., Badre and D'Esposito, 2007), as well as in regions sensitive to memory demands (i.e., bilateral hippocampus, for a review see Rajah et al., 2015). These findings suggest that age-related increases in word retrieval activation involve both increased reliance on language specific processes, as well as domain-general executive function and memory resources.

An increased reliance on domain-general neural resources is consistent with age-related increased difficulty in adjusting to task demands. Although we adopted a relatively straightforward task (picture naming) with an implicit manipulation (varying PND of the items), behavioral and neural data suggest that tasks become increasingly difficult with age. Consistent with this, several studies have demonstrated that older adults are more sensitive to task demands during language comprehension (Davis et al., 2014; Kennedy et al., 2015; Zhang et al., 2019) and during non-language tasks (e.g., Braver and Barch, 2002; Cappell et al., 2010; Reuter-Lorenz and Cappell, 2008; Rieck et al., 2017). Our work has observed such task difficulty effects specifically during picture naming (Zhang et al., 2019). We have also found that older adults may be slower to adjust to task demands, even when practice trials are provided. For example, in a blocked picture naming paradigm, older adults were less accurate compared to younger adults, but only during the first task run (Gertel et al., 2020).

One other novel aspect of our approach was the inclusion of middle-aged adults, as very few language studies have examined this age range. Behaviorally, we found that middle-aged adults (roughly 40 – 60 years old) performed on par with younger adults, suggesting that although aging is associated with declines in language production, these declines may not emerge until later in life. Of note, our participants represent a relatively well-educated sample (average years of education = 16.2), and education was significantly positively correlated with age. Additional studies are needed to establish the role of education in this context. Individual differences, such as these, may have also been reflected in the variability in responses, as acceptable alternatives accounted for ~16% of responses. It is possible that there were differences in how individuals approached the naming task and how they selected their responses. These kinds of cohort effects may have influenced our results and because we incorporated a cross-sectional design, we were unable to assess how our effects might have changed within individuals across time. As a future direction we would like to investigate how neural sensitivity to PND across the lifespan may interact with other factors such as frequency, neighborhood frequency, and name agreement, as previous investigations have shown that these factors can interact in picture naming (Karimi and Diaz, 2020; Vitevitch and Sommers, 2003)

Despite these limitations, we were able to measure behavioral and neural sensitivity to phonological characteristics in a large lifespan sample of adults. Our results suggest that, overall, sensitivity to phonological features is stable across the lifespan. However, with increasing age, difficulty in overt production increases, and older adults rely on both increased activation of language-relevant regions as well as domain-general executive and memory resources. These results demonstrate consistency in the effect of PND on language production across the lifespan.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This publication was supported by NIH R01 AG034138 to Michele T. Diaz. SBWT was also supported by a T32 fellowship (NIH T32 AG049676 to David Almeida). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. We thank Anna Eppes for help creating the stimuli and programming the experiment, and the staff and scientists at the Social, Life, & Engineering Sciences Imaging Center and the Center for Language Science, where the experiment was conducted. Stimuli, analysis scripts, behavioral and fMRI data are available via the Open Science Framework at: <https://osf.io/7bc8w>

References

- Abel S, Dressel K, Bitzer R, Kummerer D, Mader I, Weiller C, Huber W, 2009 The separation of processing stages in a lexical interference fMRI-paradigm. *Neuroimage* 44 (3), 1113–1124. [PubMed: 19015036]
- Acheson DJ, Wells JB, MacDonald MC, 2008 New and updated tests of print exposure and reading abilities in college students. *Behav. Res. Methods* 40 (1), 278–289 [PubMed: 18411551]
- Badre D, D'Esposito M, 2007 Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *J. Cogn. Neurosci* 19 (12), 2082–2099. doi:10.1162/jocn.2007.19.12.2082. [PubMed: 17892391]
- Balota DA, Yap MJ, Hutchison KA, Cortese MJ, Kessler B, Loftis B, ..., Treiman R, 2007 The english lexicon project. *Behav. Res. Methods* 39 (3), 445–459 [PubMed: 17958156]
- Barr DJ, Levy R, Scheepers C, Tily HJ, 2013 Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Memory Language* 68 (3), 255–278
- Bates D, Mächler M, Bolker B, & Walker S (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823
- Baus C, Costa A, Carreiras M, 2008 Neighbourhood density and frequency effects in speech production: A case for interactivity. *Language Cognitive Processes* 23 (6), 866–888. doi:10.1080/01690960801962372.
- Beckmann CF, Jenkinson M, Smith SM, 2003 General multilevel linear modeling for group analysis in fMRI. *Neuroimage* 20 (2), 1052–1063 [PubMed: 14568475]
- Boersma P, & Weenink D (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.08. Retrieved from <http://www.praat.org/>
- Braver TS, Barch DM, 2002 A theory of cognitive control, aging cognition, and neuromodulation. *Neurosci. Biobehav. Rev* 26 (7), 809–817. doi:10.1016/S0149-7634(02)00067-2. [PubMed: 12470692]
- Brodeur MB, Guérard K, Bouras M, 2014 Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLoS One* 9 (9), e106953, [PubMed: 25211489]
- Burke DM, Locantore JK, Austin AA, Chae B, 2004 Cherry pit primes Brad Pitt. *Psychol. Sci* 15 (3), 164–170. doi:10.1111/j.0956-7976.2004.01503004.x. [PubMed: 15016287]
- Burke DM, Mackay DG, Worthley JS, Wade E, 1991 On the tip of the tongue: What causes word finding failures in young and older adults? *J. Memory Language* 30, 542–579.
- Burke DM, White H, Diaz DL, 1987 Semantic priming in young and older adults: Evidence for age constancy in automatic and attentional processes. *J. Exp. Psychol. Hum. Percept. Perform* 13, 79–88 [PubMed: 2951489]
- Cappell KA, Gmeindl L, Reuter-Lorenz PA, 2010 Age differences in prefrontal recruitment during verbal working memory maintenance depend on memory load. *Cortex* 46, 462–473. doi:10.1016/j.cortex.2009.11.009. [PubMed: 20097332]
- Christensen KJ, Moye J, Armson RR, Kern TM, 1992 Health screening and random recruitment for cognitive aging research. *Psychol. Aging* 7 (2), 204 [PubMed: 1610509]

- Conway AR, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW, 2005 Working memory span tasks: a methodological review and user's guide. *Psychon. Bull. Rev* 12 (5), 769–786 [PubMed: 16523997]
- Dale AM, 1999 Optimal experimental design for event-related fMRI. *Hum. Brain Mapp* 8 (2-3), 109–114 [PubMed: 10524601]
- Davis SW, Zhuang J, Wright P, Tyler LK, 2014 Age-related sensitivity to task-related modulation of language-processing networks. *Neuropsychologia* 63, 107–115. doi:10.1016/j.neuropsychologia.2014.08.017 [PubMed: 25172389]
- de Zubicaray GI, McMahon KL, 2009 Auditory context effects in picture naming investigated with event related fMRI. *Cognitive, Affect. Behav. Neurosci* 9, 260–269. doi:10.3758/CABN.9.3.260.
- de Zubicaray GI, McMahon KL, Eastburn MM, Wilson SJ, 2002 Ortho-graphic/phonological facilitation of naming responses in the picture–word task: An event-related fMRI study using overt vocal responding. *Neuroimage* 16, 1084–1093. doi:10.1006/nimg.2002.1135. [PubMed: 12202095]
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, ..., Hyman BT, 2006 An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31 (3), 968–980 [PubMed: 16530430]
- Diaz MT, Hogstrom LJ, Zhuang J, Voyvodic JT, Johnson MA, Camblin CC, 2014 Written distractor words influence brain activity during overt picture naming. *Front. Human Neurosci* 8, 167. doi:10.3389/fnhum.2014.00167.
- Diaz MT, Johnson MA, Burke DM, Madden DJ, 2014 Age-related differences in the neural bases of phonological and semantic processes. *J. Cogn. Neurosci* 26 (12), 2798–2811. doi:10.1162/jocn_a_00665. [PubMed: 24893737]
- Diaz MT, Johnson MA, Burke DM, Truong T, Madden DJ, 2018 Age-related differences in the influence of task-irrelevant information on the neural bases of phonological and semantic processes. *Cognitive, Affective, Behav. Neurosci* 19 (4), 829–844. 10.3758/s13415-018-00671-2
- Folstein MF, Folstein SE, McHugh PR, 1975 “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res* 12 (3), 189–198. [PubMed: 1202204]
- Fox NP, Reilly M, Blumstein SE, 2015 Phonological neighborhood competition affects spoken word production irrespective of sentential context. *J. Memory Language* 83, 97–117. doi:10.1016/j.jml.2015.04.002.
- Fricke M, Baese-Berk MM, Goldrick M, 2016 Dimensions of similarity in the mental lexicon. *Lang., Cognit., Neurosci.* 31 (5), 639–645. doi:10.1080/23273798.2015.1130234.
- Gahl S, Yao Y, Johnson K, 2012 Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *J. Memory Lang* 66 (4), 789–806. doi:10.1016/j.jml.2011.11.006.
- Gerlach C, Aaside CT, Humphreys GW, Gade A, Paulson OB, Law I, 2002 Brain activity related to integrative processes in visual object recognition: bottom-up integration and the modulatory influence of stored knowledge. *Neuropsychologia* 40 (9), 1254–1267. doi:10.1016/S0028-3932(01)00222-6, doi:<https://doi.org/>. [PubMed: 11931928]
- Gertel VH, Karimi H, Dennis NA, Neely KA, & Diaz MT (2020). Lexical frequency affects functional activation and accuracy in picture naming among older and younger adults *Psychology & Aging*, Advance online publication, doi:10.1037/pag0000454
- Glover GH, Mueller BA, Turner JA, Van Erp TG, Liu TT, Greve DN, ..., Keator DB, 2012 Function biomedical informatics research network recommendations for prospective multicenter functional MRI studies. *J. Magn. Reson. Imaging* 36 (1), 39–54. [PubMed: 22314879]
- Goldrick M, Folk JR, Rapp B, 2010 Mrs. malaprop's neighborhood: using word errors to reveal neighborhood structure. *J. Memory Lang* 62 (2), 113–134. doi:10.1016/j.jml.2009.11.008.
- Gollan TH, Montoya RI, Cera C, Sandoval TC, 2008 More use almost always a means a smaller frequency effect: aging, bilingualism, and the weaker links hypothesis. *J. Memory Language* 58 (3), 787–814. doi:10.1016/j.jml.2007.07.001.
- Greve DN, Fischl B, 2009 Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* 48 (1), 63–72. 10.1016/j.neuroimage.2009.06.060. [PubMed: 19573611]

- Harley TA, Brown HE, 1998 What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *Br. J. Psychol* 89 (1), 151–174. doi:10.1111/j.2044-8295.1998.tb02677.x.
- Hasher L, Stoltzfus ER, Zacks RT, Rypma B, 1991 Age and inhibition. *J. Exper. Psychol* 17 (1), 163–169.
- Hasher L, Zacks RT, 1988 Working memory, comprehension, and aging: a review and a new view In: Bower GH (Ed.). In: *The Psychology of Learning and Motivation*, 22 Academic Press, San Diego, CA, pp. 193–225.
- Hayasaka S, Nichols TE, 2003 Validating cluster size inference: random field and permutation methods. *Neuroimage* 20 (4), 2343–2356. [PubMed: 14683734]
- Heim S, Opitz B, Müller K, Friederici AD, 2003 Phonological processing during language production: fMRI evidence for a shared production-comprehension network. *Cognitive Brain Res.* 16 (2), 285–296. doi:10.1016/s0926-6410(02)00284-7, doi:.
- Huettel SA , Song AW, McCarthy G, 2014 *Functional Magnetic Resonance Imaging*, 3rd ed. Sinauer Associates.
- Karimi H, Diaz MT, 2020 When phonological neighborhood density both facilitates and impedes: age of acquisition and name agreement interact with phonological neighborhood during word production. *Mem. Cogn* doi:10.3758/s13421-020-01042-4.
- Kavé G, Knafo A, Gilboa A, 2010 The rise and fall of word retrieval across the lifespan. *Psychol. Aging* 25 (3), 719–724. 10.1037/a0018927. [PubMed: 20853975]
- Kennedy KM, Rodrigue KM, Bischof GN, Hebrank AC, Reuter-Lorenz PA, Park DC, 2015 Age trajectories of functional activation under conditions of low and high processing demands: an adult lifespan fMRI study of the aging brain. *Neuroimage* 1 (104), 21–34. doi:10.1016/j.neuroimage.2014.09.056.
- Koutstaal W, Wagner AD, Rotte M, Maril A, Buckner RL, Schacter DL, 2001 Perceptual specificity in visual object priming: functional magnetic resonance imaging evidence for a laterality difference in fusiform cortex. *Neuropsychologia* 39 (2), 184–199. 10.1016/s0028-3932(00)00087-7. [PubMed: 11163375]
- LaGrone S, Spieler DH, 2006 Lexical competition and phonological encoding in young and older speakers. *Psychol. Aging* 21 (4), 804–809. doi:10.1037/0882-7974.21.4.804. [PubMed: 17201499]
- Landauer TK, Streeter LA, 1973 Structural differences between common and rare words: failure of equivalence assumptions for theories of word recognition. *J. Verbal Learn. Verbal Behav* 12 (119-131).
- Loboda TD (2012). Reading Span (RSPAN) Task [Web application]. Retrieved from <https://ubiq-x.github.io/rspan>
- Lutkenhoff ES, Rosenberg M, Chiang J, Zhang K, Pickard JD, Owen AM, Monti MM , 2014 Optimized brain extraction for pathological brains (optiBET). *PLoS One* 9 (12), e115551. [PubMed: 25514672]
- Madden DJ, Pierce TW, Allen PA, 1993 Age-related slowing and the time course of semantic priming in visual word identification. *Psychol. Aging* 8 (4), 490–507. doi:10.1037/0882-7974.8.4.490. [PubMed: 8292278]
- Maylor EA, 1990 Age, blocking and the tip of the tongue state. *Br. J. Psychol* 81, 123–134. [PubMed: 2364243]
- Moreno-Martínez FJ, Montoro PR, 2012 An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms for seven psycholinguistic variables. *PLoS One* 7 (5), e37527. [PubMed: 22662166]
- Mortensen L, Meyer AS, Humphreys GW, 2006 Age-related slowing of object naming: a review. *Lang. Cognit. Processes* 21, 238–290.
- Munson B, Solomon NP, 2004 The effect of phonological neighborhood density on vowel articulation. *J. Speech Lang. Hear. Res* 47, 1048–1058. doi:10.1044/1092-4388(2004/078). [PubMed: 15605431]
- Newman RS, German DJ , 2005 Life span effects of lexical factors on oral naming. *Lang. Speech* 48 (2), 123–156. [PubMed: 16411502]

- Ossher L, Flegal KE, Lustig C, 2013 Everyday memory errors in older adults. *Aging, Neuropsychol. Cognition* 20, 220–242.
- Patterson J, 2011 Controlled oral word association test In: Kreutzer JS, DeLuca J, Caplan B (Eds.), *Encyclopedia of Clinical Neuropsychology*. Springer, New York, NY.
- Peramunage D, Blumstein SE, Myers EB, Goldrick M, Baese-Berk M, 2011 Phonological neighborhood effects in spoken word production: an fMRI study. *J. Cogn. Neurosci* 23 (3), 593–603. doi:10.1162/jocn.2010.21489. [PubMed: 20350185]
- Perez MA, 2007 Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *Q. J. Exp. Psychol* 60 (1), 32–42. doi:10.1080/17470210600577423.
- Rajah MN, Maillet D, Grady CL, 2015 Episodic memory in healthy older adults In: Addis DR, Barense M, Duarte A (Eds.), *The Wiley Handbook on the Cognitive Neuroscience of Memory*. Wiley-Blackwell, Oxford, pp. 347–370.
- Rastle KG, Burke DM, 1996 Priming the tip of the tongue: Effects of prior processing on word retrieval in young and older adults. *J. Memory Language* 35, 585–605.
- Reuter-Lorenz PA, Cappell KA, 2008 Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci* 17 (3), 177–182. doi:10.1111/j.1467-8721.2008.00570.x.
- Rieck JR, Rodrigue KM, Boylan MA, Kennedy KM, 2017 Age-related reduction of BOLD modulation to cognitive difficulty predicts poorer task accuracy and poorer fluid reasoning ability. *Neuroimage* 15 (147), 262–271. doi:10.1016/j.neuroimage.2016.12.022.
- Sadat J, Martin CD, Costa A, Alario FX, 2014 Reconciling phonological neighborhood effects in speech production through single trial analysis. *Cognit. Psychol* 68, 33–58. doi:10.1016/j.cogpsych.2013.10.001. [PubMed: 24291531]
- Sakai K, Hikosaka O, Takino R, Miyauchi S, Nielsen M, Tamada T, 2000 What and when: parallel and convergent processing in motor control. *J. Neurosci* 20 (7), 2691–2700. 10.1523/JNEUROSCI.20-07-02691.2000 [PubMed: 10729350]
- Scarborough R, 2013 Neighborhood-conditioned patterns in phonetic detail: Relating coarticulation and hyperarticulation. *J. Phonetics* 41 (6), 491–508. doi:10.1016/j.wocn.2013.09.004.
- Scarborough R, Zellou G, 2013 Clarity in communication: “clear” speech authenticity and lexical neighborhood density effects in speech production and perception. *J. Acoust. Soc. Am* 134 (5), 3793–3807. doi:10.1121/1.4824120. [PubMed: 24180789]
- Shafto MA, Burke DM, Stamatakis EA, Tam PP, Tyler LK, 2007 On the tip-of-the-tongue: neural correlates of increased word-finding failures in normal aging. *J. Cogn. Neurosci* 19 (12), 2060–2070. doi:10.1162/jocn.2007.19.12.2060. [PubMed: 17892392]
- Shafto MA, Stamatakis EA, Tam PP, Tyler LK, 2010 Word retrieval failures in old age: the relationship between structure and function. *J. Cogn. Neurosci* 22 (7), 1530–1540. doi:10.1162/jocn.2009.21321. [PubMed: 19642890]
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, ..., Flitney DE, 2004 Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–S219. [PubMed: 15501092]
- Snodgrass JG, Vanderwart M, 1980 A standardized set of 200 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. [Hum. Learn.]* 6, 174–215.
- Stemberger JP, 2004 Neighbourhood effects on error rates in speech production. *Brain Lang.* 90, 413–422. doi:10.1016/S0093-934X(03)00452-8. [PubMed: 15172557]
- Stojanoski B, Cusack R, 2014 Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of Vision* 14 (12), 1–16 doi:10.1167/14.12.6.
- Stroop JR, 1935 Studies of interference in serial verbal reactions. *J. Exp. Psychol* 18, 643–662.
- Tabak W, Schreuder R, Baayen RH, 2010 Producing inflected verbs: a picture naming study. *Mental Lexicon* 5 (1), 22–46. doi:10.1075/ml.5.1.02tab.
- Team RC (2017). R: A language and environment for statistical computing. Retrieved from <https://www.R-project.org/>
- Tsai P, 2018 Phonological neighborhood effect in spontaneous speech in adults who stutter. *J. Fluency Disord* 58, 86–93. doi:10.1016/j.jfludis.2018.08.005. [PubMed: 30180994]

- Vaden KI, Muftuler LT, Hickok G, 2010 Phonological repetition-suppression in bilateral superior temporal sulci. *Neuroimage* 49 (1), 1018–1023. doi:10.1016/j.neuroimage.2009.07.063, doi: [PubMed: 19651222]
- Vitevitch MS, 1997 The neighborhood characteristics of malapropisms. *Lang. Speech* 40 (Pt 3), 211–228. doi:10.1177/002383099704000301. [PubMed: 9509578]
- Vitevitch MS, 2002 The influence of phonological similarity neighborhoods on speech production. *J. Exper. Psychol* 28 (4), 735–747. doi:10.1037/0278-7393.28.4.735.
- Vitevitch MS, Luce PA, 2016 Phonological neighborhood effects in spoken word perception and production. *Ann. Rev. Linguistics* 2, 75–94. doi:10.1146/annurev-linguistics-030514-124832.
- Vitevitch MS, Sommers MS, 2003 The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Mem. Cogn* 31 (4), 491–504.
- Vitevitch MS, Stamer MK, 2006 The curious case of competition in Spanish speech production. *Lang. Cognitive Processes* 21, 760–770.
- Wechsler D. Coalson DL, Raiford SE, 1997 WAIS-III: Wechsler Adult Intelligence Scale. Psychological Corporation San Antonio, TX.
- Woods SP, Delis DC, Scott JC, Kramer JH, Holdnack JA, 2006 The California verbal learning test—second edition: test-retest reliability, practice effects, and reliable change indices for the standard and alternate forms. *Arch. Clin. Neuropsychol* 21 (5), 413–420. [PubMed: 16843636]
- Woolrich MW, Behrens TE, Beckmann CF, Jenkinson M, Smith SM, 2004 Multilevel linear modelling for FMRI group analysis using Bayesian inference. *Neuroimage* 21 (4), 1732–1747. [PubMed: 15050594]
- Worsley K. 2001 Statistical analysis of activation images. *Funct. MRI* 14, 251–270.
- Wright RE, 2004 Factors of lexical competition in vowel articulation In: Local J, Ogden R (Eds.), *Papers in Laboratory Phonology VI*. Cambridge University Press, Cambridge, pp. 75–87.
- Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO, 1982 Development and validation of a geriatric depression screening scale: a preliminary report. *J. Psychiatr. Res* 17 (1), 37–49. [PubMed: 7183759]
- Zhang H, Carlson MT, Diaz MT, 2019 Investigating the effects of phonological neighbours on word retrieval and phonetic variation in word naming and picture naming paradigms. *Lang. Cognit. Neurosci* doi:10.1080/23273798.2019.1686529.
- Zhang H, Eppes A, Diaz MT, 2019 Task difficulty modulates age-related differences in the behavioral and neural bases of language production. *Neuropsychologia* 124, 254–273. 10.1016/j.neuropsychologia.2018.11.017 [PubMed: 30513288]

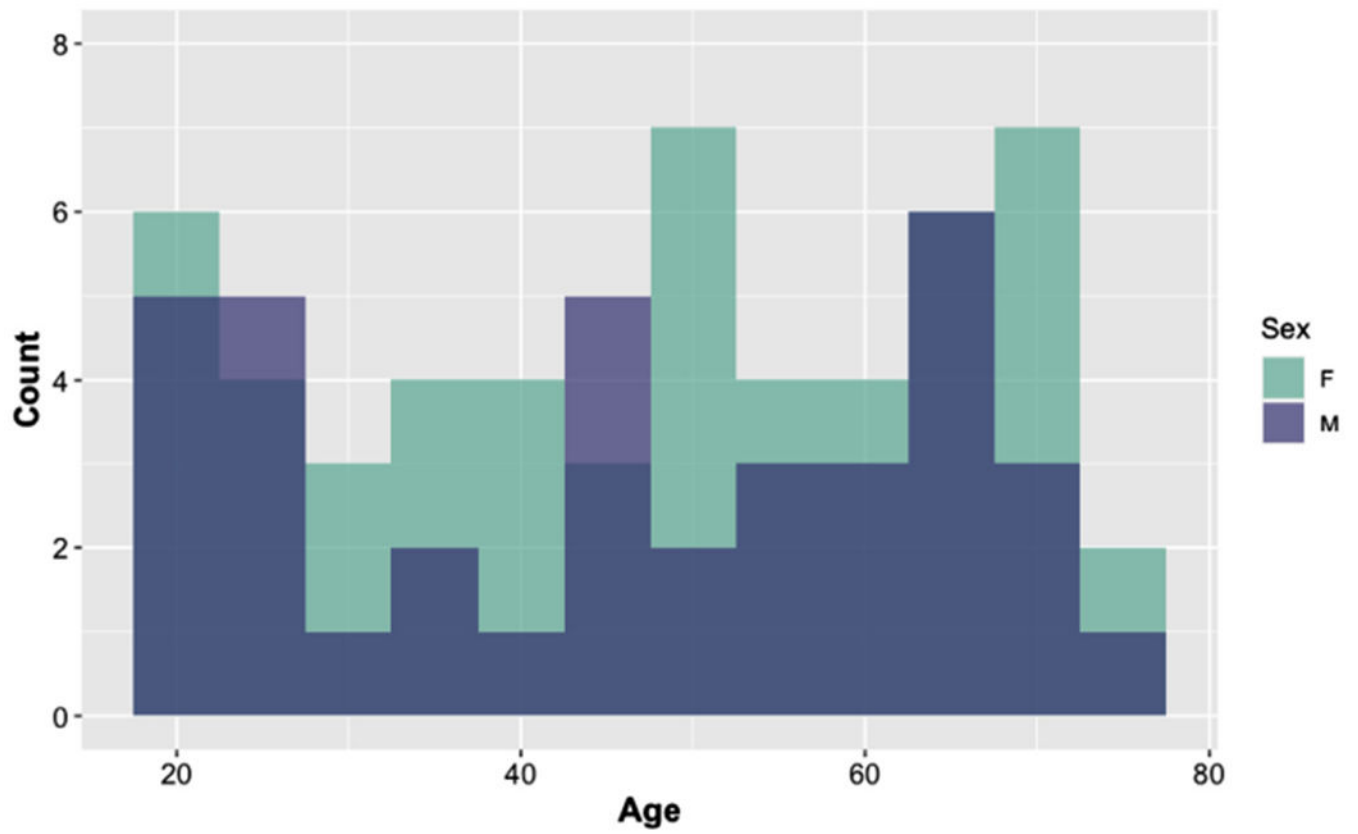


Fig. 1. Distribution of participant ages. Ninety-one adults, roughly evenly distributed across the lifespan, participated in the experiment (age range = 20 – 75, mean age = 47.40, SD = 17.45, 54 Females).

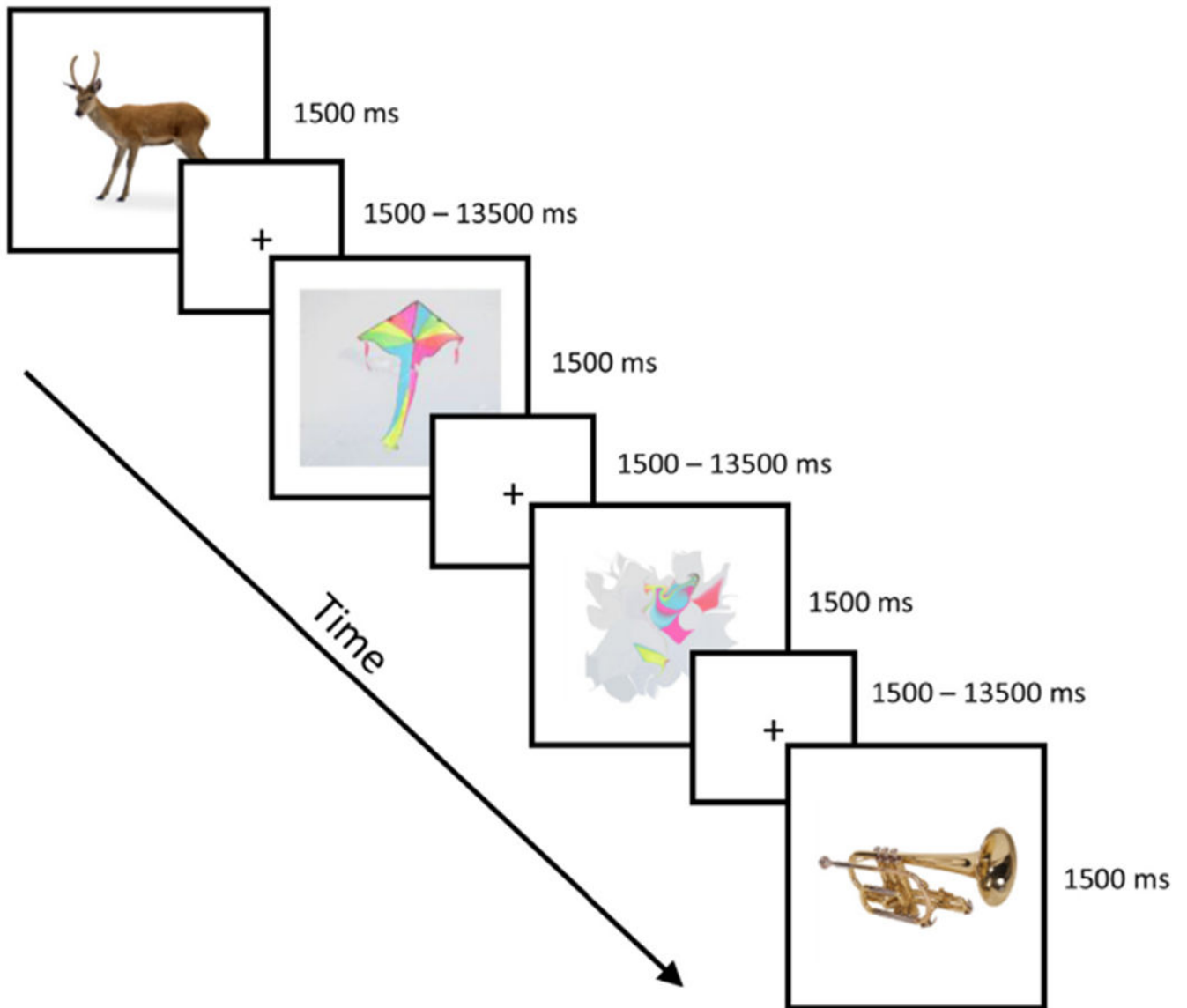


Fig. 2. An overview of the experimental procedure. Participants overtly named pictures that varied in phonological neighborhood density (e.g., deer, kite = high PND, trumpet = low PND), as well as abstract pictures.

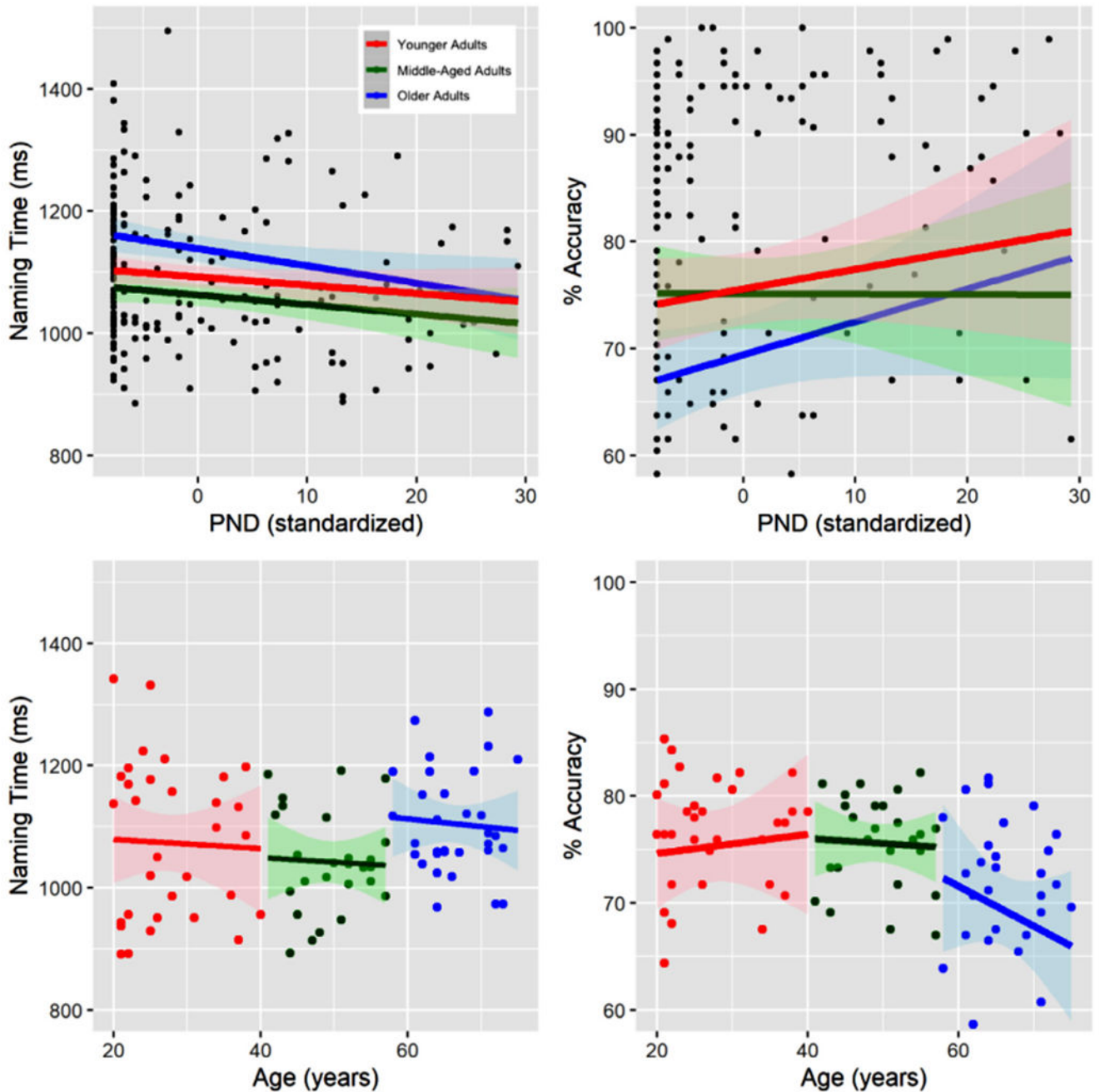


Fig. 3. Behavioral results for picture naming. The colors reflect the data split across three age groups (see the methods for full details, red dots and lines = younger adults, green dots and lines = middle-aged adults, and blue dots and lines = older adults). Top panels: Effects of PND on naming time and accuracy are shown (grouped by item, $N = 191$), increasing PND is associated with faster naming times and higher accuracies. Additionally, there was an interaction between PND and Age on naming time suggesting that older adults were more impaired by items with small PNDs. Bottom panels: Effects of Age on naming time and

accuracy are shown (grouped by individual, $N = 91$). There were significant effects of age on naming time and accuracy, and a marginally significant quadratic trend of age on accuracy.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

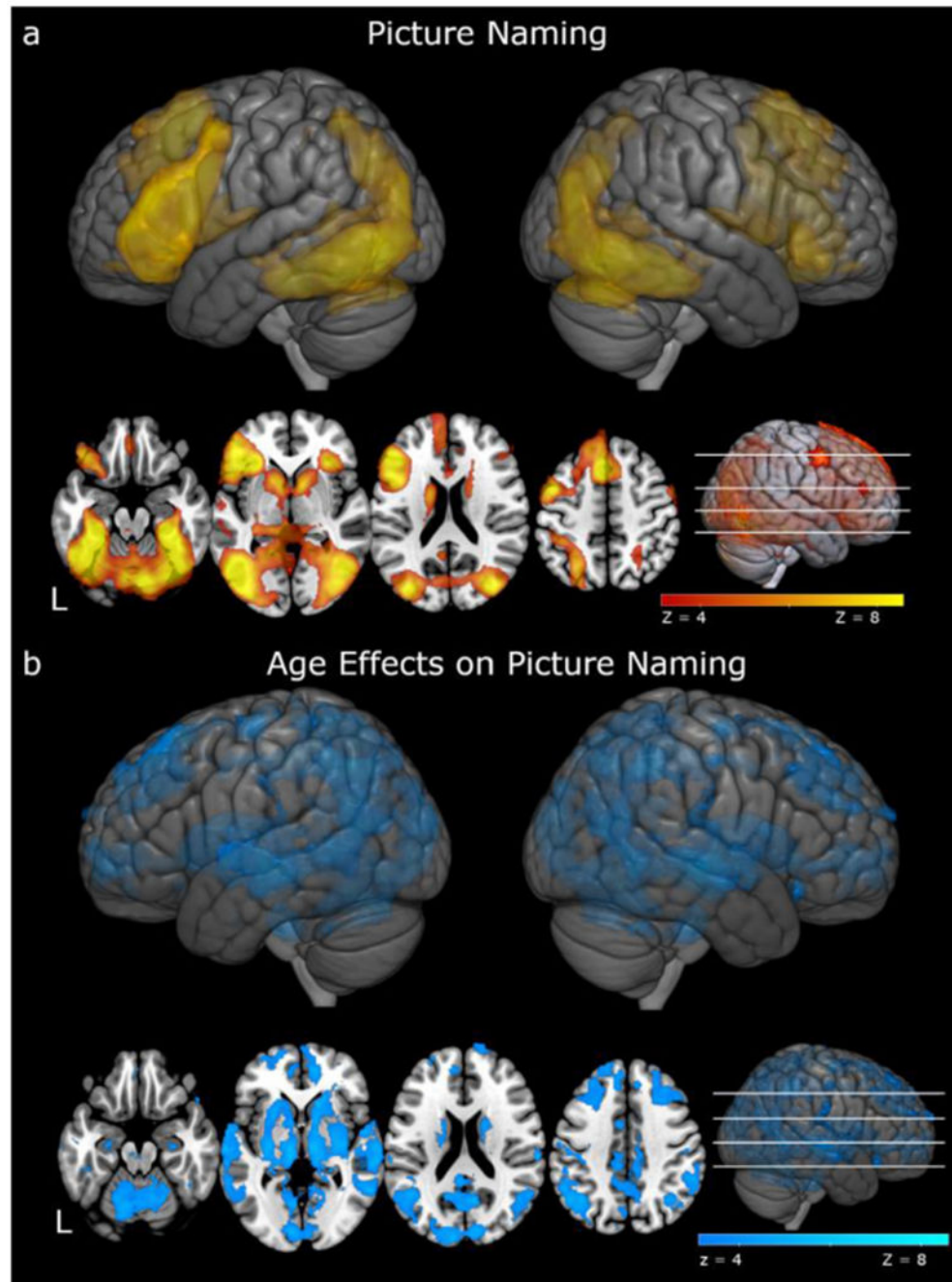


Fig. 4. fMRI activation to picture naming. a. red-yellow activations represent regions of significant activation to picture naming, collapsing across age. b. blue activations illustrate regions where age was positively correlated with picture naming activation. There were no regions in which age was negatively correlated with picture naming. In both a and b, the color bar represents statistically significant activations from $Z = 3.1$ to $Z = 8.6$. Axial slices show activation at $Z = -20, 0, 20,$ and 40 .

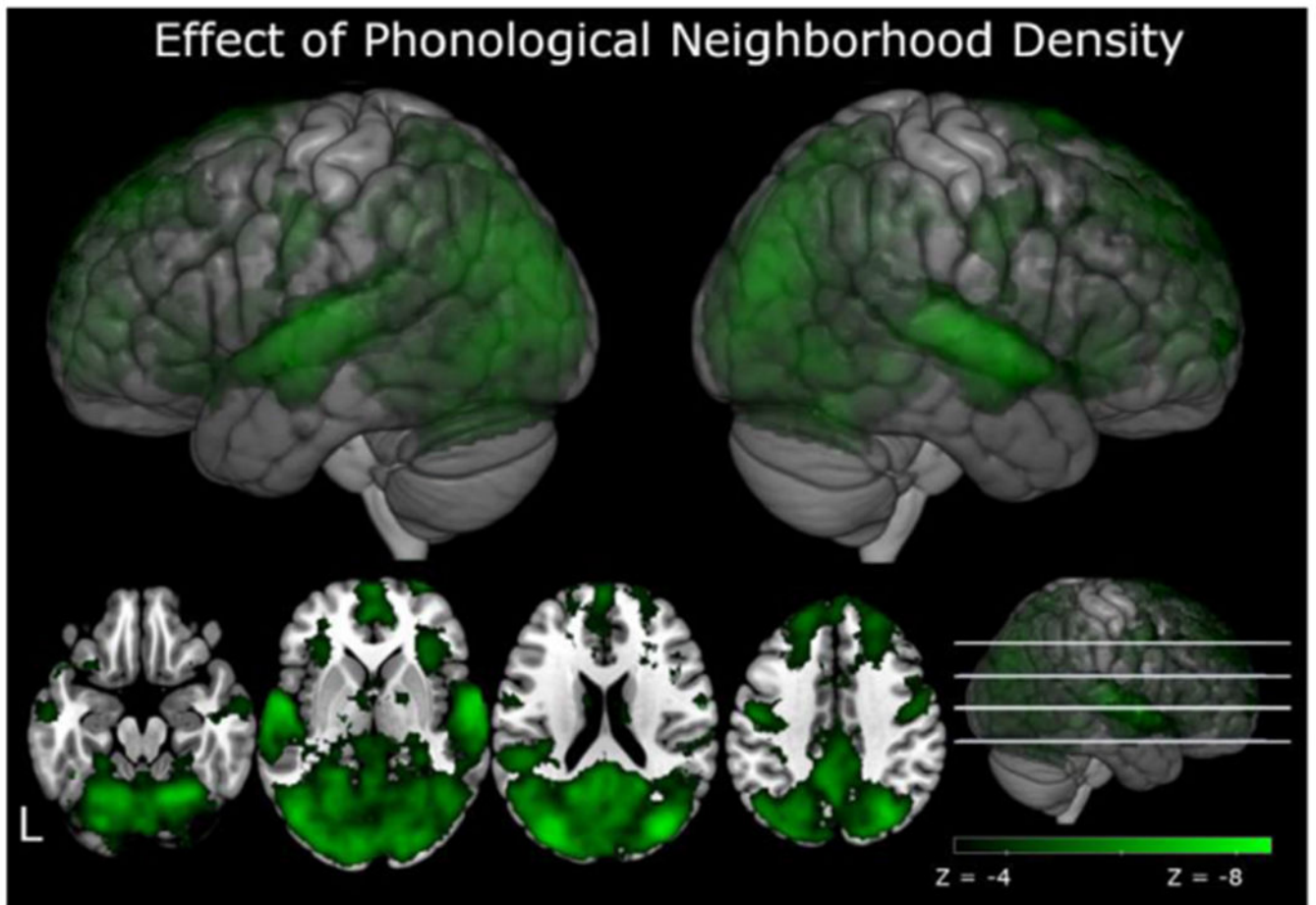


Fig. 5. fMRI activation to PND. Green activations represent regions in which PND was significantly negatively correlated with fMRI activation. That is, as PND increased activation decreased. There were no significant positive correlations between PND and fMRI activation. The color bar represents statistically significant activations from $Z = -3.1$ to $Z = -8.6$. Axial slices show activation at slices $Z = -20, 0, 20,$ and 40 .

Table 1

Participant demographic and neuropsychological testing scores.

Demographic information	
N	91
Age (Years)	47.40 (17.45)
Gender (M/F)	37/54
Education (Years)	16.9 (2.5)
Correlation of age with cognitive assessments Education	0.24*
MMSE	-0.19
Depression (GDS)	-0.13
Speed RT (choice)	0.56***
Digit Symbol RT	0.69***
Digit Span Forward	-0.20
Digit Span Backward	-0.26*
Stroop Effect	0.33***
Verbal Working Memory	-0.40***
Immediate Recall	-0.27*
Delayed Recall	-0.26*
Category Fluency	-0.30**
Phonemic Fluency	-.11
WAIS Vocabulary	0.07
Author Recognition	0.47***

Values provided are means, with standard deviation in parentheses.

* $p < .05$,

** $p < .01$,

*** $p < .001$.

Table 2

Coordinates for peak and sub-peaks of picture naming activation.

Region	Hemisphere	Voxels	Max Z	X	Y	Z
<i>Picture-Naming Activation</i>						
Lateral occipital cortex	Left	61,921	12.2	-46	-66	-6
Lateral occipital cortex	Right	—	11.4	50	-64	-10
Inferior temporal gyrus	Left	—	12.0	-48	-54	-14
Parahippocampus	Left	—	9.4	-30	-30	-19
Parahippocampus	Right	—	8.7	28	-36	-14
Hippocampus	Left	—	7.2	-23	-20	-12
Hippocampus	Right	—	5.9	22	-20	-12
Anterior cingulate	—	—	7.3	0	6	32
Precentral gyrus	Left	—	8.9	-48	2	52
Superior frontal gyrus	Left	—	11.4	-4	14	56
Insula	Left	—	9.9	-30	24	3
Insula	Right	—	8.5	32	26	2
IFG, pars opercularis	Left	—	10.8	-41	11	28
IFG, pars triangularis	Left	—	10.2	-48	34	16
Orbital frontal cortex	Left	—	10.9	-30	34	-12
Orbital frontal cortex	Right	—	8.9	34	28	-6
Caudate	Left	—	7.9	-10	8	6
Caudate	Right	—	7.7	9	8	4
Cerebellum	Right	—	10.7	10	-75	-24
Medial Frontal Cortex	Left	586	7.1	-6	54	-12
Frontal pole	Left	—	4.8	-18	50	-14
Precentral gyrus	Right	433	7.4	58	-8	50
Postcentral gyrus	Right	—	4.9	60	-18	52
<i>Age Effects in Picture-Naming</i>						
Thalamus	Right	40,872	11.7	6	-18	-4
Thalamus	Left	—	7.1	-14	-16	6
Lateral occipital cortex	Left	—	4.8	-51	-64	16

Region	Hemisphere	Voxels	Max Z	X	Y	Z
Lateral occipital cortex	Right	—	4.7	47	-63	16
Planum temporale	Left	—	9.9	-54	-32	8
Hippocampus	Left	—	5.3	-30	-24	-11
Hippocampus	Right	—	8.8	28	-16	-18
Parietal operculum	Right	—	4.4	62	-23	18
Middle temporal gyrus	Left	—	5.3	-64	-23	-6
Middle temporal gyrus	Right	—	5.5	64	-18	-5
Precentral gyrus	Left	—	5.2	-54	-16	11
Cerebellum	Right	—	9.3	28	-44	-28
Frontal pole	Left	7,200	7.9	-18	56	38
Frontal pole	Right	—	6.8	12	64	-4
Superior frontal gyrus	Left	—	7.9	-22	22	58
Precentral gyrus	Right	—	7.5	40	-12	64

Note. The voxel-wise Z threshold = 3.1, $p < .001$. Clusters with a corrected significance of $p < .05$ were retained. Sub-peaks are included to provide a more thorough characterization of the functional activation. IFG = inferior frontal gyrus

Table 3

Coordinates for peak and sub-peaks of PND activation.

Region	Hemisphere	Voxels	Max Z	X	Y	Z
Heschl's gyrus	Right	87,207	8.9	52	-16	6
Heschl's gyrus	Left	—	8.7	-48	-22	8
Lateral occipital cortex	Left	—	8.1	-30	-81	24
Lateral occipital cortex	Right	—	8.0	34	-79	20
Fusiform gyrus	Left	—	6.2	-23	-61	-11
Fusiform gyrus	Right	—	5.6	30	-56	-8
STG – posterior	Left	—	6.7	-66	-24	6
STG – posterior	Right	—	6.9	63	-24	6
Precentral gyrus	Left	—	6.1	-48	-9	37
Precentral gyrus	Right	—	5.7	44	-9	36
STG – anterior	Left	—	7.6	-58	-1	-2
STG – anterior	Right	—	7.7	64	-3	1
Orbital frontal cortex	Right	—	5.4	35	32	-4
Frontal pole	Left	—	5.4	-1	66	-1

Note. The voxel-wise Z threshold = 3.1, $p < .001$. Clusters with a corrected significance of $p < .05$ were retained. Sub-peaks are included to provide a more thorough characterization of the functional activation. STG = Superior Temporal Gyrus