nature communications

Article

https://doi.org/10.1038/s41467-024-51514-3

Underlying dimensions of real-time word recognition in cochlear implant users

Received: 7 June 2023

Accepted: 8 August 2024

Published online: 29 August 2024

Check for updates

Bob McMurray[1,2,3,4], Francis X. Smith[1,2], Marissa Huffman[3], Kristin Rooff[3], John B. Muegge[1], Charlotte Jeppsen[1], Ethan Kutlu[1,4] & Sarah Colby[1,3]

Word recognition is a gateway to language, linking sound to meaning. Prior work has characterized its cognitive mechanisms as a form of competition between similar-sounding words. However, it has not identified dimensions along which this competition varies across people. We sought to identify these dimensions in a population of cochlear implant users with heterogenous backgrounds and audiological profiles, and in a lifespan sample of people without hearing loss. Our study characterizes the process of lexical competition using the Visual World Paradigm. A principal component analysis reveals that people's ability to resolve lexical competition varies along three dimensions that mirror prior small-scale studies. These dimensions capture the degree to which lexical access is delayed ("Wait-and-See"), the degree to which competition fully resolves ("Sustained-Activation"), and the overall rate of activation. Each dimension is predicted by a different auditory skills and demographic factors (onset of deafness, age, cochlear implant experience). Moreover, each dimension predicts outcomes (speech perception in quiet and noise, subjective listening success) over and above auditory fidelity. Higher degrees of *Wait-and-See* and *Sustained-Activation* predict poorer outcomes. These results suggest the mechanisms of word recognition vary along a few underlying dimensions which help explain variable performance among listeners encountering auditory challenge.

As the world's population ages, hearing loss and cognitive decline are becoming issues of major importance. These issues are intertwined: mounting evidence suggests that hearing loss – and the consequent loss of speech understanding – is a crucial (but remediable) factor in cognitive decline[1–6]. Despite the importance of speech comprehension as a product of hearing, our understanding of the cognitive mechanisms of speech understanding has not yielded theories that generalize across variation in hearing loss, age, or other demographic factors. Nowhere is this variation more apparent than in people who are profoundly deaf and use cochlear implants (CIs). These devices convert the natural acoustic input to a pattern of electrical activity across a small number of stimulating sites on the cochlea. This is a profoundly different input than the auditory system typically receives, requiring some adaptation at central levels[7,8]. The present study thus leverages a highly variable sample of CI users, using new approaches to uncover the fundamental dimensions of a critical aspect of language processing.

We focus on the recognition of isolated words. Words lie at the core of language, and recognizing a word allows access to its pronunciation, syntax, and meaning. Thus, it is not surprising that deficits in word recognition are observed in language disorders like dyslexia, developmental language disorder, and autism[9,10]. Isolated word recognition is a common basis of clinical tests of hearing[11] and vocabulary[12], and presenting words without context allows us to

[1]Dept. of Psychological & Brain Sciences, University of Iowa, Iowa City, IA, USA. [2]Dept. of Communication Sciences & Disorders, University of Iowa, Iowa City, IA, USA. [3]Dept. of Otolaryngology—Head and Neck Surgery, University of Iowa, Iowa City, IA, USA. [4]Dept. of Linguistics, University of Iowa, Iowa City, IA, USA. e-mail: bob-mcmurray@uiowa.edu

isolate true lexical processing (a potential bottleneck in language) from broader sentence or discourse processes that may compensate.

Moreover, isolated word recognition offers a clear theoretical platform for this investigation. While word recognition is a product of many processes (e.g., auditory and semantic), and is affected by many factors (properties of the words, and properties of the listener), decades of work in cognitive science has focused on a key aspect of the problem: the problem of of hearing a word as it unfolds over time and selecting the target word from an array of similar sounding competitors. This work has established mechanisms by which modal listeners (neurotypical, monolingual, normal hearing adults) map auditory input to stored representations of the sound pattern of a word (a wordform) in the mental lexicon[13–15]. Research on small samples of various specialized populations has observed that this process differs across a variety of populations (e.g., language disorders, development[16]), and listening contexts (e.g., noise or quiet speech[17–19]). However, at a theoretical level we have not yet identified the underlying dimensions along which these mechanisms vary; that is, we do not yet know what aspects of the mechanisms identified for modal listeners vary systematically across listeners, or even how many dimensions there may be.

Understanding these dimensions is necessary for more universal, inclusive, and generalizable theories that describe how basic cognitive mechanisms vary across people[16]. That is, if the mechanisms of word recognition vary along a small number of dimensions, these are the necessary degrees of freedom in cognitive models. Moreover, the mechanisms that CI users employ to adapt to degraded input may be similar to those used by normal hearing (NH) listeners in challenging listening situations like noise[18–20]. This creates further opportunities to generalize theories of language processing. Understanding these dimensions may also help clinical care of people with speech and language disorders; it can inform novel assessments of language and hearing as well as inform treatment. Finally, identifying these dimensions is relevant to understanding the relationship between hearing loss and dementia[1–6]. One prominent account suggests that social engagement plays a crucial role in maintaining cognitive function[21–29]. However, difficulties in language processing, which declines with age[30], could compound the impact of hearing loss on social disengagement; alternatively, strong language skills could buffer the functional consequences of hearing loss.

The present study thus sought to uncover the dimensions by which the basic processes of wordform recognition vary in a population that is (1) highly relevant to ongoing concerns about aging; (2) is characterized by high heterogeneity in outcomes; and (3) which faces significant barriers to language comprehension: profoundly deaf individuals who use cochlear implants.

About 15% of U.S. adults are affected by hearing loss[31], which impedes social functioning and can lead to cognitive decline[32]. CIs restore access to sound and improve social function for most profoundly deaf listeners[33–35]. However, not all CI users perform well in the real world and there is substantial unexplained variability in outcomes[36–43]. A key predictor is peripheral auditory factors. The health of the auditory nerve, the nature of implantation, and access to residual acoustic hearing all affect how well sound is transmitted via the CI to the auditory system. However, the peripheral auditory system alone cannot fully explain differences in outcomes. CIs provide a profoundly different input than a Normal Hearing (NH) ear. The CI collapses thousands of frequencies into a small number of electrodes; it also eliminates temporal fine structure and even some entire frequency ranges. Consequently, successful speech perception requires CI users to learn to cope with this new form of degraded input and the fundamental uncertainty it imposes at more central or cognitive levels over the first year of CI experience[7,8,44].

Several studies link variation in general cognitive abilities (e.g., working memory) to speech perception[45–48] and self-report measures of real world success[49] (though see ref. 50). There is also mounting evidence that people with hearing impairment engage more "cognitive effort" to perceive speech[51–53]. This work emphasizes the importance of cognition but does not provide a clear explanation for how fundamental language comprehension processes differ for people confronting the challenge of listening with a CI.

Cognitive science frames the mechanisms of isolated wordform recognition in terms of temporary ambiguity. Because words unfold over time, all listeners face a brief period of ambiguity. For instance, at the onset of *basket*, the input (*ba-*) is consistent with hundreds of words (*batter, back, bathtub*, and so forth). In NH adults, ambiguity is resolved via a process of immediate competition. At the beginning of the word, listeners activate a set of candidates that match the partial input. This set is continuously winnowed until only one remains Fig. 1A[13];. This competition does not derive solely from the accruing input. Words are considered that do not entirely match the input or should have been ruled out by earlier information[54–56] and are further affected by inhibition from neighboring words[57–59]. Thus, lexical competition is a cognitive mechanism that balances demands of accuracy, efficiency, and flexibility[10].

The real-time dynamics of competition are commonly studied using eye-tracking in the Visual World Paradigm (VWP)[60]. In the VWP, listeners hear a spoken word (e.g., *basket*) and select its referent from an array of pictures including the target (*basket*), potential competitors (e.g., onset competitors [cohorts] like *batter*, or rhymes like *casket*), and an unrelated word. To perform this task, listeners must find the target. Listeners typically make 3-5 eye-movements before responding. As eye-movements unfold, fixations to different competitors reveal the likelihood of considering various classes of words over time (Fig. 1A). These patterns of fixations align closely with computational models of word recognition[60,61].

While the VWP can characterize many aspects of word recognition[62,63], this variant—which focuses on competition among candidates—has been influential because of its ability to capture the most important mechanism that undergird most theories of word recognition: competition[15,64,65]. It has been in wide use across populations including children[66,67], older adults[30], people with developmental language disorder[61], and multilinguals[68], as well as NH listeners in challenging conditions[17,18]. Thus, it offers a consensus diagnostic of how the competition process that underlies word recognition varies across listeners.

Several small-scale studies[19,20,69] have used the VWP to characterize the dynamics of lexical competition in CI users. These illustrate two processing profiles, termed *Sustained Activation* and *Wait-and-See*, which offer initial hypotheses. Both profiles are also observed in other hearing impaired populations[70] and NH listeners in challenging conditions[17–20]. Thus, they may offer a generalized description of how word recognition adapts to challenge.

Post-lingually deaf CI users often exhibit a profile—now termed *Sustained Activation*—in which word recognition is slowed and competition never fully resolves (Fig. 1B). Even late in processing, these listeners do not fully commit to the target, and continue to fixate competitors[19,71]. Similar profiles are observed in NH listeners in moderate levels of noise[17], distortion[19], or with unfamiliar dialects[72]. This profile does not entirely derive from poor encoding. CI users can accurately encode fine-grained speech cues, and still show *Sustained-Activation*[69].

It is unclear if this is functional. One possibility is that *Sustained-Activation* is an automatic consequence of poor input: the degraded input from the CI does not afford enough information to fully discriminate words, so competitors cannot be fully ruled out. Alternatively, *Sustained-Activation* may support more flexible listening. Listeners may keep candidates available, in case they need to revise an earlier choice and reactivate a competitor c.f.[73,74],. Supporting this, post-lingually deaf CI users show less disruption than NH listeners

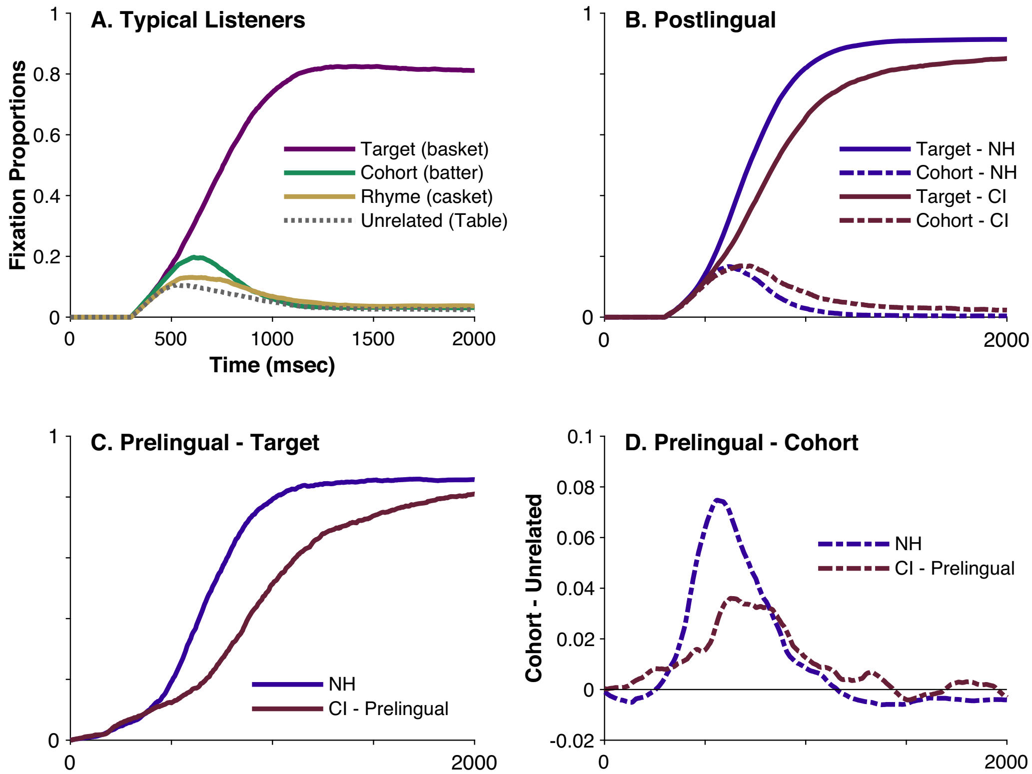

Fig. 1 | Typical results of VWP experiments for Normal Hearing (NH) listeners and Cochlear Implant (CI) users. A Fixations to targets and competitors as a function of time for modal adult listeners. At each moment, the amount of fixations reflects the degree to which the listener is considering (activating) that class of word as they settle on the correct item[19]. **B** Fixations to targets and cohort competitors in NH adults and in post-lingually deaf adult CI users[19]. CI users are slower to fully commit to the target and rule out competitors, and they continue fixating competitors even when they've selected the target, a *Sustained Activation* profile. **C**, **D** Pre-lingually deaf adolescent CI users show a *Wait and See* profile[20], with much larger delays in target fixations (**C**). Because lexical access is delayed, cohorts show less competition (**D**) – by the time they begin lexical access for *wizard* they have heard some information to rule out *window*.

when recognizing speech that mismatches its expected form (e.g., hearing *tog* instead of *dog*). It is not yet clear if *Sustained-Activation* is helpful for more general outcomes.

This profile known as *Wait-and-See* was first observed in pre-lingually deaf CI users[20,70]. Listeners delay lexical access by nearly the length of a word (Fig. 1C). By the time lexical access begins, they have thus heard most of the word, and consequently show less onset competition (Fig. 1D). *Wait-and-See* has also been seen in children with moderate hearing loss who use hearing aids[70], and in NH adults hearing severely distorted[20] or quiet[18] speech. It is unclear why listeners wait and see. One possibility is that the input is so degraded, there is not enough information to begin lexical access. However, children with mild-to-moderate hearing loss also show *Wait-and-See*, despite nearly perfect accuracy[70]. Alternatively, *Wait-and-See* may support more accurate listening. By delaying lexical access, listeners accrue more input before committing to a word, minimizing competition and the chance of an error.

The previously discussed studies generate hypotheses for how the mechanisms of word recognition differ in CI users. However, their small samples precluded any analysis of individual differences that could link these profiles to outcomes (e.g., to determine if it is beneficial to wait-and-see) or identify factors that lead listeners to sustain activation or wait and see. The present study thus incorporated a larger and thoroughly characterized sample of CI users ($N = 101$), alongside new analyses of a previously reported lifespan sample of listeners without major hearing loss ($N = 107$)[30], to address three questions.

First, we sought to characterize the dimensionality of these differences. One hypothesis is that listeners show *Sustained-Activation* in response to moderate challenge and *Wait-and-See* for more severe challenge. For example, prior work has shown *Sustained-Activation* with 8-channel vocoding but *Wait-and-See* with 4-channel[19,20]. Under this view, these profiles are two points along one dimension of difficulty. Alternatively, these profiles may represent independent dimensions derived from different sources and serving different functional goals[71], or these profiles may not characterize the underlying dimensions at all—word recognition may vary in ways not previously detected.

Second, we asked if listeners without major hearing loss can be described along the same dimensions by leveraging a sample tested with identical tasks as part of an independent study on lifespan aging[30]. This sample was not intended as a direct comparison to the sample of CI users (participants were not purposely matched to the CI users on factors like age), but rather an opportunity to extend the analysis to a new group (see Supplementary Note 5).

Third, it is unclear what leads listeners to exhibit variation along these dimensions or to exhibit each profile. Likely factors include deafness onset (pre- vs. post-lingual)[70], development and aging[30,66], as well as auditory factors (e.g., how well the CI encodes spectral or temporal differences). Critically, we must also rule out non-linguistic differences in factors like oculomotor control to document that these reflect dimensions of word recognition.

Fourth, it is unclear whether these profiles are functional or reflect a bottom-up response to poor input. To address this requires an analysis that accounts for the quality of the auditory periphery, which likely impacts both real-time word recognition and outcomes (e.g., people with poor spectral resolution may be more likely to wait and see and be more likely to show poor outcomes). Specifically, we ask if the degree to which a CI user exhibits *Wait-and-See* or *Sustained-Activation* predicts their overall ability to perceive speech (outcomes) *over and above* auditory fidelity. If these profiles are functionally adaptive, a stronger profile (e.g., more *wait-and-see*) should predict better outcomes. In contrast, these profiles could negatively predict outcomes. This could occur if these profiles are not causally related to outcomes, but instead are a marker that listening is challenging in general. A negative relationship could also be observed if these profiles do relate to outcomes, but listeners adopt them when they should not.

We examined 101 CI users (Table 1, Supplementary Table S1) including both pre-lingually and post-lingually deaf individuals who used a variety of device configurations, including 57 with Functional Acoustic Hearing (FAH) in at least one ear (which benefits speech perception[75,76]). Participants were tested in a standard version of the VWP that captured the dynamics of competition that underlie the recognition of isolated wordforms. This was modeled after our prior work[19,20] and extensive work by others[60], and a broadly similar approach has been applied to other groups[67,68,77] and listening conditions[17,78]. Our specific paradigm underwent extensive psychometric validation (see Methods). We characterized the fidelity of peripheral encoding along several dimensions: pure-tone audiometry to assess functional acoustic hearing, a spectral ripple task to assess spectral fidelity[79,80], an envelope task to assess temporal fidelity[81], and common device factors (e.g., the use of one vs. two CIs). We assessed outcomes with standard clinical measures: CNC words in quiet[11] and AzBio sentences in noise[82], and with self-report measures of listening success[83].

Using this dataset, the present study identified three underlying dimensions that account for the majority of the variance in real-time lexical competition in cochlear implant users and in listeners without hearing loss. The first two dimensions showed a close correspondence to the *Wait-and-See* and *Sustained-Activation* profiles identified by prior work[19,20] and the third (*Slow-Activation-Rate)* corresponded to previously observed changes due to aging[30]. The degree to which an individual listener shows each dimension was predicted by distinct combinations of demographic and audiological factors. Moreover, each dimension predicted outcomes, even after accounting for the auditory periphery. This suggests that real-time lexical processing may be a unique locus for explaining clinical outcomes, whose cognitive mechanisms may vary in a small number of ways.

Table 1 | Distribution of participants

	Factor	N	Age (SD)
Biological Sex	Female	56	55.6 (15.3)
	Male	45	59.5 (15.2)
Listening Configuration	Bilateral	22	51.5 (17.0)
	Unilateral	79	58.9 (14.5)
	No FAH	44	53.7 (14.9)
	FAH – 1 ear	47	59.5 (15.0)
	FAH – 2 ears	10	53.7 (16.0)
DeafnessOnset	Pre-Lingual	17	39.5 (18.5)
	Intermediate	9	51.0 (16.2)
	Post-Lingual	75	62.3 (10.1)

Results

Figure 2A shows fixations to each of the four types of competitors over time in the VWP task. Early fixations were directed to the target (e.g., *basket*) and cohort (*batter*), before separating around 600 msec. Given that it takes 200 ms to plan and launch and eye-movement, and there was 100 ms between trial onset and the stimulus, this means that these curves functionally separate at around 300 ms after word onset. Fixations to the rhyme (*casket*) did not reach the same peak and were generally slower to rise and fall. As a whole, CI users showed a pattern of lexical competition qualitatively similar to the incremental processing of NH listeners[60]. However, there were large differences across listeners (Fig. 2, bottom row). For example, 699 showed robust cohort fixations and early target fixations, implying more immediate (NH-like) competition. In contrast, 592 showed delayed target fixations, and practically no competition (a hallmark of *Wait-and-See*), and Participant 1517 showed differences at the end of processing with reduced

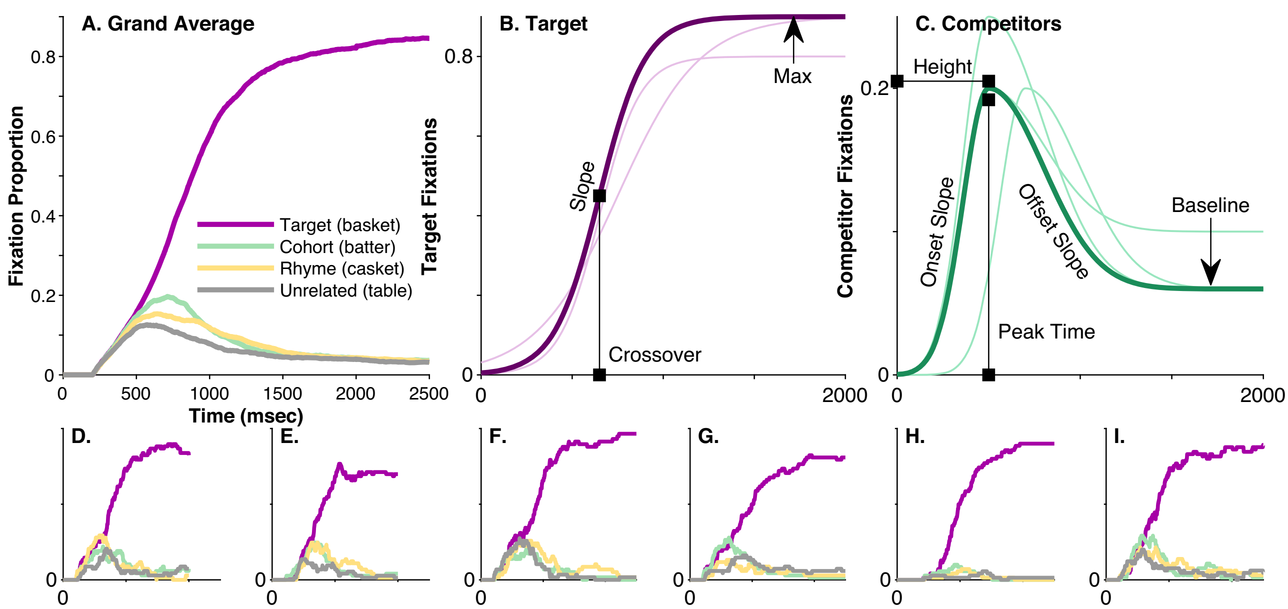

Fig. 2 | Overview of VWP results. A Fixations to the target word (*basket*), an onset competitor (*batter*), a rhyme (*casket*) and an unrelated item as a function of time averaged across all the CI users. **B** Target fixations take the form of a sigmoid that can vary in three dimensions: the *slope* of the transition, the crossover (the time at which the function its midpoint) and the *maximum* at asymptote; **C** Fixations to competitors can vary in five parameters including the onset and offset slopes, the height and time of the peak and the baseline at asymptote. Bottom row) Results from six representative CI users. **D**–**I** Individual listeners (participant # is noted for reference to the shared datasets); **D** Bilateral CI – Post-lingually deaf (#699); **E** Unilateral CI with bilateral FAH – Post-lingually deaf (#1517); **F** Bimodal listener – Post-lingually deaf (#1486); **G** Unilateral CI with bilateral FAH – Post-lingually deaf (#673); **H** Bilateral – Prelingually deaf (#592); **I** Unilateral – Prelingually deaf (#1308).

https://doi.org/10.1038/s41467-024-51514-3

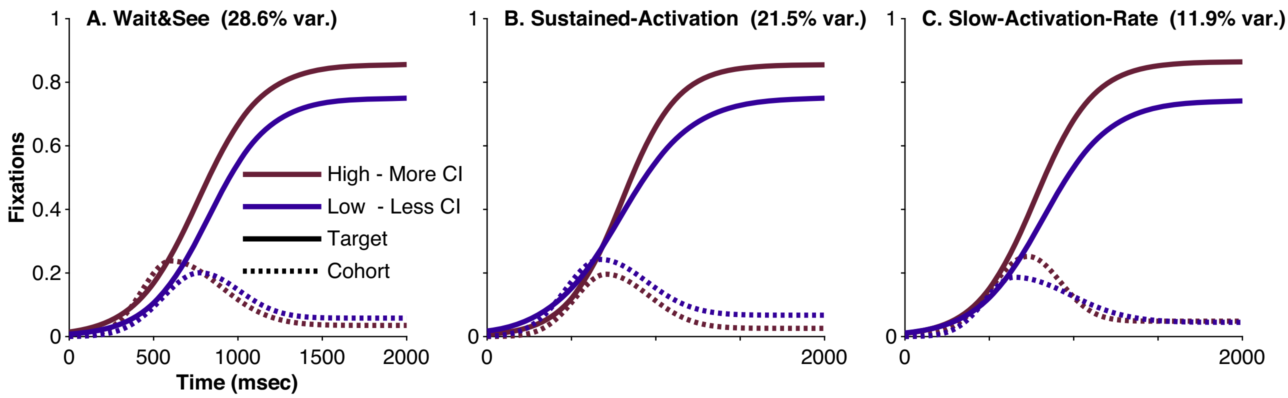

Fig. 3 | Fixation curves reconstructed from the PCA for the lexical processing indices. Shown are the predicted fixations for a listener that is high (+1.5 SD) or low (−1.5 SD) along each discovered dimension of processing. **A** The first principal component described a *Wait&See* profile with delayed fixations to both targets and competitors and a reduced cohort peak. **B** The second component reflect *Sustained-Activation* with slower growth of the function and reduced separation of the target and cohort asymptotes. **C** The final component reflected a slower growth of activation.

target fixations and increased competitor fixations (a hallmark of *Sustained-Activation*).

Dimensions of individual differences

To identify the underlying dimensions of processing, non-linear functions were fit to the time course of fixations to each candidate (e.g., target, cohort) for each participant. These functions were based on prior work[19,61], and captured factors like the slope, crossover, and asymptote of target fixations (Fig. 2B), the height and timing of the peak, the slope to and from the peak, and the asymptotes for competitor fixations (Fig. 2C). Collectively, these curves accounted for 99.4% of the variance in target fixations (SD = .43%), and over 94% of the variance in competitor fixations (Cohort: M = 96.2%, SD = 2.2%; Rhyme: 93.9%, SD = 3.9%; unrelated: M = 94.4%, SD = 3.6%).

We then submitted the 13 parameters describing target, cohort, and unrelated fixations to a principal component analysis (PCA). This identified five Principal Components (PCs; Eigenvalues: 1.93, 1.67, 1.25, 1.14, 1.06) that collectively accounted for 80.6% of the variance (Supplementary Note 1). Monte Carlo analyses using a parallel PCA approach (Supplementary Note 2) validated that these specific results (both in terms of the amount of variance and the specific components that were identified) were highly unlikely to be achieved by chance.

Our subsequent analyses focused on the first three components (62.1% of the variance) for three reasons. First, as we describe next, these three clearly mapped onto pre-existing theoretical constructs (e.g., Fig. 1), while the remaining two did not reflect any pattern in the literature (Supplementary Note 3, Section 3.1). Second, the fourth component had only a small relationship with audiological and demographic factors, and the fifth component was not related to them at all. Third, neither factor predicted outcomes (Supplementary Note 3, Section 3.2). Finally, as described shortly, these components were clearly related to visual and oculomotor processes involved in the VWP, and likely did not reflect language (Supplementary Note 4).

Figure 3 shows reconstructed fixation curves for a hypothetical listener with lower or higher than average values (±1.5 SD) for the first three PCs (Supplementary Note 3 for the others). Each is scaled such that a low value on a given PC represents typical NH processing in ideal conditions and a higher value represents a CI user or a NH listener in challenging conditions.

These PCs (which we term lexical competition dimensions/indices) were clearly interpretable in terms of our hypotheses (see Supplementary Note 1 for a discussion). The first (28.6% of variance) reflected *Wait&See* (compare Figs. 3A to 1C/D). At high values of *Wait&See*, both the rise in target fixations and the cohort peak showed a fixed delay, and there was a *reduced* cohort peak. The second (21.5% of variance) reflected *Sustained-Activation* (Figs. 3B vs. 1B). At high values of *Sustained-Activation*, the overall rate of activation (target slope) was slower, and at asymptote there were *more* competitor fixations and fewer target fixations. The third (11.9% of variance), reflecting a *Slow-Activation-Rate*, showed a pattern we have observed in NH listeners over development[66,84] and aging[30] [16, for a review]. At high values, targets and onset competitors were activated and cohorts were suppressed more slowly.

The concordance between these PCs and existing theoretical proposals, as well as the fact that the PCs are orthogonal, supports three dimensions of lexical competition. These dimensions are continuous: people show differing degrees of *Wait&See* (etc).

Figure 4A illustrates this with the distribution of listeners across the first two PCs as a function of language status at the onset of deafness (*deafness onset*). While prelingually deaf CI users (red) show more *Wait&See* (they are right-shifted), some post-lingually deaf CI users (blue) appear in a similar region. Figure 4B shows the same participants grouped by functional acoustic hearing (FAH; Fig. 4B). The availability of FAH in *both ears* (green) reduces the variability along both dimensions, and these listeners are least likely to show high *Wait&See* or *Sustained-Activation*. Thus, a listener's unique profile in this multi-dimensional space of lexical competition may span both *Wait&See* and *Sustained-Activation* and be driven by many factors.

Are the same dimensions relevant for people without major hearing loss?

We next asked whether this processing space differs in listeners without major hearing loss. We examined a sample of age-typical hearing (ATH) listeners ($N = 107$, ages 11–78 years), tested with identical tasks as part of an independent study on lifespan aging[30]. We projected their VWP results onto the PCA defined for the CI users. Figure 5A suggests that ATH listeners show overall lower values on both *Wait&See* ($t(177.5) = 6.30$, $p < 0.0001$, $g = 0.88$, $CI_{95\%} = [1.01, 1.93]$) and *Sustained-Activation* ($t(177.0) = 8.7$, $p < 0.0001$, $g = 1.22$, $CI_{95\%} = [1.36, 2.16]$), and substantially less variance along both (*Wait&See:* $F_{Levene}(1200) = 10.6$, $p = .001$; *Sustained-Activation:* $F_{Levene}(1200) = 6.2$, $p = 0.013$).

In contrast, Fig. 5B suggests that ATH listeners and CI users do not show significantly different variation along the *Slow-Activation-Rate* dimension ($F_{Levene} < 1$), even as their mean is lower ($t(196.6) = 4.42$, $p < 0.001$, $g = 0.62$, $CI_{95\%} = [0.46\ 1.20]$). Variation on this dimension was significantly related to age and a quadratic effect of age: ($R^2 = 0.162$; Age: $B = 0.0232$, $t(196) = 3.57$, $p < 0.001$, $\beta = 0.302$, $CI_{95\%} = [0.010, 0.036]$; Age^2: $B = 0.0011$, $t(196) = 3.37$, $p = 0.001$, $\beta = 0.307$, $CI_{95\%} = [0.00046, 0.0018]$), and there was not a significant interaction

https://doi.org/10.1038/s41467-024-51514-3

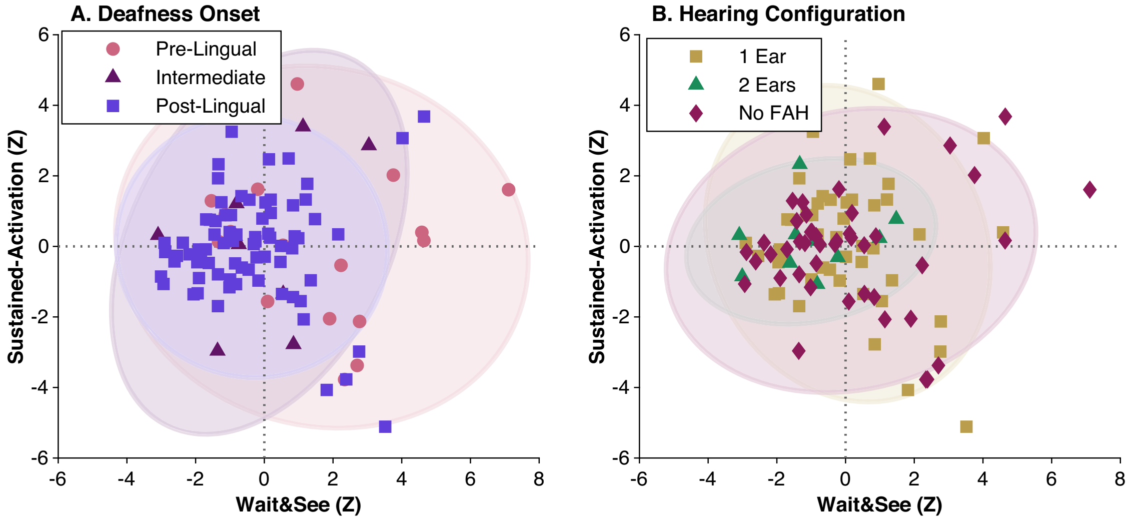

Fig. 4 | Distribution of individual participants across the *Wait&See* and *Sustained-Activation* dimensions. A Distribution of participants as a function of language ability. **B** As a function of Functional Acoustic Hearing (FAH). Each point is one participant. Shaded ellipses represent the 95% confidence interval around the mean of each group.

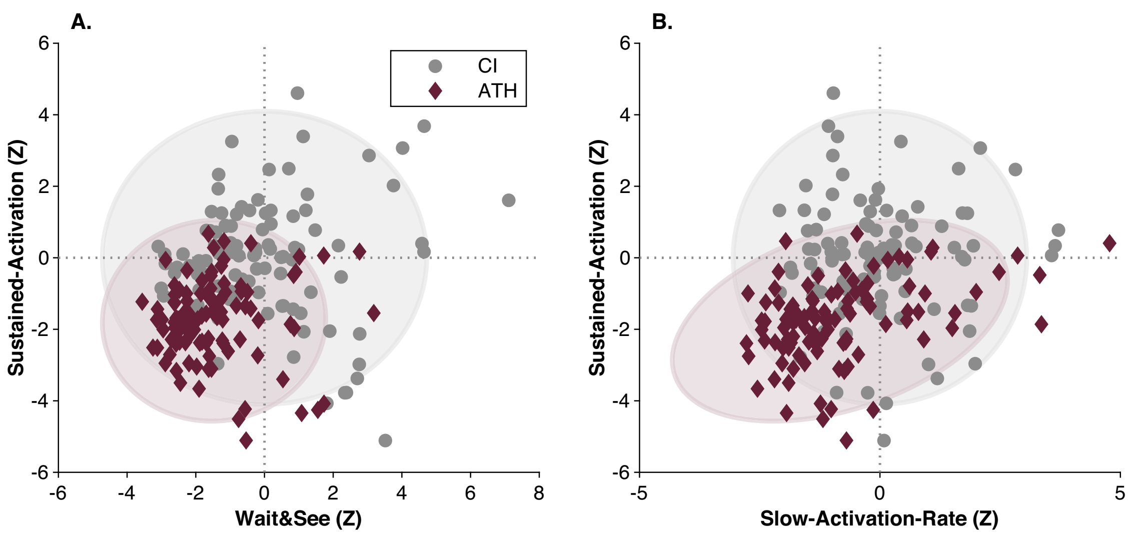

Fig. 5 | A projection of a sample of Age Typical Hearing (ATH) listeners from ref. 30 projected onto the processing space of the Cochlear Implant (CI) users. ATH listeners ranged in age from 11–80 and were run on the same experiment under identical conditions as part of a study on aging. **A** A space defined by the *Wait&See* and *Sustained-Activation* dimensions; **B** A space defined by *Slow -Activation-Rate* and *Sustained-Activation*. Ellipses represent 95% confidence intervals centered at a point defined by the mean on each dimension of the corresponding group.

with listener group (CI vs. NH) (Age × Group: B = −0.0080, t(196) = 0.61, p = 0.543, β = −0.052, $CI_{95\%}$ = [−0.033, 0.018]; Age^2 × Group: B = 0.00058, t(196) = 0.88, p = 0.380, β = 0.080, $CI_{95\%}$ = [−0.0007, 0.0019]). This pattern of results supports the idea that *Slow-Activation-Rate* largely reflects a more general age-related processing dimension, not hearing loss (though one that may be relevant for outcomes). Thus, CI users show substantially more variation in two dimensions of lexical competition, while the third may reflect natural age-related variation. We also conducted an independent PCA on the ATH listeners alone (Supplementary Note 5). Despite reduced variation, this again found three PCs that reflected *Wait&See, Sustained-Activation* and *Slow-Activation Rate*.

Predictors of lexical competition indices

We next asked what factors drive variability in these dimensions of lexical competition by using regressions to predict each listener's location along a lexical competition dimension as a function of demographic and peripheral auditory factors. Demographic factors included age and a quadratic effect of age (motivated by ref. 30), length of experience with a CI, biological sex, and deafness onset (pre- vs. post-lingually deaf). Peripheral auditory variables included the use of one or two CIs, the availability of Functional Acoustic Hearing [FAH] in one or both ears, pure tone average (PTA) in the lower frequencies (to capture FAH), and measures of spectral and temporal fidelity. We followed a model selection approach, so not all factors appear in every regression. Figure 6 summarizes the results (Supplementary Note 6, Table S7 for complete results). It reveals markedly different set of predictive factors for each dimension. Similar analyses on the 4th and 5th PCs in Supplementary Note 3, Section 3.1 do not show strong effects, suggesting these PCs reflect task-specific or visual-cognitive factors involved in the VWP, not lexical competition.

Wait&See was strongly linked to deafness onset (B = −2.380, t(95) = 4.30, p < 0.001, β = −0.479, $CI_{95\%}$ = [−3.466, −1.294]), which interacted with CI experience (B = 0.139, t(95) = 2.45, p = 0.016,

https://doi.org/10.1038/s41467-024-51514-3

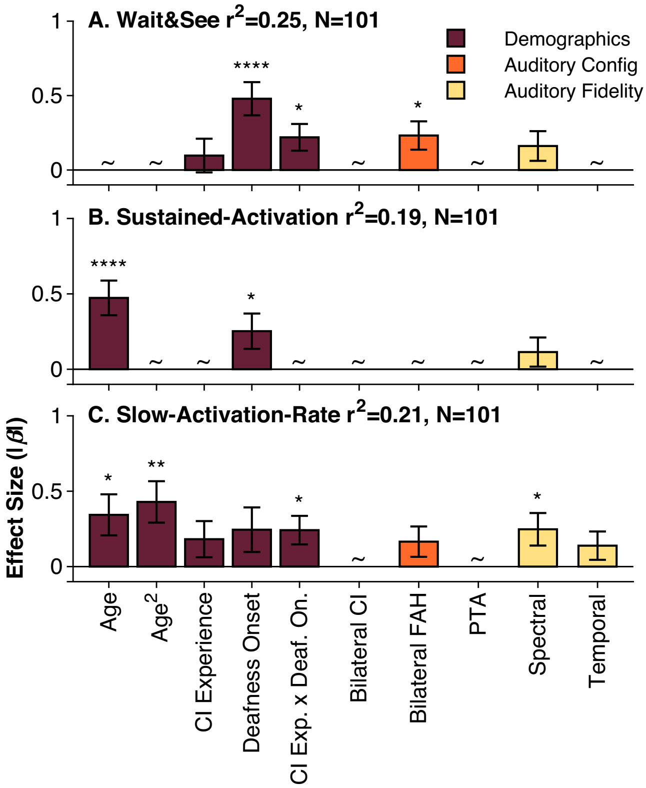

Fig. 6 | Effect size for each of the factors of the linear regressions predicting the first three principal components of lexical processing. The height of the bar represents the absolute size of the effect (the absolute value of the standardized regression coefficients); error bars represent standard error of the estimate or the uncertainty in estimating the effect size. *N* provided in each panel indicates the number of individual participants contributing to that regression. Lighter bars represent those same estimates in a separate model that did not include demographic factors. FAH Functional Acoustic Hearing, PTA Pure Tone Average. Significance is the statistical significance of each term in the regression (two-tailed, not corrected for multiple comparisons; full results in Supplementary Note 6, Table S8). $^*p < 0.05$, $^+p < 0.1$; ~: Effect did not survive model selection. **A** Effects on *Wait&See*. Exact *p*-values are CI Experience: $p = 0.39$; Deafness Onset: $p < 0.001$; CI Experience x Deafness Onset: $p = 0.016$; Bilateral FAH: $p = 0.017$; Spectral Fidelity: $p = 0.108$. **B** Effects on *Sustained-Activation*. Exact *p*-values are Age: $p < 0.001$; Deafness Onset: $p = 0.033$; Spectral Fidelity: $p = 0.241$. **C** Effects on *Slow-Activation-Rate*. Exact *p*-values are: Age: $p = 0.013$; Age^2: $p = 0.002$; CI Experience: $p = 0.134$; Deafness Onset: $p = 0.100$; CI Experience x Deafness Onset: $p = 0.012$; Bilateral FAH: $p = 0.102$; Spectral Fidelity: $p = 0.024$; Temporal Fidelity: $p = 0.142$.

$\beta = 0.220$, $CI_{95\%} = [0.027, 0.251]$; Fig. 7A). Pre-lingually deaf CI users with more experience exhibited less *Wait&See*, converging with post-lingually deaf CI users as they aged. Additionally, people with bilateral FAH exhibited less *Wait&See* ($B = -1.346$, $t(95) = 2.44$, $p = 0.017$, $\beta = -0.232$, $CI_{95\%} = [-2.446, -0.266]$; Fig. 4B).

In contrast, *Sustained-Activation* was almost entirely driven by a linear effect of age ($B = 0.052$, $t(95) = 4.11$, $p < 0.001$, $\beta = 0.473$, $CI_{95\%} = [0.027, 0.077]$; Fig. 7B), with a smaller influence of deafness onset ($B = -1.09$, $t(95) = 2.16$, $p = 0.033$, $\beta = -0.253$, $CI_{95\%} = [-2.080, -0.100]$). None of the other demographic or peripheral auditory variables significantly predicted *Sustained-Activation*.

Finally, for *Slow-Activation-Rate*, there was a significant interaction of device experience and deafness onset ($B = -0.242$, $t(95) = 2.55$, $p = 0.012$, $\beta = -0.242$, $CI_{95\%} = [-0.175, -0.023]$): Pre-lingually deaf CI users tended to speed up with device experience (overcoming their propensity to *Wait-and-See*; Fig. 7C); whereas, post-lingual CI users showed a small slowing. There was a large quadratic effect of age (Fig. 7C, Age: $B = 0.028$, $t(95) = 2.52$, $p = 0.013$, $\beta = 0.343$, $CI_{95\%} = [0.006, 0.050]$; Age^2: $B = 0.002$, $t(95) = 3.12$, $p = 0.002$, $\beta = 0.429$, $CI_{95\%} = [0.000, 0.004]$). This effect and the effect of age on Sustained-Activation (Fig. 7B) matches the results in Colby and McMurray[30] for ATH listeners. Specifically, *Sustained-Activation* matches what they termed "competitor resolution"[10], which declines with age. *Slow-Activation-Rate* matches their "timing" index, exhibiting a developmental profile with gains up to 30 years of age followed by a decline.

In only two cases did peripheral auditory factors predict any of the real-time lexical competition indices. First, spectral fidelity predicted *Slow-Activation-Rate* ($B = 0.016$, $t(95) = 2.29$, $p = 0.024$, $\beta = 0.247$, $CI_{95\%} = [0.002, 0.030]$)—less spectral fidelity predicted slower activation rate. Second, bilateral FAH predicted *Wait&See* ($B = -1.356$, $t(95) = 2.44$, $p = 0.017$, $\beta = -0.232$, $CI_{95\%} = [-2.446, -0.266]$)—people with bilateral FAH show less *Wait&See*. The smaller effects for peripheral auditory factors (as a whole) relative to other factors was surprising. One explanation was that effects of auditory function may have been masked by correlated demographic variables (e.g., poorer hearing with age). Indeed, comprehensive analyses in Supplementary Note 7 show moderate relationships among these variables. However, separate regressions containing only the auditory measures (the pale bars in the background in Fig. 6) showed only moderate effects. In these, spectral fidelity predicted *Wait&See* ($B = 0.02$, $t(95) = 2.05$, $p = 0.043$, $\beta = 0.225$, $CI_{95\%} = [0.001, 0.043]$), and was correlated with *Sustained-Activation* in a similar direction but was not significant ($B = 0.017$, $t(95) = 1.77$, $p = 0.080$, $\beta = 0.198$, $CI_{95\%} = [-0.002, 0.035]$). The specific profile of lexical competition shown by any listener is not robustly related to their auditory periphery.

Finally, we asked whether any of the five PCs may reflect more general processes (e.g., speed of processing) or factors like visual search or oculomotor performance that are relevant for the VWP (but not language). These skills were indexed with a non-linguistic analog of the VWP (nlVWP) in which participants matched a centrally presented shape to one of four competitors while eye-movements were monitored[85]. Supplementary Note 4 presents regressions relating indices from this task to each of the five PCs. These found no significant effects for the first three PCs, but significant effects for the fourth and fifth. This pattern of results provides a form of discriminant validity: there is clear statistical support that the first three reflect mechanisms relevant to lexical processing whereas there is little statistical support for the hypothesis that fourth and fifth are relevant for language and hearing.

The relationship of lexical competition to speech perception outcomes

Next, we asked if the dimensions of lexical competition predicted speech perception outcomes. Here, we considered three standard audiological measures: word recognition in quiet the Consonant Nucleus Coda [CNC] words[11], sentence recognition in +10 dB SNR AzBio Sentences[82], and a retrospective evaluation of listeners' real-world speech perception the Speech subscale of the Speech-Spatial-Qualities [SSQ] questionnaire[83].

Our prior analyses showed that the real-time lexical competition indices were moderately affected by demographic and auditory factors. This was also true for speech perception outcomes (Supplementary Note 7). Thus, we evaluated the relationship between real-time lexical competition and outcomes while controlling for these factors. We conducted the analysis in two stages: first identifying the optimal model for each outcome based solely on demographic and auditory variables, and then adding all three VWP indices.

Figure 8 shows effect sizes for each variable (Table S8 for numerical results). As in prior work[79,80], CNC word recognition was strongly related to spectral resolution ($B = -0.207$, $t(93) = 2.40$, $p = 0.002$, $\beta = -0.227$, $CI_{95\%} = [-0.376, -0.038]$) but not significantly related to temporal fidelity ($B = -0.942$, $t(93) = 1.79$, $p = 0.076$, $\beta = -0.166$, $CI_{95\%} = [-1.971, 0.087]$) with better fidelity predicting more

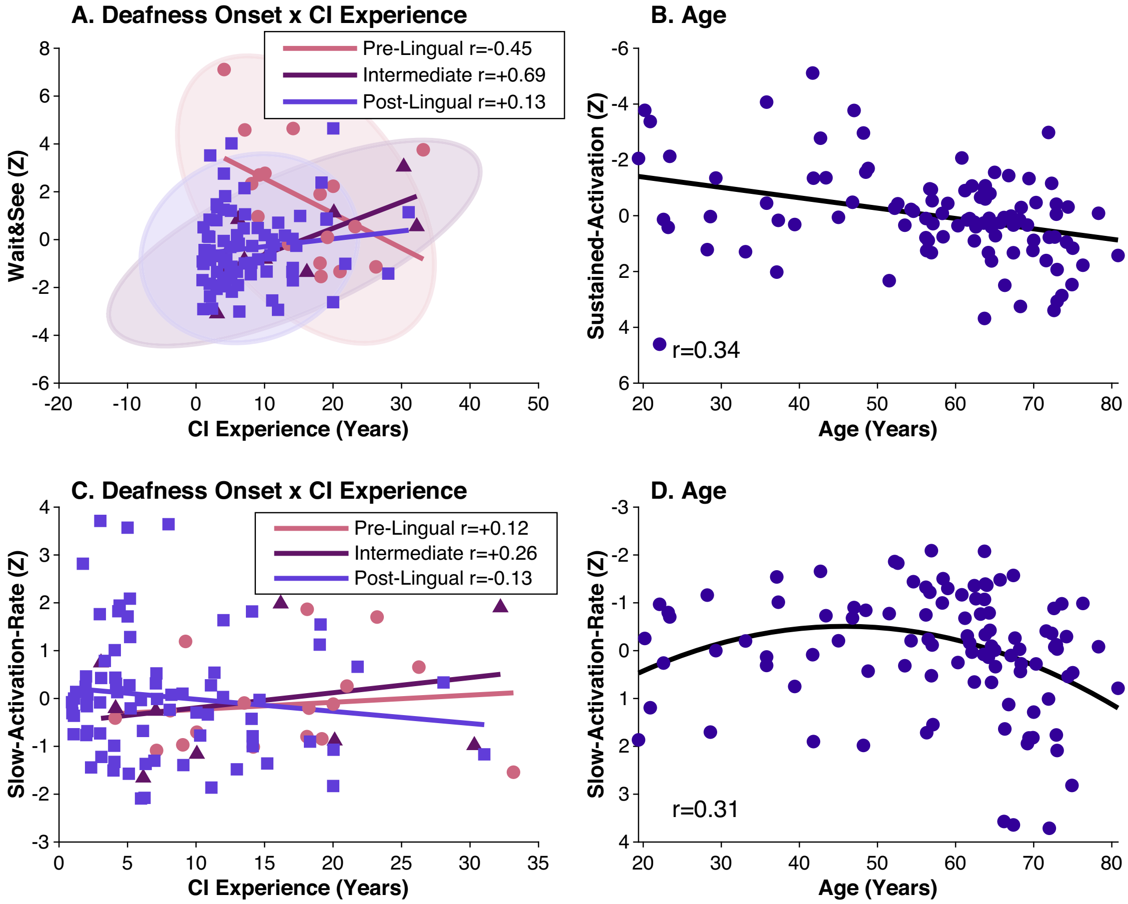

Fig. 7 | Effect of various factors on lexical processing indices. **A** Effect of Device experience and language status on the *Wait&See* dimension of lexical processing. Ellipses represent 1 SD from mean of each group. **B** Effect of Age on *Sustained-Activation* dimension. **C** Effect of device experience and language status on *Slow-Activation-Rate*. **D** Effect of age on *Slow-Activation-Rate*. *Y* axis in panels **B**–**C** is reversed so better performance is higher.

accurate performance. Critically, word recognition was significantly related to *Wait&See* ($B = -2.133$, $t(93) = 2.41$, $p = 0.018$, $\beta = -0.223$, $CI_{95\%} = [-3.832, -0.394]$), and *Sustained-Activation* ($B = -2.162$, $t(93) = 2.17$, $p = 0.033$, $\beta = -0.198$, $CI_{95\%} = [-4.120, -0.204]$), even after accounting for these factors. Sentence recognition in noise was not strongly related to spectral fidelity ($B = -0.239, t(55) = 2.00$, $p = 0.051$, $\beta = -0.233$, $CI_{95\%} = [-0.474, -0.004]$). However, it was significantly related to *Wait&See* ($B = -3.152$, $t(53) = 2.55$, $p = 0.014$, $\beta = -0.290$, $CI_{95\%} = [-5.575, -0.729]$) and *Sustained-Activation* ($B = -3.115$, $t(53) = 2.08$, $p = 0.042$, $\beta = -0.231$, $CI_{95\%} = [-6.047, -0.183]$). Finally, real world performance was predicted both by acoustic hearing (PTA: $B = 0.021$, $t(60) = 2.21$, $p = 0.031$, $\beta = 0.325$, $CI_{95\%} = [0.001, 0.041]$) and *Sustained-Activation* ($B = -0.284$, $t(60) = 2.25$, $p = 0.028$, $\beta = -0.271$, $CI_{95\%} = [-0.531, -0.037]$). Neither the fourth nor fifth PC predicted outcomes (Supplementary Note 3, Section 3.2).

We were concerned that controlling for so many factors in these analyses may underestimate the degree to which lexical competition predicted outcomes. Thus, we repeated the regressions with only the real-time lexical competition indices (Fig. 8, pale bars). These showed more widespread and robust effects. Word recognition accuracy was significantly predicted by *Wait&See* ($B = -2.68$, $t(95) = -3.06$, $p = 0.003$, $\beta = -0.28$, $CI_{95\%} = [-4.40, -0.97]$), *Sustained-Activation* ($B = -2.77$, $t(95) = -2.75$, $p = 0.007$, $\beta = -0.25$, $CI_{95\%} = [-4.74, -0.81]$) and *Slow-Activation-Rate* ($B = -2.74$, $t(95) = -2.02$, $p = 0.046$, $\beta = -0.19$, $CI_{95\%} = [-5.39, -0.10]$).

Sentence recognition was also negatively related to all three (*Wait&See*: $B = -4.07$ $t(57) = -3.38$, $p = 0.001$, $\beta = -0.37$, $CI_{95\%} = [-6.42, -1.72]$; *Sustained-Activation*: $B = -4.19$, $t(57) = -2.80$, $p = 0.007$, $\beta = -0.31$, $CI_{95\%} = [-7.12, -1.27]$; *Slow-Activation-Rate*: $B = -4.94$, $t(57) = -2.23$, $p = 0.029$, $\beta = -0.25$, $CI_{95\%} = [-9.25, -0.63]$).

Real-world outcomes continued to be predicted only by *Sustained-Activation* ($B = -0.30$, $t(63) = -2.43$, $p = 0.018$, $\beta = -0.28$, $CI_{95\%} = [-0.54, -0.06]$), but they did not have a significant relationship to *Wait&See* ($B = -0.11$, $t(63) = 1.14$, $p = 0.257$, $\beta = -0.13$, $CI_{95\%} = [-0.30, 0.08]$) or *Slow-Activation-Rate* ($B = -0.18$, $t(63) = 1.17$, $p = 0.245$, $\beta = -0.14$, $CI_{95\%} = [-0.47, 0.12]$). As a whole, these results suggest a robust relationship between indices of real-time lexical competition and multiple outcomes. Crucially, these effects are seen even when controlling for peripheral auditory and demographic factors.

Do Wait&See, Sustained-Activation and Slow-Activation-Rate benefit listeners?

Finally, we asked whether these differences in lexical competition are adaptive. When listeners wait and see, they accumulate more information before beginning lexical access. This could improve accuracy. Similarly, *Sustained-Activation* may help listeners maintain flexibility, keeping options open in case later information requires them to update a decision. These hypotheses predict that these indices will be positively correlated to outcomes: more waiting or sustaining leads to better speech perception.

This is not what was found. The regression coefficients (Supplementary Note 6, Table S9) suggest that in every case, less NH-like processing (higher values on the lexical competition indices) reflected *poorer* outcomes. One possibility is that the lexical competition indices reflect poorer auditory fidelity in addition to any functional benefits. That is, poor auditory fidelity leads to more *Waiting-and-Seeing*, but the benefits

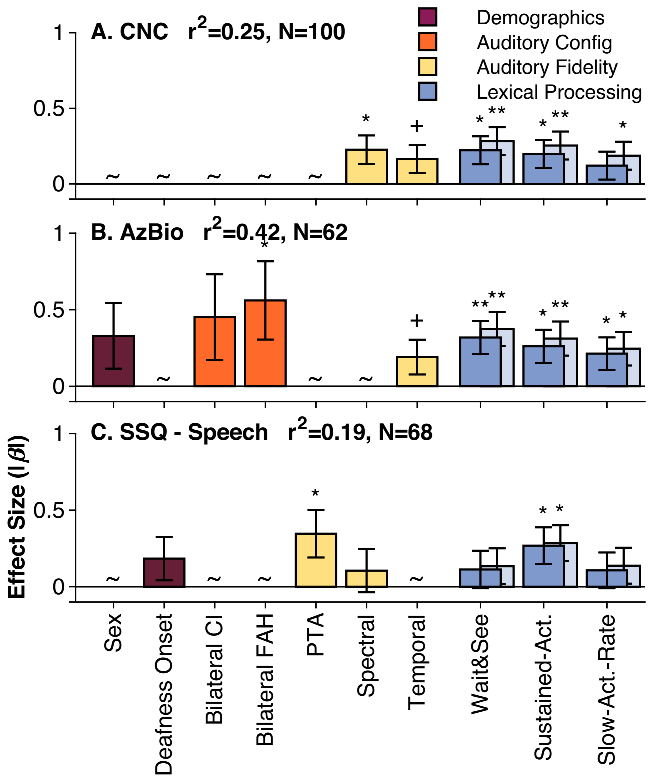

Fig. 8 | Effect size for each of the factors of the regressions predicting each outcome measure from the demographic, auditory factors and the first three principal components of lexical processing. The height of the bar represents the absolute size of the effect (the absolute value of the standardized regression coefficients); error bars represent standard error of the estimate or the uncertainty in estimating the effect size. Lighter bars represent those same estimates in a separate model that did not include demographic and auditory factors. PTA Pure Tone Average. FAH Functional Acoustic Hearing. *N* provided in each panel indicates the number of individual participants contributing to that regression. Significance is the statistical significance of each term in the regression (two-tailed, not corrected for multiple comparisons; full results in Supplementary Note 6, Table S9). *$p < 0.05$, $^+p < 0.1$; ~: Effect did not survive model selection. **A** Effect of demographic factors, auditory function and lexical processing indices on CNC word recognition. Exact *p*-values are Spectral Fidelity: $p = 0.018$; Temporal Fidelity: $p = 0.077$; *Wait&See:* $p = 0.018$; *Sustained-Activation:* $p = 0.033$; *Slow-Activation-Rate:* $p = 0.193$. **B** Same effects on Az-Bio Sentence Recognition. Exact *p* values are Sex: $p = 0.082$; Bilateral CI: $p = 0.400$; Spectral Fidelity: $p = 0.051$; *Wait&See*: $p = 0.014$; *Sustained-Activation:* $p = 0.042$; *Slow-Activation-Rate:* $p = 0.105$. **C** Same effects on real-world speech perception (SSQ). Exact *p*-values are Deafness Onset: $p = 0.220$; PTA: $p = 0.031$; Spectral Fidelity: $p = 0.487$, *Wait&See*: $p = 0.336$; *Sustained-Activation:* $p = 0.028$; *Slow-Activation-Rate:* $p = 0.372$.

of the lexical competition profile do not outweigh the costs of the poor fidelity. To address, we estimated the degree of *Wait&See* (or the other real-time lexical competition indices) relative to the quality of each listener's auditory periphery (the residuals of the lexical processing after regressing out the periphery). This analysis indexes whether each listener was more or less *Wait&See* (or *Sustained-Activation*) than expected given their hearing ability. People who exhibit more of a profile may exhibit better than expected speech perception. Even with this additional control, we still found robust negative relationships between lexical processing indices and outcomes (Fig. 9, Supplementary Note 6, Table S10). Thus, there is little evidence that these profiles are adaptive. They may instead represent three dimensions of challenged processing.

Discussion

This study identified three basic dimensions that underlie individual differences in real-time lexical competition. Each was related to distinct constellations of predictive factors, and each was predicted outcomes in CI users, even controlling for the nature and quality of the auditory periphery. The same dimensions—though with a reduced range—were also found in a separate sample of people without significant hearing loss. To see this, we related VWP indices of real-time lexical competition to traditional audiological measures of hearing and outcomes.

We note that these latter measures (e.g., CNC word recognition, Spectral Ripple, etc.) represent important variables in their own right, particularly given our large and diverse sample. Thus, Supplementary Note 7 describes a series of analyses on these factors alone. These analyses (briefly) showed that bilateral CI user appears to improve spectral fidelity as a form of redundancy gain; acoustic hearing can contribute to frequency separation but may make it more difficult to track temporal changes in the envelope; and that pre- and post-lingually deaf listeners have similar degrees of spectral resolution, but pre-lingually deaf individuals in this sample may have reduced temporal fidelity.

Before discussing the theoretical and clinical implications of the present study, we start by noting its limitations and scope. Word recognition encompasses many sub-systems such as speech perception and semantic processing; it typically occurs in sentence contexts which can further constrain it; and it may be affected by a variety of properties of the input and of the words (e.g., frequency, length). Nonetheless, our single-word indices were still related to outcomes in sentences and the real-world, validating the importance of this level of analysis. Additionally, while wordform recognition is multifaceted, our study focused only on one aspect of the process. This aspect was selected as it is the aspect that most mechanistic theories[14,86] have emphasized—competition among similar sounding words. In doing so, it illustrates the ways in which this slice of the problem can vary systematically and how this may relate to outcomes. Our conclusions should be narrowly construed in terms of variation in the way lexical competition is resolved, and we are not presuming that there is not any meaningful variation at other levels of the system, or in listeners' responses to other variables. Indeed, a clear extension of this work would be to conduct a similar individual differences approach but using new paradigms based on the VWP that tap other aspects of word recognition (e.g., refs. 63,87), or using other paradigms entirely[88,89]. In that way, our work has offers clear conceptual and statistical tools that may help identify the relevant dimensions of these other aspects of word recognition.

Turning to the most important VWP results, our PCA identified three key dimensions (Fig. 3), all of which were predicted by prior theoretical and empirical work[19,20,71]. The first was *Wait&See*, in which lexical access undergoes a fixed (and somewhat large) delay, which reduces cohort competition. The second showed a *Sustained-Activation* profile in which lexical activation builds slowly and competitors are not fully ruled out at asymptote. These two profiles have been observed in smaller-scale studies of CI users[19,20,71], in children with hearing aids[70], and in NH listeners experiencing challenging listening conditions[17,18,72]. The third reflected the overall rate of activation buildup and decay (*Slow-Activation-Rate*)/ This has been linked to both development[66,84] and aging[30]. These three dimensions were not strongly related to general visual/cognitive processing (Supplementary Note 4), suggesting they uniquely reflect word recognition, not ancillary processes that are engaged in the VWP (e.g., visual search, oculomotor planning).

Our large sample and individual differences approach allowed us to extend our understanding of these profiles beyond prior work. Considering lexical competition dimensions as outcomes, our study revealed three important findings. First, we demonstrated that *Wait&See* and *Sustained-Activation* are not two ends of a single dimension; they are independent of each other. Individual CI users (and listeners without major hearing loss) adopt a unique combination

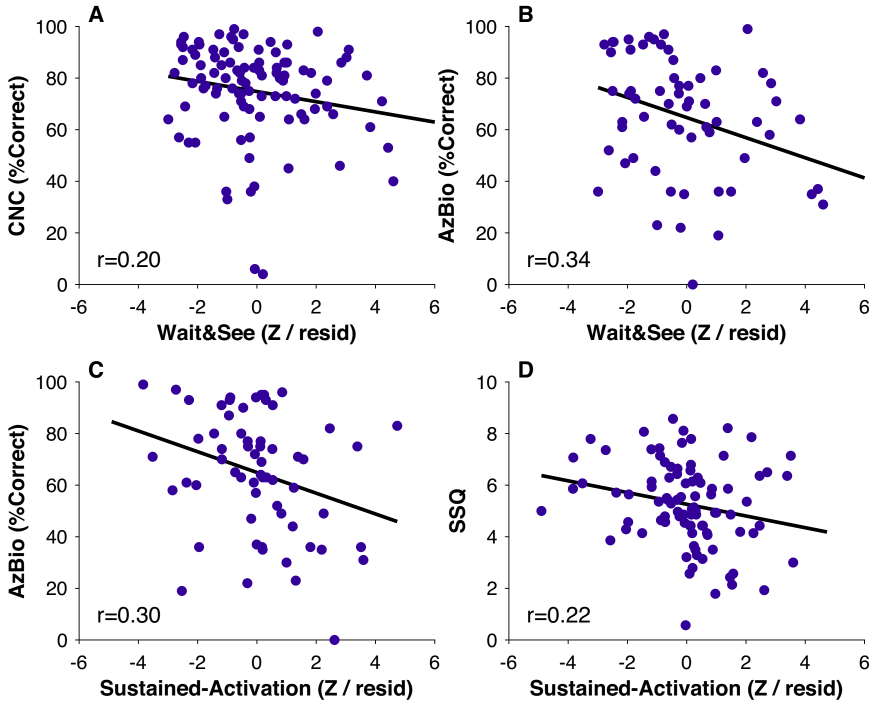

Fig. 9 | Relationship between speech perception outcomes and indices of real-time language processing which have been residualized against peripheral auditory factors. A Word recognition in quiet (CNC accuracy) as a function of *Wait&See*; **B** Sentence recognition in noise (AzBio) as a function of *Wait&See;* **C** Sentence recognition in noise (AzBio) as a function of *Sustained-Activation*; **D** Subjective real world speech perception (SSQ) as a function of *Sustained-Activation*.

of these continuous dimensions (Fig. 4). Second, while *Wait&See* was strongly associated with early deafness, it was also observed in a sizable number of post-lingually deaf individuals. It also accounted for the bulk of the variance in word recognition across the sample, even though our sample was heavily weighted toward post-lingually deaf CI users (N = 75/101). Thus, *Wait&See* was a substantial component underlying most CI users' performance. Third, a fair number of CI users appeared in the ATH portions of the space (Fig. 5A). This was not strongly predicted by peripheral auditory function or demographics, raising the possibility of factors that insulate listeners from adopting maladaptive processing profiles. One likely candidate is language skill prior to hearing loss[90,91]. Future work should identify these factors, particularly those that can be assessed prior to implantation to shape outcomes.

It was surprising that the auditory periphery was not a particularly strong predictor of the lexical competition indices relative to other factors (Fig. 6). Bilateral FAH was associated with *Wait&See* and spectral fidelity with *Slow-Activation-Rate*. However, non-auditory factors, like deafness onset (*Wait&See*, *Slow-Activation Rate)*, age (*Sustained-Activation, Slow-Activation Rate)*, and their interaction (*Wait&See*) had larger and more robust effects. One possibility is that our peripheral measures were not sufficiently sensitive. However, these same measures significantly related to outcomes (particularly when the VWP was not in the analysis) in the predicted ways (Supplementary Note 7). This supports the hypothesis that the dimensions of processing found by the PCA are unique cognitive differences that do not passively reflect the quality of the input. Ongoing work by our team is asking whether these differences derive from language abilities and cortical structure prior to hearing loss (and intervention), and/or from the distribution of listening experiences that people have (e.g., the diversity of talker voices they face every day, the amount of listening practice [device utilization] that they engage in).

Notably, *Wait&See* was much more strongly tied to deafness onset than hearing factors. This is not likely mediated by auditory fidelity: Supplementary Note 7 shows that pre- and post-lingually deaf listeners do not differ in spectral fidelity. This conclusion is further supported by Klein et al.[70] who showed that children with mild-to-moderate hearing loss (who have good auditory fidelity with their hearing aids) still showed *Wait&See*. Thus, hearing loss during early development is likely an important factor leading to *Wait&See*. However, *Wait&See* is not limited to pre-lingually deaf individuals: 29 of the 75 post-lingually deaf CI users showed a *Wait&See* index greater than 0 (and 5 of the 17 pre-linguals showed a value less than 0), as did 12 people without hearing loss. This implies additional unknown factors may lead listeners to wait and see.

We also found that both *Sustained-Activation* and *Slow-Activation-Rate* were strongly tied to aging. This was particularly apparent for *Slow-Activation-Rate*, where listeners without major hearing loss showed almost as much variation as CI users (Fig. 5B). This mirrors findings with ATH listeners across the lifespan[30], and makes an important point that the natural aging of lexical competition skills impacts CI users. However, in CI users, the normal slowdown in language skills with age could compound with *Wait&See*-induced delays or poorer resolution to make everyday language processing quite challenging. This raises the need for assessments of the *efficiency of language processing* as part of both standard neuropsychiatric and audiological care, even for normal hearing neurotypical adults c.f.[92].

A critical goal in this study was to determine if these dimensions of lexical competition related to outcomes, and if so, whether profiles like *Wait&See* are functional. The latter hypothesis was supported: both *Wait&See* and *Sustained-Activation-Rate* predicted outcomes even after accounting for the periphery and demographic differences, and this is underscored by their relative insensitivity to oculomotor or general cognitive differences. However, different factors appeared to be related to different outcomes. All three dimensions were important for the most complex measure (sentences in noise) (Fig. 8). This may reflect that perceiving sentences in noise demands both efficient processing (*Wait&See* and *Slow-Activation-Rate*) and the ability to fully suppress competitors (*Sustained-Activation*). In contrast, for real-world outcomes, individuals may be able to ask others to slow down or use context to fill in missing words. Consequently, efficiency is less important, but the ability to fully suppress competitors may be more important (*Sustained-Activation*).

However, in all cases, there was not strong evidence that these profiles were adaptive (Fig. 9). Listeners who showed more *Wait&See* or more *Sustained-Activation* generally had poorer outcomes—even accounting for their auditory fidelity. It appears that the people who do not *need* to adapt for accuracy do better at word recognition, and people who do not need to adapt for flexibility do better in sentences. One concern might be that our outcome measures do not require complex integration *across* sentences or a discourse. Thus, it is possible that these effects will change when related to more demanding outcomes.

Nevertheless, if we take these effects at face value, there remains an open question. If these profiles are not a simple product of auditory fidelity and they are not adaptive, then what drives listeners to these differences in the basic process of lexical access? First, we note that listeners without hearing loss also show considerable variation in this space (Fig. 5A). Perhaps hearing loss simply expands on whatever natural proclivity a listener has toward *Wait&See* or *Sustained-Activation*. One possibility in this regard, is that these profiles are akin to an allergy—an overreaction to a mild insult. That is, perhaps these profiles are "intended" to be adaptive (much like the immune response) but are too extreme to be beneficial. If so, this kind of over-compensation may relate to listeners' anxiety or meta-cognition about their language comprehension skills (e.g., about their ability to keep up or hear everything correctly), or to the diversity of talkers and language tasks they do every day (which may make them better "prepared" for the laboratory tasks).

Of note is that *Slow-Activation-Rate* did not predict outcomes. However, *Slow-Activation-Rate* was primarily associated with age, to which none of the outcome measures were strongly related (Fig. 8). Moreover, none of our outcome measures were timed. Therefore, *Slow-Activation-Rate* may be more relevant for more specific challenges like dealing with fast speech.

At the broadest level, this study illustrates the importance of real-time language processing for understanding the success and challenges in language comprehension amongst people with hearing impairment (as well as people undergoing typical aging). This is clear if we compare the amount of variance accounted for by models predicting outcomes from the auditory and demographic factors (Supplementary Note 7) to those models when we add the indices of lexical competition from the VWP (Fig. 8, Supplementary Note 6, Table S8). For all three outcomes, the models examining demographic and auditory factors alone showed small to moderate effect sizes. This reflects the persistent and difficult-to-explain variability in performance among CI users in both lab-based and real-world measures. However, the addition of the lexical competition indices led to large gains in the amount of variance that was accounted for (see Supplementary Note 6, Table S10 for numerical results, and Fig. S4 for a visualization). For example, regressions predicting CNC accounted for 16% of the variance using only the typical variables, but 25% when lexical competition indices were added. Similarly, 21% of the variance in sentence in noise recognition was due to demographic and audiological variables alone, but the regression accounted for 41% when the lexical indices were added. Finally, for the more difficult to predict SSQ (subjective real world speech perception), regressions accounted for 11% of the variance with only auditory and demographic variables, but 19% with the lexical competition indices. In each case, predictive power almost doubled by adding indices of lexical competition (with substantial shared variance). Therefore, real-time processing measures capture unique variance that cannot be attributed to the auditory periphery and may be uniquely important for outcomes.

For a listener to achieve successful speech perception with a CI, or in other challenging conditions, it is not sufficient to accurately encode the signal. People must be able to efficiently use whatever input they have to access meaning. This efficiency may be particularly important in situations where speech is fast, or when it is not part of an interactive conversation in which the partner can pace themselves to the needs of the listener (e.g., in a radio broadcast). Moreover, a lack of efficiency may require listeners to exert more effort, an important real-word issue as many people with hearing impairment report that language comprehension is fatiguing[93,94]. There has been considerable recent emphasis in audiology on more naturalistic texts such as sentence processing in noise. However, our results raise the need to consider other challenging listening conditions, particularly, the need to "keep up" with rapid speech in context. Similarly, this raises a need for assessments that stress efficient processing, not just accuracy.

There has been considerable emphasis in the recent literature on the link between hearing and cognitive decline[1,6,22,27,95,96]. Here, we see substantial variance in traditional hearing outcomes (e.g., CNC word recognition accuracy) that are uniquely linked to cognitive processes that are specific to language (and not to domain general visual or decision-making processes: Supplementary Note 4), but which are also not strongly related to auditory fidelity. That is, variation in language plays an independent role in speech perception outcomes that is at least as big as the auditory periphery. Notably, we only examined one aspect of word recognition – it is likely that we would see even more gains by considering additional aspects of word recognition, or by expanding this approach to sentence processing. Such work blurs the line between hearing and cognition: to the degree that social isolation and deprivation are critical factors in hastening decline, it may not be pure hearing ability that matters, as much as functional hearing—the combination of hearing and language. People may struggle to access language efficiently for either peripheral auditory reasons or due to difficulties in the cognitive processes of language. However, it is the fact that they are struggling to access language (and the resulting difficulty in social engagement) that matters more than whether this is specifically due to hearing loss. Neuropsychological approaches should consider language (particularly processes related to efficiency) as a key factor in social engagement for both cognitive decline and hearing loss. However, efficiency is not just another factor that can be retained or lost with age. Rather it is a potential mediator of the link between hearing loss and cognitive decline.

Theoretically, the present study demonstrates that the cognitive mechanisms underlying a key component of real-time lexical processing—dynamically unfolding competition—have lawful individual differences across listeners with, and without, hearing loss that can be characterized in a small number of meaningful dimensions. These dimensions were detected by combining the tools of individual differences (PCA) with the tools of cognitive science (e.g., the VWP). They are not just abstractions—they are theoretically meaningful and predicted by prior work, and they matter for real-world success. Moreover, these mechanistic differences may be amenable to training[97], raising the possibility of moving people to new regions of the processing space.

Cognitive science has traditionally sought to unpack basic mechanisms in modal listeners (normal hearing, monolingual, neurotypical adults), in part because we did not have the tools to characterize lawful differences in processing. This difficulty has only scaled with the advent of both models[64] and measures like the VWP that make precise predictions at a millisecond timescale. In the face of such complexity, it can be difficult to identify a few degrees of freedom to characterize variation beyond the modal listener. However, modal listeners are rare. Many people struggle with hearing loss, most people undergo development and aging, and multilingualism is the norm worldwide. Thus, theories of language processing must encompass not only the 'ideal' case, but also the underlying dimensions of variation in processing mechanisms. The critical issue facing the next generation of cognitive models is to identify lawful degrees of freedom by which basic mechanisms can vary to describe variation across people. This study pushes the field toward the use of tools like the VWP that are well-established and linked to basic mechanisms, but to use them in a way that characterizes the diversity of mechanisms, rather than assuming that any differences from the modal listener represent a deficit. Such an investigation points to a cognitive science that can be equitable for all people and not only the "modal listener", and one that captures variation in basic mechanisms across individuals and within individuals across contexts to facilitate more flexible language processing.

Methods

Participants

This study tested 114 CI users. All participants were monolingual English speakers with at least one year of CI experience, normal speech motor control, normal or corrected-to-normal vision, no history of developmental or neurological disorder, and at least some hearing loss in both ears (CI users with single side of deafness were excluded). Participants were categorized by self-reported biological sex, and our study design attempted to sample sufficient people from both sexes to permit it to be used as a factor in analysis. Thirteen people were tested but excluded from analysis for not meeting eligibility criteria ($N = 9$) or for not completing the VWP task ($N = 4$). This left 101 in the final analysis (45 male, Age: M = 57.4, SD = 15.26 years, range 80.8–19.4).

Participants were recruited from a large registry of CI users through the Iowa Cochlear Implant Clinical Research Center at the University of Iowa Hospital and Clinics. This was not part of a clinical trial, but rather was part of an ongoing clinical research project examining outcomes in a sample that was being treated for hearing loss. Participants ran this study on the same or following day as their routine audiological checkup and programming. Most patients had their devices tuned prior to testing. Participants with any acoustic hearing received a full audiogram in the clinic using a Grayson Stadler GSI Audiostar Pro audiometer with sounds presented over the included headphones.

The sample was highly variable across a number of factors (Table 1 main text; complete details in Supplementary Table S1). There was large variability in device configurations. There were unilateral and bilateral CI users, many of the CI users were implanted with hybrid CIs[76] that preserved acoustic hearing in the implanted ear(s) and some had usable acoustic hearing on a non-implanted ear (bimodal). We thus characterized listeners along three dimensions.

First, we documented whether the listener used one CI or two (Unilateral: $N = 79$, Bilateral: $N = 22$). Second, we assessed whether the listener had functional acoustic hearing (FAH). This was based on pure tone audiometry on the day of test, using the average of the low frequencies (250, 500, 1000, 1500 Hz) in the better ear as a continuous index of acoustic hearing. Listeners were classified as having FAH if they had better than 85 dB on a non-implanted ear, or better than 65 dB in the implanted ear.

The particular selection of frequencies does not represent the full range of useful acoustic hearing. It was motivated by the large number CI users who were implanted with hearing preservation (Hybrid) CIs that retain acoustic hearing in the implanted ear. For many of these listeners, cochlear implantation results in the loss of acoustic hearing above 1000 Hz; we included 1500 to catch the few who may have a little more hearing. Work with hybrid listeners shows that even at these low frequencies this acoustic hearing is helpful[98,99]. We retained this same threshold for the large number of bimodal listeners to ensure a common metric for evaluating functional acoustic hearing. Since most of these listeners have typical profiles of age-related high frequency hearing loss, this also captures the frequency range they are most likely to have access to via acoustic hearing.

By this standard, 57 listeners had FAH in one ear and 44 did not. Finally, we classified each listener as having functional acoustic hearing (FAH) in one or both ears: 10 CI users had FAH in both ears, 47 in one, and 44 had none.

Seventeen CI users were classified as having pre-lingual onset of deafness (profound deafness before age 5); 75 had clear post-lingual onset of deafness (after age 18); and nine were labeled as intermediate with deafness occurring in childhood (often progressive). With respect to audiological factors, age and gender this sample is representative of CI users. With respect to race, we had only a single non-white listener. Though this is not representative of the population as a whole, epidemiological work on CI users suggests that the population of CI users is heavily skewed toward white individuals[100,101].

All recruitment and experimental procedures were approved by the University of Iowa Institutional Review Board (IRB# 202210440), with separate protocols for the CI users tested here and the listeners without major hearing loss. Prior to implantation, CI users provided written informed consent, with the opportunity to ask questions and view the laboratory. CI users were compensated $50 for a half day of testing and $75 for a full day.

We also describe results from listeners without major hearing loss, with *age typical hearing* (ATH). This data came from a separate study[30]. This sample included 107 participants (39 male, 68 female) between the ages of 11–78 (Age: M = 47.8 years, SD = 19.5 years, range = 11.2–78.1 years). ATH listeners met the same criteria as CI users with the exception that they had normal hearing. All ATH participants received a full audiogram with a calibrated Grayson Stadler GS-61 audiometer, with sounds presented over Grayson Stadler DD-45 headphones. Participants were required to have a pure tone average of less than 30 dB in at least one ear (average calculated at 0.25, 0.5, 1, 2, 4, and 6 kHz). We relaxed the typical criteria for ATH because (1) it was necessary to obtain sufficient older adults; (2) if requested, participants were allowed to slightly adjust the volume for comfort; and (3) in the original report of that study[30], we found no relationship between minor between-subject variation in hearing and VWP performance. ATH listeners were tested under a separate IRB protocol which covered both minors and adults (IRB# 200902782). They provided written informed consent on the day of testing with an opportunity for questions if they had any. For minors, their parent or guardian signed a written informed consent document, and the participant underwent an additional verbal assent procedure with the experimenter.

Power. The size of the sample of CI users was determined by participant availability, not by an a priori prior analysis. Our plan was to test all CI users that were available for testing during a 3.5-year period (timed to the end of the grant). To understand power, we conducted sensitivity analyses based on this sample which computed the Minimum Detectable Effect (MDE) for variants of a linear regression. These assumed $\alpha = 0.05$, $1\text{-}\beta = 0.8$, and a two-tailed test. The MDE for a simple correlation given these assumptions was $|r| \geq 0.271$. The MDE for detecting a single significant effect in a regression with 5 parameters was $r^2 \geq 0.073$. Finally, the MDE for detecting a change in variance for a regression that started with 4 parameters (e.g., auditory fidelity and demographic factors) and added three more (the VWP indices) was $r^2 \geq 0.101$. Thus, this sample was sufficient for detecting small-to-medium effects.

General procedures

This study was conducted as part of a large clinical research project that included a large battery of experiments and standardized measures lasting about two hours. Only a subset of these tests is reported here (though this study reports all auditory fidelity measures and all speech perception outcomes that were available). For all participants, eye-tracking in the Visual World Paradigm (VWP) was conducted in a soundproof booth by the McMurray lab team. Other measures were collected by audiologists that were members of the Iowa Cochlear Implant Clinical Research Center (ICICRC). This was done in a separate soundproof booth in the Department of Otolaryngology. For the CI sample, audiometry was always conducted by the ICICRC, and spectral and temporal fidelity tests were also generally conducted by that group. For some CI users, scheduling constraints meant that auditory fidelity measures were collected in the McMurray Lab. For ATH listeners, audiometry and fidelity measures were all tested by the McMurray lab. The order of procedures was the same for each participant.

Auditory fidelity tasks

We assessed the fidelity of auditory encoding along both spectral and temporal dimensions. Note that these tests are usually conducted using restricted listening conditions. For example, spectral fidelity may be tested using a single CI and without any acoustic hearing to determine how well the CI separates frequencies. However, in this study, listeners were tested in their full everyday listening configuration (e.g., if they used two CIs and a HA in their day-to-day life, they were tested with these devices here). Thus, these measures reflect functional auditory fidelity, not the performance of a single listening device. Tests were designed to be conducted in the audiology clinic by the audiologists to support our multiple-lab center. Thus, they used presentation settings that were common to multiple labs, and thus may not match settings used for the VWP paradigms that were conducted by our team. Testing was conducted in a double-walled sound booth using a single loudspeaker located 1.5 m in front of the participant. Stimuli were played at 60 dB SPL intensity (measured with a handheld sound level meter in dBA weighting), which was fixed for all participants.

General procedures. Participants performed a 3-alternative forced choice oddball task in which they heard three stimuli and selected which differed. Sounds were 500 ms long with a 50 ms ramp at onset and offset and were generated uniquely for each trial by the control software. To deter listeners from using loudness as a cue to detect the oddball stimulus, root mean square values were first equalized among the three stimuli, and the presentation level was then roved randomly between −3 and +3 dB within the three sounds on a trial, and across all the stimuli. Stimuli were played with a 750 ms inter-stimulus interval. As each sound played, a numbered box appeared (1-3) on the screen. After hearing all three, the listener chose the oddball by clicking on a numbered box or typing the number.

This task was embedded in a Bayesian adaptive procedure using the Updated Maximum Likelihood (UML) algorithm[102]. In this procedure, the algorithm estimated a psychophysical function on each trial. It was updated after each response, and then used to select a stimulus for the next trial that will be most informative (given the current estimated function). The algorithm ran for a fixed number of 70 trials, and typically yields convergence faster and more reliably than traditional staircase procedures. The task was constrained by priors estimated from previous CI users' performance. We used this procedure to estimate a 3-parameter logistic function with a slope, threshold and guess-rate parameters. The threshold parameter was used as the estimate of performance for that dimension.

Each task started with four practice trials. These included feedback as to the accuracy of the response and did not contribute to the estimates. Testing trials did not include feedback. The entire procedure took approximately 7 min for each dimension. This was implemented with custom experimental control software developed in the MATLAB Psychophysics 3 Toolbox.

Spectral fidelity. Spectral fidelity has been strongly linked to speech perception accuracy in CI users[79,80] as it reflects the degree of separation between frequency bands. We assessed this with spectral ripples[79]. These were full-frequency stimuli that consisted of broad band noise whose spectrum contained peaks at specific frequencies (analogous to a vowel), evenly spaced on a log-frequency scale. We used a low ripple density (1.25 ripples / octave), which is more characteristic of human speech and does not lead to artifacts from the CI processor[103]. The UML procedure held the density of the ripples (ripples / octave) constant and manipulated the depth on each trial to determine the minimum depth at which frequencies could be separated. For each trial, the standard sounds were created with a random starting location for the spectral peak, and the oddball was created with an inverted phase.

Temporal modulation. While CI processing often loses spectral fine detail, it is thought to preserve differences in the amplitude envelope; thus CI users may have more access to temporal cues[103]. Stimuli for this task consisted of a 500 ms tone with five-component frequencies (at 1515, 2350, 3485, 5045, and 6990 Hz) whose amplitude was modulated at 20 Hz[104]. Stimuli to be discriminated differed by the presence of an amplitude modulation (either two modulated and one unmodulated sound, or one modulated and two unmodulated). The UML procedure manipulated the depth of the amplitude modulation.

Speech perception and outcomes

Speech perception outcomes were assessed by the audiological team in three ways. Testing was conducted in a double-walled sound booth using a sound field presentation. The loudspeaker was located 1.5 m in front of the participant.

Word recognition in quiet. Word recognition in quiet was assessed with the Consonant Nucleus Coda (CNC) words[11]. Participants heard mono-syllabic words from the loudspeaker at 60 dB SPL and repeated the word. A response was correct if it was repeated correctly in its entirety. Participants were tested on two lists, each with fifty words in a fixed order, and the average was recorded.

Sentence recognition in noise. For a more ecological outcome, we used AzBio sentences in noise[82]. Participants heard a semantically unpredictable sentence in multi-talker babble and repeated the entire sentence. Accuracy was based on the number of correctly repeated words and scored in real-time by the audiologist. Sentences were presented at 60 dB SPL and noise consisted of multi-talker babble at 50 dB SPL (for a + 10 dB SNR), presented through the same speaker as the target speech. Participants were tested on two lists out of the thirty-three available, each list containing twenty sentences in a fixed order.

Retrospective real-world speech perception. We also evaluated how listeners felt they were performing in the real world with a retrospective survey, the Speech-Spatial-Quality Scale (SSQ)[83]. The SSQ is a 49-item survey with items assessing speech perception, auditory localization and spatial processing, and overall sound quality. As our emphasis was on speech perception, we examined only the speech subscale.

Visual world paradigm

Design. The VWP task was modeled after prior work[60], including work with CI users[19,20]. Participants heard a spoken word (e.g., *rocket*) accompanied by pictures on a computer of the target, a cohort or onset competitor (*rocker*), a rhyme competitor (*pocket*) and a phonologically unrelated word (*bubble*). Items consisted of 60 sets of four words, each set containing a target, cohort (onset competitor), rhyme, and an unrelated competitor. Each word was easily picturable and piloted beforehand to ensure that they were readily understood. There were 30 monosyllabic sets and 30 bisyllabic sets (Supplementary Table S2).

Sets were developed over a series of pilot studies intended to build a canonical VWP task. We started with 120 sets which were developed and piloted with 68 NH young adults. We then selected the 60 item-sets 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: $r = 0.54$).

Each item in a set was used as the auditory target once. To discourage participants from adopting a process of elimination strategy (e.g., "I heard *rocker* on the last trial, so this word must be *rocket*"), one item from each set was randomly selected to serve as the target word on an additional trial. This led to a total of 300 trials (60 sets x 4 targets/set x 1.25 repetitions/set). Image placement was pseudo-randomized across trials and participants, such that each image type was equally likely to appear in any quadrant of the computer screen.

Given the structure of these item sets, not all types of competitors were present on any trial. For example, when *rocket* was the target there would be a cohort *(rocker*) and rhyme (*pocket*), what was termed a TCR trial. However, when *rocker* was the target there was only a cohort (*rocket*), as *pocket* was now mostly unrelated (a TC trial); and when *pocket* was the target there was a rhyme (*rocket*) but no cohort (a TR trial). When computing fixations to the cohort or rhyme only the relevant trial types were included. This effectively counterbalances any frequency difference between items (*rocket* serves as both a target and cohort). Looks to the target were averaged across all three types of trials which had competitors.

Procedure. The experiment was presented on a computer with a 17″ (5:4) monitor operating at 1280 × 1024 resolution and a standard keyboard and mouse. Audio signals were played on a SoundBlaster soundcard on the PC at a sample rate of 44,100 Hz, low-pass filtered at 6.5 kHz, and subsequently fed to a Samson C-que8 headphone amplifier and then two Boston Acoustics speakers in the soundproof booth. The loudspeakers were approximately one meter from the participant. Their volume was set to achieve 60 dB SPL for the recordings being tested, calibrated with a handheld sound level meter (dBA weighting) held at approximately the location of the participant's head. Participants were tested with whatever hearing devices they normally used (their CI(s) plus any hearing aids).

The experiment was run using Experiment Builder (Version 2.4.193, SR Research, Oakville, ON, Canada). Participants first placed their chin on a padded chinrest at the end of the testing table (55 cm from the monitor) and the experimenter adjusted its height to a comfortable position. The eye-tracker was then calibrated using a standard nine-point calibration.

Next, participants began the experimental phase. 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. This pre-scan period was intended to familiarize the participants with the locations of pictures to minimize the role of visual search on fixations after the target word was presented[105]. Pictures were 300×300 pixels, separated by 580 pixels horizontally and 324 pixels vertically. After 500 ms, the circle changed from blue to red, indicating that the participants could use the mouse to click on the dot and play the auditory stimulus. After hearing the target word, participants clicked on the picture that best represented the auditory word.

Every 30 trials a drift correction was performed to update the calibration for any drift of the eyes/head during the experiment. If the participant failed a drift correction, the eye tracker was recalibrated. The experiment lasted about 25 minutes and participants were permitted to take a break at any drift correction.

Stimuli. Auditory stimuli were recorded from a monolingual female English talker who spoke with the American Midwest dialect in a natural cadence. Words were recorded in a sound-attenuated room with a M-Audio 2×2 External Audio Interface with a head mounted microphone at a sampling rate of 44.1 kHz. For each word, the talker produced four to five exemplars both in isolation and in a neutral carrier sentence (*He said*…), to ensure a more uniform prosody across exemplars. We then selected the best exemplar (i.e., the one that had a falling prosody and the fewest auditory artifacts, such as clicks, creaky voice, etc.). This was excised from the sentences for use in the experiment. These tokens were then edited to reduce noise and remove any remaining artifacts, and 100 ms of silence was appended to the onset of each stimulus. Stimuli were amplitude normalized in Praat to 70 dB. The average duration of the experimental words was 710 ms (not including the silent period).

Visual stimuli consisted of 240 pictures constructed using a standard lab protocol[61]. For each word, 5–10 pictures were downloaded from a commercial clipart database. These were reviewed by a focus group of graduate students, undergraduate students, and research staff which selected the image that was most prototypical for that stimulus and recommended any modifications. Pictures were then edited for consistency with other visual stimuli, to ensure prototypical color and/or orientation or to eliminate unnecessary features, or for cultural considerations (e.g., reducing stereotype, ensuring a representative mix of genders and races in pictures depicting humans). The final images were approved (or sent back for modification) by an independent lab member with extensive VWP experience.

Eye-tracking. Eye movements were recorded with an SR Research Eyelink 1000 desktop mounted eye tracker. Both eyes were tracked if possible, and the better eye was selected after the fact for analysis. Pupil and corneal reflection were sampled at 1000 Hz to determine point-of-gaze.

Pre-processing of the fixation data was done using EyelinkAnalysis (version 4.211)[106]. This works on the basis of "events" (saccades, fixations and blinks), which grounds the analysis in more physiologically realistic data than working with 4 ms samples[107]. Fixations, saccades, and blinks were identified by the Eyelink control software using the default "cognitive" parameters set. Adjacent saccades and fixations were combined into a single look, which began at the onset of the saccade and terminated at the end of the fixation. Looks were assigned to one of four regions of interest, where regions were defined as the 300×300 area of the image, extended by 100 pixels to account for any noise in the estimation of gaze position. This did not result in any overlap between areas of interest. Looks were time-locked to the onset of the auditory stimulus. These looks were the basis of analysis. Looks launched before the onset of the target word (accounting for a 200 ms oculomotor delay) were ignored.

Accuracy of the final mouse click was generally high (M = 92.3%, SD = 7.5%, Range = 48.0–99.7%). Only trials where the correct target image was selected were included in further analyses, since the analyses sought to identify differences in the processes by which CI users arrived at the correct word.

Non-linguistic VWP (nlVWP)

We used a modified—visual-only—variant of the VWP (the non-linguistic VWP or nlVWP) to assess general speed of processing, as well as visual/cognitive factors (like eye-movement control and visual search). This was available for 91 participants. Participants saw a target shape (e.g., a *maroon hourglass*) in the center of the screen accompanied by four potentially matching shapes in the corners. Their task was to click on the shape that matched the target (see Supplementary Note 4 for an example). One of the shapes was a direct match, another matched its color but not shape (a *maroon chevron*), and the other two were unrelated. As in the standard VWP, eye-movements were monitored to yield a real-time index of the decision process, but without any contribution from language processing. This was originally developed by ref. 85, but we modified it here to use less nameable colors and shapes, to use color contrasts that were less susceptible to some color-blindness, and to shorten it for use with larger clinical sample.

Design. The nlVWP used 16 sets of four colored shapes. Shapes and colors were designed to be difficult to name (e.g., burgundy or lavender instead of red and purple; shapes like a chevron rather than a square). Each set consisted of two pairs of items that matched in color (e.g., a *burgundy arrow* and a *burgundy moon* paired with a *lavender parallelogram* and a *lavender trefoil*). Consequently, when one of the burgundy objects was the target, the two lavender objects served as unrelated foils. Each item in each set was presented as the target three times for a total of 190 trials.

Procedure. Each trial started with a preview, in which each of the four objects appeared in their respective corners. This was accompanied by a small blue dot at screen center. After 500 ms, the dot turned red, and the participant clicked on it. After 100 ms delay, the target stimulus was shown for 100 ms before it disappeared. The participant then clicked on the matching object to advance to the next trial. Eye-movements were recorded and pre-processed using identical procedures to the VWP experiments.

Statistical methods

Analysis of fixations: VWP. Analysis started from curves similar to Fig. 1. For each participant we computed the proportion of trials on which the participant was fixating the target, cohort and unrelated items at each 4 ms time slice. Fixation curves were generated from trials which contained the relevant competitor types. For example, target fixations were based on TCRU, TC, and TR trials while cohort fixations were based on TCR and TC trials. Rhymes were not included in this analysis as with 101 participants, the PCA could not reasonably accommodate the additional 4 parameters. Additionally, rhymes often receive few fixations[108], making them less suitable to index of general competition than cohorts. Moreover, in this sample, rhymes patterned with cohorts (e.g., a person with higher rhyme fixations also had higher cohort fixations), suggesting rhymes captured similar variance.

The fixation curves (e.g., Fig. 1A) were then characterized by fitting non-linear functions to them. The parameters of this function were then used to compactly describe an individual participant's data in terms of properties like the slope or asymptote (Fig. 2B, C). Functions were based on prior work[61,109]. Targets used a four-parameter logistic with parameters for the lower and upper asymptotes, the crossover (when in time the function transitioned between asymptotes) and the slope (the derivative at the crossover). Competitors and unrelated objects used a six-parameter asymmetric Gaussian with parameters for the initial and final asymptotes, the height of the peak, the location of the peak, and the onset and offset slopes. Curves were fit using a constrained gradient descent method that minimized the RMS error between the data and the function while obeying constraints to keep the function within reasonable bounds[110, version 30.0]. Fits were conducted separately for each participant. Fits were evaluated by hand and refit with new starting parameters if needed.

The parameters were then submitted to a Principal Components Analysis (PCA) to identify a smaller number of dimensions. To avoid overfitting the PCA, we dropped the onset asymptotes for the target and competitors, and only examined target, cohort, and unrelated fixations (rhymes tended to pattern with cohorts), leaving 13 parameters. Data were z-scored prior to the PCA. PCAs used the prcomp() function in R (version 4.2.2 2022-10-31 ucrt) and were conducted without rotation, as we embraced potential cross-loading of the factors as theoretically meaningful. PCs were visually inspected and scaled (by multiplying loadings by −1) such that a high value of that PCA meant a more CI-like pattern (e.g., more waiting), and a low value meant more NH-like. We retained five PCs. To compute each participant's score on these PCs, we used the get_pca_ind() function of the factoextra library in R (version 1.0.7).

To construct visualizations such as Fig. 3 (main text), we multiplied the eigenvectors (the loadings) by +/− 1.5 to create low and high values for each parameter (a difference of 3 SD between high and low). We then undid the Z transformation by multiplying the estimated parameters by the observed SD of that parameter and adding the observed mean, to compute each parameter under a low or high value of that PCA. These were then used to construct predicted timecourse functions.

Analysis of Fixations: nlVWP. Analysis started by fitting non-linear functions to the fixations to the target and color-matching competitor over time, using the logistic and asymmetric Gaussian respectively (similar to the linguistic VWP). Following prior work[111], these parameters were combined into two indices. *Target Timing* reflects the speed of fixating the target (the slope and crossover). Slope was log-scaled, and z-scored. Crossover was z-scored and multiplied by −1 (since an earlier crossover predicts a higher slope). These were then averaged. *Resolution* reflects the ultimate separation between the target and competitors. It was the maximum of the target minus the offset asymptote of the color competitor.

Missing data. A critical goal of our project was to relate VWP indices to outcomes even after accounting for peripheral auditory function. However, auditory fidelity was evaluated as a part of clinical care and was occasionally missing for some participants. We opted to fill in these missing values for two reasons. First, our goal was to examine the heterogeneity among CI users and even a small number of excluded subjects could eliminate valuable subsets (younger or older, pre- vs. post-lingually deaf, different hearing configurations). Second, our model-space regression approach did not presume any one auditory variable was critical, and even if temporal fidelity was missing (for example) any given participant would have many other variables of potential interest (e.g., FAH, bilateral hearing, etc.).

Missing data were filled in according to the following procedure. First, in some cases, audiograms were missing because a CI user was documented by the audiologist to exhibit profound deafness and did not conduct an audiogram. For these, thresholds on the implanted ear were replaced by 115 dB. We were all missing measures of Spectral Fidelity for 9 participants, and Temporal Fidelity for 11 (8 were missing both). These were presumed to be missing completely at random as they were missing for things like time constraints or technical errors. Missing completely at random was verified with a Little test[112]: $\chi^2(20) = 27.3$, $p = 0.126$). Thus, we used imputation (using the mice package in R, version 3.16.0) to compute these values. Scores were imputed from Age, Biological Sex, CI experience, PTA, and bilateral CI use with 200 imputations. This became our final dataset which is posted in the OSF page associated with this project. For analyses in which spectral or temporal fidelity was the outcome measure (Supplementary Note 7), we did not use imputed values, but excluded participants who were missing these measures.

General statistical approach. All significance tests assumed two-tailed distributions. When *T*-tests were employed, Levene's test of equality of variance was conducted, and *T*-tests used Welch's corrections for unequal variance. *T*-tests are reported with Hedge's *g* as a measure of effect size, confidence intervals represent 95% confidence interval of the mean differences. For regressions, effect sizes are reported as r^2 for the overall regression, r^2_Δ for commonality analyses, and standardized regression coefficients (β) for individual effects. Confidence intervals always reflect the 95% interval around the estimate of the unstandardized coefficient (B).

Regression approach. Results were analyzed as a series of regressions in R using lm() with the jtools package (version 2.2.2) to compute standardized regression coefficients. The distribution of the residuals and linearity of the relationships were assessed by examining scatterplots of the critical analyses. Collinearity was assessed by examining the correlations among independent variables and handled by the model space approach described below. We avoided overfitting the data by using the model-space search to limit models to in general 6-7 predictors.

Our first series of regressions asked what factors shaped each VWP component. We were interested in 6 demographic factors: biological sex (Male = 0.5, Female = −0.5), CI experience in years (centered), deafness onset (prelingual = −0.5, intermediate = 0, postlingual = +0.5), age (centered), and age^2 since our prior work on typical aging found a strong quadratic trend[30]. We also included a CI experience × Deafness onset interaction as we expected device experience to play different roles for pre- and post-lingually deaf listeners. There were three peripheral auditory factors (all centered)—spectral fidelity, temporal fidelity, acoustic hearing (better ear PTA)—as well as device configuration which was characterized with two additional factors: the use of two CIs (bilateral = +0.5, unilateral = −0.5), and the presence of functional acoustic hearing across ears (one ear FAH = −0.5, no FAH = 0, two ears = +0.5).

We did not have sufficient data to include all 11 possible factors in the model (not to mention the real-time lexical competition indices that would be needed for some analyses). Thus, to avoid overfitting the regressions, we conducted a full search of the possible permutations using the lmSubsets toolbox (version 0.5-2). This search was constrained such that (a) if the quadratic effect of age was in the model, the linear effect must also be included; and (b) if the device experience × language interaction was present, both main effects must be included; and (c) all models must include at least one of the peripheral auditory factors (since our primary research question was whether there was an effect of real-time lexical competition over and above auditory factors). We then used the Akaike Information Criteria (which penalizes model fit based on the number of degrees of freedom) to select the most parsimonious model.

Our second set of regressions then used speech perception outcomes (CNC, AzBio and SSQ) as the dependent measure, and added the real-time lexical competition indices derived from the PCA. These used a similar model selection approach. First, for each outcome measure, we found the most parsimonious model based on demographic and audiological factors alone. Next, we added all three VWP indices to the model.

Finally, we conducted commonality analyses using the yhat library in R (version 2.0-4). Complete results of this analysis are shown in Supplementary Note 6, Table S10, Fig. S4. This conducts a series of models with and without each combination of factors and uses r^2_{change} metrics to determine how much variance is uniquely attributable to each factor and how much variance is shared between subsets. We started by computing this separately for each variable and combination, but then subsequently added up the variance to compile the variance that was uniquely due to a class of factors. For example, to obtain the unique variance due to the auditory periphery, we summed the r^2 of any individual auditory variable and any shared variance that was shared between only auditory variables. Similarly, shared variance was pooled as either reflecting only shared variance among demographic or auditory variables (no lexical competition variables) or shared with lexical competition.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

This manuscript uses both newly collected data and existing data from ref. 30. The data collected for this study (the sample of CI users) is available at the Open Science Framework, accession code https://doi.org/10.17605/OSF.IO/K32FT (https://osf.io/k32ft/). This dataset includes all of the individual curvefits and all participant-level data. The raw eye-tracking data is too large to be conveniently shared. It is available by request to the first author. Existing data (the sample of ATH listeners) is available at the Open Science Framework, accession code https://doi.org/10.17605/OSF.IO/ZTHBW (https://osf.io/zthbw/).

Code availability

Code is available at the Open Science Framework, https://doi.org/10.17605/OSF.IO/K32FT (https://osf.io/k32ft/. This includes analysis code and code for generating all the figures. In addition, we provide scripts for the temporal and spectral fidelity tasks on the Open Science Framework https://doi.org/10.17605/OSF.IO/MC4FN (https://osf.io/mc4fn/). Eye-tracking processing and done with a separate utility available at the Open Science Framework (https://doi.org/10.17605/OSF.IO/C35TG https://osf.io/c35tg/). Curvefitting was done using publicly available software (https://doi.org/10.17605/OSF.IO/4ATGV https://osf.io/4atgv/).

References

1. Lin, F. R. Hearing loss and cognition among older adults in the United States. *J. Gerontol. A* **66A**, 1131–1136 (2011).
2. Lin, F. R. et al. Hearing loss and incident dementia. *Arch. Neurol.* **68**, 214–220 (2011).
3. Amieva, H., Ouvrard, C., Meillon, C., Rullier, L. & Dartigues, J. F. Death, depression, disability, and dementia associated with self-reported hearing problems: a 25-year study. *J. Gerontol. A Biol. Sci. Med Sci.* **73**, 1383–1389 (2018).
4. Heywood, R. et al. Hearing loss and risk of mild cognitive impairment and dementia: findings from the Singapore longitudinal ageing study. *Dement. Geriatr. Cogn. Disord.* **43**, 259–268 (2017).
5. Huang, A. R., Jiang, K., Lin, F. R., Deal, J. A. & Reed, N. S. Hearing loss and dementia prevalence in older adults in the US. *JAMA* **329**, 171–173 (2023).
6. Yeo, B. S. Y. et al. Association of hearing aids and cochlear implants with cognitive decline and dementia: a systematic review and meta-analysis. *JAMA Neurol.* https://doi.org/10.1001/jamaneurol.2022.4427 (2022).
7. Oh, S.-h et al. Speech perception after cochlear implantation over a 4-year time period. *Acta oto-laryngologica* **123**, 148–153 (2003).
8. Hamzavi, J., Baumgartner, W.-d, Pok, S. M., Franz, P. & Gstoettner, W. Variables affecting speech perception in postlingually deaf adults following cochlear implantation. *Acta Otolaryngol.* **123**, 493–498 (2003).
9. Nation, K. Lexical learning and lexical processing in children with developmental language impairments. *Philos. Trans. R. Soc. B Biol. Sci.* **369**, 20120387 (2014).
10. McMurray, B., Apfelbaum, K. S. & Tomblin, J. B. The slow development of real-time processing: spoken Word Recognition as a crucible for new about thinking about language acquisition and disorders. *Curr. Dir. Psychol. Sci*. https://psyarxiv.com/uebfc/. https://doi.org/10.1177/09637214221078325 (2022).
11. Peterson, G. E. & Lehiste, I. Revised CNC lists for auditory tests. *J. Speech Hear. Disord.* **27**, 62–70 (1962).
12. Dunn, L. M. & Dunn, L. M. *Examiner's manual for the PPVT-III Peabody Picture Vocabulary Test*. 3rd edn (American Guidance Service, 1997).
13. Dahan, D. & Magnuson, J. S. in *Handbook of Psycholinguistics* (eds M. J. Traxler & M. A. Gernsbacher) 249–283 (Academic Press, 2006).
14. Weber, A. & Scharenborg, O. Models of spoken-word recognition. *Wiley Interdiscip. Rev. Cogn. Sci.* **3**, 387–401 (2012).
15. Marslen-Wilson, W. D. Functional parallelism in spoken word recognition. *Cognition* **25**, 71–102 (1987).

16. McMurray, B., Apfelbaum, K. S., Colby, S. & Tomblin, J. B. Understanding language processing in variable populations on their own terms: towards a functionalist psycholinguistics of individual differences, development and disorders. *Appl. Psycholinguist*. **44**, 565–592, https://psyarxiv.com/zp564aw/ (2023).
17. Brouwer, S. & Bradlow, A. R. The temporal dynamics of spoken word recognition in adverse listening conditions. *J. Psycholinguist. Res.* 1–10 (2015).
18. Hendrickson, K., Spinelli, J. & Walker, E. Cognitive processes underlying spoken word recognition during soft speech. *Cognition* **198**, 104196 (2020).
19. Farris-Trimble, A., McMurray, B., Cigrand, N. & Tomblin, J. B. The process of spoken word recognition in the face of signal degradation: Cochlear implant users and normal-hearing listeners. *J. Exp. Psychol. Hum. Percept. Perform.* **40**, 308–327 (2014).
20. McMurray, B., Farris-Trimble, A. & Rigler, H. Waiting for lexical access: cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition* **169**, 147–164 (2017).
21. Lin, F. R. & Albert, M. Hearing loss and dementia—who is listening? *Aging Ment. Health* **18**, 671–673 (2014).
22. Griffiths, T. D. et al. How can hearing loss cause dementia? *Neuron* **3**, 401–412 (2020).
23. Wu, Y. H. & Bentler, R. A. Do older adults have social lifestyles that place fewer demands on hearing? *J. Am. Acad. Audio.* **23**, 697–711 (2012).
24. Armstrong, N. M. et al. Association of Midlife Hearing Impairment With Late-Life Temporal Lobe Volume Loss. *JAMA Otolaryngol. Head Neck Surg*. https://doi.org/10.1001/jamaoto.2019.1610 (2019).
25. Mick, P., Kawachi, I. & Lin, F. R. The association between hearing loss and social isolation in older adults. *Otolaryngol. Head. Neck Surg.* **150**, 378–384 (2014).
26. Shukla, A. et al. Hearing loss, loneliness, and social isolation: a systematic review. *Otolaryngol. Head. Neck Surg.* **162**, 622–633 (2020).
27. Uchida, Y. et al. Age-related hearing loss and cognitive decline—the potential mechanisms linking the two. *Auris Nasus Larynx* **46**, 1–9 (2019).
28. Kuiper, J. S. et al. Social relationships and risk of dementia: a systematic review and meta-analysis of longitudinal cohort studies. *Ageing Res. Rev.* **22**, 39–57 (2015).
29. Livingston, G. et al. Dementia prevention, intervention, and care. *Lancet* **390**, 2673–2734 (2017).
30. Colby, S. & McMurray, B. Efficiency of spoken word recognition slows across the adult lifespan. *Cognition* **240**, 105588, https://psyarxiv.com/gcj105576 (2023).
31. Blackwell, D. L., Lucas, J. W. & Clarke, T. C. Summary health statistics for US adults: national health interview survey, 2012. *Vital and health statistics. Series 10, Data from the National Health Survey*, 1–161 (2014).
32. Wie, O. B., Hugo Pripp, A. & Tvete, O. Unilateral deafness in adults: effects on communication and social interaction. *Ann. Otol. Rhinol. Laryngol.* **119**, 772 (2010).
33. Bond, M. et al. The effectiveness and cost-effectiveness of cochlear implants for severe to profound deafness in children and adults: a systematic review and economic model. *Health Technol. Assess.* **13**, 1–330 (2009).
34. Francis, H. W., Chee, N., Yeagle, J., Cheng, A. & Niparko, J. K. Impact of cochlear implants on the functional health status of older adults. *Laryngoscope* **112**, 1482–1488 (2002).
35. Saunders, J. E., Francis, H. W. & Skarzynski, P. H. Measuring success: cost-effectiveness and expanding access to cochlear implantation. *Otol. Neurotol.* **37**, e135–e140 (2016).
36. Hast, A., Schlücker, L., Digeser, F., Liebscher, T. & Hoppe, U. Speech perception of elderly cochlear implant users under different noise conditions. *Otol. Neurotol.* **36**, 1638–1643 (2015).
37. Lenarz, M., Sönmez, H., Joseph, G., Büchner, A. & Lenarz, T. Long-term performance of cochlear implants in postlingually deafened adults. *Otolaryngol. Head. Neck Surg.* **147**, 112–118 (2012).
38. Mahmoud, A. F. & Ruckenstein, M. J. Speech perception performance as a function of age at implantation among postlingually deaf adult cochlear implant recipients. *Otol. Neurotol.* **35**, e286–e291 (2014).
39. Fetterman, B. L. & Domico, E. H. Speech recognition in background noise of cochlear implant patients. *Otolaryngol. Head. Neck Surg.* **126**, 257–263 (2002).
40. Noble, W., Tyler, R. S., Dunn, C. C. & Bhullar, N. Younger-and older-age adults with unilateral and bilateral cochlear implants: speech and spatial hearing self-ratings and performance. *Otol. Neurotol.* **30**, 921 (2009).
41. Humes, L. E., Kidd, G. & Lentz, J. Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Front. Syst. Neurosci.* **7** https://doi.org/10.3389/fnsys.2013.00055 (2013).
42. Rowland, J. P., Dirks, D. D., Dubno, J. R. & Bell, T. S. Comparison of speech recognition-in-noise and subjective communication assessment. *Ear Hear.* **6**, 291–296 (1985).
43. Hustedde, C. G. & Wiley, T. L. Consonant-recognition patterns and self-assessment of hearing handicap. *J. Speech Lang. Hear. Res.* **34**, 1397–1409 (1991).
44. Horwitz, A. R. & Turner, C. W. The time course of hearing aid benefit. *Ear Hear.* **18**, 1–11 (1997).
45. Moberly, A. C., Houston, D. M., Harris, M. S., Adunka, O. F. & Castellanos, I. Verbal working memory and inhibition-concentration in adults with cochlear implants. *Laryngoscope Investig. Otolaryngol.* **2**, 254–261 (2017).
46. Heydebrand, G., Hale, S., Potts, L., Gotter, B. & Skinner, M. Cognitive predictors of improvements in adults' spoken word recognition six months after cochlear implant activation. *Audiol. Neurotol.* **12**, 254–264 (2007).
47. Hua, H., Johansson, B., Magnusson, L., Lyxell, B. & Ellis, R. J. Speech recognition and cognitive skills in bimodal cochlear implant users. *J. Speech Lang. Hear. Res.* **60**, 2752–2763 (2017).
48. O'Neill, E. R., Kreft, H. A. & Oxenham, A. J. Cognitive factors contribute to speech perception in cochlear-implant users and age-matched normal-hearing listeners under vocoded conditions. *J. Acoust. Soc. Am.* **146**, 195–195 (2019).
49. Skidmore, J. A., Vasil, K. J., He, S. & Moberly, A. C. Explaining speech recognition and quality of life outcomes in adult cochlear implant users: complementary contributions of demographic, sensory, and cognitive factors. *Otol. Neurotol.* **41**, e795–e803 (2020).
50. Heinrich, A., Henshaw, H. & Ferguson, M. A. Only behavioral but not self-report measures of speech perception correlate with cognitive abilities. *Front. Psychol.* **7**, 576 (2016).
51. Phillips, N. A. The implications of cognitive aging for listening and the framework for understanding effortful listening (FUEL). *Ear Hear* **37**, 44s–51s (2016).
52. Van Engen, K. J. & McLaughlin, D. J. Eyes and ears: using eye tracking and pupillometry to understand challenges to speech recognition. *Hear. Res.* **369**, 56–66 (2018).
53. Winn, M. B. & Teece, K. H. Listening effort is not the same as speech intelligibility score. *Trends Hear.* **25**, 23312165211027688 (2021).
54. Toscano, J. C., Anderson, N. D. & McMurray, B. Reconsidering the role of temporal order in spoken word recognition. *Psychonom. Bull. Rev.* **20**, 1–7 (2013).
55. Connine, C. M., Blasko, D. & Titone, D. Do the beginnings of spoken words have a special status in auditory word recognition? *J. Mem. Lang.* **32**, 193–210 (1993).
56. Luce, P. A. & Cluff, M. S. Delayed commitment in spoken word recognition: evidence from cross-modal priming. *Percept. Psychophys.* **60**, 484–490 (1998).

57. Dahan, D., Magnuson, J. S. & Tanenhaus, M. K. Time course of frequency effects in spoken-word recognition: evidence from eye movements. *Cogn. Psychol.* **42**, 317–367 (2001).
58. Dahan, D., Magnuson, J. S., Tanenhaus, M. K. & Hogan, E. Subcategorical mismatches and the time course of lexical access: evidence for lexical competition. *Lang. Cogn. Process.* **16**, 507–534 (2001).
59. Luce, P. A. & Pisoni, D. B. Recognizing spoken words: the neighborhood activation model. *Ear Hear.* **19**, 1–36 (1998).
60. Allopenna, P., Magnuson, J. S. & Tanenhaus, M. K. Tracking the time course of spoken word recognition using eye-movements: evidence for continuous mapping models. *J. Mem. Lang.* **38**, 419–439 (1998).
61. McMurray, B., Samelson, V. S., Lee, S. H. & Tomblin, J. B. Individual differences in online spoken word recognition: Implications for SLI. *Cogn. Psychol.* **60**, 1–39 (2010).
62. McMurray, B., Clayards, M., Tanenhaus, M. K. & Aslin, R. N. Tracking the time course of phonetic cue integration during spoken word recognition. *Psychon. Bull. Rev.* **15**, 1064–1071 (2008).
63. Yee, E. & Sedivy, J. C. Eye movements to pictures reveal transient semantic activation during spoken word recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* **32**, 1–14 (2006).
64. McClelland, J. L. & Elman, J. L. The TRACE model of speech perception. *Cogn. Psychol.* **18**, 1–86 (1986).
65. Hannagan, T., Magnuson, J. & Grainger, J. Spoken word recognition without a TRACE. *Front. Psychol.* **4** https://doi.org/10.3389/fpsyg.2013.00563 (2013).
66. Rigler, H. et al. The slow developmental timecourse of real-time spoken word recognition. *Dev. Psychol.* **51**, 1690–1703 (2015).
67. Sekerina, I. A. & Brooks, P. J. Eye movements during spoken word recognition in Russian children. *J. Exp. Child Psychol.* **98**, 20–45 (2007).
68. Spivey, M. J. & Marian, V. Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychol. Sci.* **10**, 281–284 (1999).
69. McMurray, B., Farris-Trimble, A., Seedorff, M. & Rigler, H. The effect of residual acoustic hearing and adaptation to uncertainty in Cochlear Implant users. *Ear Hear.* **37**, 37–51 (2016).
70. Klein, K., Walker, E. & McMurray, B. Delayed lexical access and cascading effects on spreading semantic activation during spoken word recognition in children with hearing aids and cochlear implants: evidence from eye-tracking. *Ear Hear.* https://psyarxiv.com/mdzn7/, https://doi.org/10.1097/AUD.0000000000001286 (2022).
71. Smith, F. X. & McMurray, B. Lexical access changes based on listener needs: real-time word recognition in continuous speech in cochlear implant user. *Ear Hear.* **43**, 1487–1501, https://psyarxiv.com/wyaxd/ (2022).
72. Clopper, C. & Walker, A. Effects of lexical competition and dialect exposure on phonological priming. *Lang. Speech* **60**, 85–109 (2017).
73. McMurray, B., Tanenhaus, M. K. & Aslin, R. N. Within-category VOT affects recovery from "lexical" garden paths: evidence against phoneme-level inhibition. *J. Mem. Lang.* **60**, 65–91 (2009).
74. Kapnoula, E. C., Edwards, J. & McMurray, B. Gradient activation of speech categories facilitates listeners' recovery from lexical garden paths, but not perception of speech-in-noise. *J. Exp. Psychol. Hum. Percept. Perform.* **47**, 578–595 (2021).
75. Dorman, M. F. et al. Bimodal cochlear implants: the role of acoustic signal level in determining speech perception benefit. *Audiol. Neurotol.* **19**, 234–238 (2014).
76. Gantz, B. J., Turner, C., Gfeller, K. E. & Lowder, M. W. Preservation of hearing in cochlear implant surgery: advantages of combined electrical and acoustical speech processing. *Laryngoscope* **115**, 796–802 (2005).
77. Yee, E., Blumstein, S. E. & Sedivy, J. C. Lexical-semantic activation in Broca's and Wernicke's aphasia: evidence from eye movements. *J. Cogn. Neurosci.* **20**, 592–612 (2008).
78. Ben-David, B. M. et al. Effects of aging and noise on real-time spoken word recognition: evidence from eye movements. *J. Speech Lang. Hear. Res.* **54**, 243–262 (2011).
79. Henry, B. A., Turner, C. W. & Behrens, A. Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. *J. Acoust. Soc. Am.* **118**, 1111–1121 (2005).
80. Litvak, L. M., Spahr, A. J., Saoji, A. A. & Fridman, G. Y. Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *J. Acoust. Soc. Am.* **122**, 982–991 (2007).
81. Chatterjee, M. & Oberzut, C. Detection and rate discrimination of amplitude modulation in electrical hearing. *J. Acoust. Soc. Am.* **130**, 1567–1580 (2011).
82. Spahr, A. J. et al. Development and validation of the AzBio sentence lists. *Ear Hear.* **33**, 112 (2012).
83. Gatehouse, S. & Noble, W. The speech, spatial and qualities of hearing scale (SSQ). *Int. J. Audiol.* **43**, 85–99 (2004).
84. Apfelbaum, K. S., Goodwin, C., Blomquist, C. & McMurray, B. The development of lexical competition in written and spoken word recognition. *Q. J. Exp. Psychol.* **76**, 196–219 (2022).
85. Farris-Trimble, A. & McMurray, B. Test-retest reliability of eye tracking in the visual world paradigm for the study of real-time spoken word recognition. *J. Speech Lang. Hear. Res.* **56**, 1328–1345 (2013).
86. Magnuson, J. S., Mirman, D. & Myers, E. B. in *The Oxford Handbook of Cognitive Psychology* (ed D. Reisberg) 412–441 (The Oxford University Press, 2013).
87. Dahan, D. & Tanenhaus, M. K. Continuous mapping from sound to meaning in spoken-language comprehension: immediate effects of verb-based thematic constraints. *J. Exp. Psychol. Learn. Mem. Cogn.* **30**, 498–513 (2004).
88. McMurray, B., Chiu, S., Sarrett, M., Black, A. & Aslin, R. N. Decoding the neural dynamics of word recognition from scalp EEG. *NeuroImage* **260**, 119457 https://psyarxiv.com/119456ng119452k/ (2022).
89. Broderick, M. P., Di Liberto, G. M., Anderson, A. J., Rofes, A. & Lalor, E. C. Dissociable electrophysiological measures of natural language processing reveal differences in speech comprehension strategy in healthy ageing. *Sci. Rep.* **11**, 4963 (2021).
90. Kaandorp, M. W., De Groot, A. M. B., Festen, J. M., Smits, C. & Goverts, S. T. The influence of lexical-access ability and vocabulary knowledge on measures of speech recognition in noise. *Int. J. Audiol.* **55**, 157–167 (2016).
91. Kaandorp, M. W., Smits, C., Merkus, P., Festen, J. M. & Goverts, S. T. Lexical-access ability and cognitive predictors of speech recognition in noise in adult cochlear implant users. *Trends Hear.* **21**, 2331216517743887 (2017).
92. Saxton, J. A. et al. Speed and capacity of language processing test: normative data from an older American community-dwelling sample. *Appl. Neuropsychol.* **8**, 193–203 (2001).
93. Pichora-Fuller, M. K. et al. Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). *Ear Hear.* **37**, 5S–27S (2016).
94. McCoy, S. L. et al. Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Q. J. Exp. Psychol. A* **58**, 22–33 (2005).
95. Lin, F. R. et al. Association of hearing impairment with brain volume changes in older adults. *NeuroImage* **90**, 84–92 (2014).
96. Sarant, J. et al. The effect of cochlear implants on cognitive function in older adults: initial baseline and 18-month follow up results for a prospective international longitudinal study. *Front Neurosci.* **13**, 789 (2019).

97. Kapnoula, E. C. & McMurray, B. Inhibitory processes are plastic: training alters competition between words. *J. Exp. Psychol. Gen.* **145**, 8–30 (2016).
98. Roland, Jr., J. T., Gantz, B. J., Waltzman, S. B., Parkinson, A. J. & Group, T. M. C. T. United States multicenter clinical trial of the cochlear nucleus hybrid implant system. *Laryngoscope* **126**, 175–181 (2016).
99. Gantz, B. J., Turner, C. W. & Gfeller, K. E. Acoustic plus electric speech processing: Preliminary results of a multicenter clinical trial of the Iowa/nucleus hybrid implant. *Audiol. Neurotol.* **11**, 63–68 (2006).
100. Tolisano, A. M. et al. Identifying disadvantaged groups for cochlear implantation: demographics from a large cochlear implant program. *Ann. Otol. Rhinol. Laryngol.* **129**, 347–354 (2020).
101. Holder, J. T., Reynolds, S. M., Sunderhaus, L. W. & Gifford, R. H. Current profile of adults presenting for preoperative cochlear implant evaluation. *Trends Hear.* **22**, 2331216518755288 (2018).
102. Shen, Y. & Richards, V. M. A maximum-likelihood procedure for estimating psychometric functions: thresholds, slopes, and lapses of attention. *J. Acoust. Soc. Am.* **132**, 957–967 (2012).
103. Winn, M. B., Won, J. H. & Moon, I. J. Assessment of spectral and temporal resolution in cochlear implant users using psychoacoustic discrimination and speech cue categorization. *Ear Hear.* **37**, e377 (2016).
104. Galvin Iii, J. J., Oba, S., Başkent, D. & Fu, Q.-J. Modulation frequency discrimination with single and multiple channels in cochlear implant users. *Hear. Res.* **324**, 7–18 (2015).
105. Apfelbaum, K. S., Klein-Packard, J. & McMurray, B. The pictures who shall not be named: empirical support for benefits of preview in the Visual World Paradigm. *J. Mem. Lang.* **121**, 104279 (2021).
106. McMurray, B. *EyelinkAnalysis [computer software]*, http://osf.io/c35tg (2022, Feb. 17).
107. McMurray, B. I'm not sure that curve means what you think it means: toward a more realistic understanding of eye-movement control in the Visual World Paradigm. *Psychon. Bull. Rev.* **30**, 102–146 (2023).
108. Hendrickson, K. et al. The profile of real-time competition in spoken and written word recognition: more similar than different. *Q. J. Exp. Psychol.* **75**, 1653–1673 (2022).
109. Oleson, J. J., Cavanaugh, J. E., McMurray, B. & Brown, G. Detecting time-specific differences between temporal nonlinear curves: analyzing data from the visual world paradigm. *Stat. Methods Med. Res.* **26**, 2708–2725 (2017).
110. McMurray, B. *Nonlinear curvefitting for Psycholinguistics* [computer software], https://osf.io/4atgv/ (2017).
111. Jeppsen, C., Apfelbaum, K. S., Tomblin, J. B., Klein, K. & McMurray, B. The development of lexical processing: real-time phonological competition and semantic activation in school age children. *Quart. J. Exp. Psychol.* https://osf.io/xbyz8 (in press).
112. Little, R. J. A. A test of missing completely at random for multivariate data with missing values. *J. Am. Stat. Assoc.* **83**, 1198–1202 (1988).

Acknowledgements

The authors would like to thank Camille Dunn for assistance with patient recruiting and coordinating the audiological testing; Bruce Gantz for leadership of the grant that supported this project; Jake Oleson for assistance with the statistics; and Camila Morales, Evita Woolsey, Nikki Chen and Kate Hinz for assistance with data collection. We particularly thank Keith Baxelbaum for consultation and support during the study design and the development of the statistical tools. We thank Richard Aslin, Michael Tanenhaus, Tim Griffiths and Chris Petkov for comments on an earlier draft of this manuscript. This project was supported by NIH grants P50 DC000242 (PD: Bruce Gantz, PI: Bob McMurray), R01 DC008086 (PI: McMurray), and NSF 2104015 (PI: Kutlu, Mentor: McMurray).

Author contributions

B.M., F.X.S., and S.C. designed the study; K.R., F.X.S., S.C. and B.M. developed the materials; K.R., M.H., S.C., J.M., C.J., and F.X.S. collected the data; B.M. performed the analysis, with consultation from S.C., F.X.S., and E.K.; All authors contributed to the selection of the analytic approach and the interpretation of the results; B.M. and M.H. wrote the first draft; All authors contributed to the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-51514-3.

Correspondence and requests for materials should be addressed to Bob McMurray.

Peer review information *Nature Communications* thanks Stefanie Kuchinsky, and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.