UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Dimension-Based Statistical Learning in Older Adults

Permalink
https://escholarship.org/uc/item/6qn2q5rc

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Hodson, Alana
DiNino, Mishaela
Shinn-Cunningham, Barbara
et al.

Publication Date
2022

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Dimension-Based Statistical Learning in Older Adults

Alana J. Hodson (ahodson@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Mishaela DiNino (mdinino@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Barbara G. Shinn-Cunningham (bgsc@andrew.cmu.edu)
Neuroscience Institute, Carnegie Mellon University, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Lori L. Holt (loriholt@cmu.edu)
Department of Psychology, Carnegie Mellon University, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Abstract

The ability to perceptually "reweight" acoustic dimensions in response to changes in distributional statistics is known as dimension-based statistical learning (DBSL). However, it is currently unknown whether DBSL imposes a cognitive load. Older adults, who typically have age-related declines in cognitive ability, may be sensitive to this load. We examined young and older adults' categorization of *beer* and *pier* sounds when the statistical relationship between VOT and F0 was consistent with that of American English, followed by a condition in which those statistics were reversed. Listeners made categorization decisions on each stimulus (Experiment 1), or after passive exposure to a string of stimuli (Experiment 2). In both experiments, younger and older participants demonstrated DBSL following exposure to the reversed statistics. Older adults tracked distributional statistics even when learning required accumulation of statistics over 8 sec, suggesting that rapid adaptation to regularities in speech input is robust across differing perceptual loads.

Keywords: aging; speech perception; cognition; perceptual learning; statistical learning; working memory

Introduction

Speech comprehension relies on an interplay of cognitive and perceptual processes to sort out the complex mapping from speech acoustics to linguistic representations. Talker characteristics, noise, and even room acoustics can contribute to speech acoustic variability, complicating the mapping. How we map acoustic structure to speech categories despite these challenges is the crux of the *lack of invariance problem* (Liberman et al., 1967). A wealth of studies demonstrate that one way speech perception accommodates systematic variability is via rapid perceptual learning whereby experiencing ambiguous acoustics in disambiguating contexts, either through lexical support or patterns of acoustic properties, leads to lasting adjustments in the mapping of speech input to pre-lexical representations (Kraljic & Samuel, 2005; Davis et al., 2005; Luthra et al., 2020; Idemaru & Holt, 2011; Lehet & Holt, 2020).

For example, individuals can draw on lexical knowledge to disambiguate a sound between /f/ and /s/ (e.g., *peace* is a word while *peef* is not; Norris, McQueen, & Cutler, 2003). Consistently experiencing an ambiguous /f/-/s/ sound in the lexical context of "pea__" results in a lasting perceptual shift such that the ambiguous sound is later heard more often as /s/ even in a lexically neutral (non-word) context. This lexically guided perceptual learning has been observed across stops, liquids, fricatives, and vowels (Kraljic & Samuel, 2006; Sharenborg, Mitterer, &McQueen, 2011; Drouin, Theodore, & Myers, 2016; McQueen & Mitterer, 2005).

Individuals can also take advantage of the covariance among acoustic speech cues to categorize speech sounds in an ever-changing listening environment. Speech sounds are signaled by multiple acoustic dimensions, but dimensions are not equally informative. For example, the sounds /b/ and /p/ are primarily distinguished by voice-onset-time (VOT) (/b/ has a shorter VOT and /p/ a longer VOT) but are also, secondarily, signaled by fundamental frequency (F0) (/b/ with a lower F0 and /p/ with a higher F0). This secondary F0 dimension, while informative, is often relied upon to a lesser degree. The primacy of VOT appears to reflect the long-term regularities present in American English speech in the assignment of a perceptual "weight", or diagnosticity, to acoustic speech dimensions.

When short-term input deviates from the long-term VOTxF0 covariation typical of English, as happens in accented speech, the perceptual weight of acoustic speech dimensions adjusts. This perceptual "reweighting" can be observed and quantified by exposing listeners to a distribution of speech tokens, such as *beer* and *pier*, that follow a standard English, or "canonical", pattern of VOT and F0 and then subsequently exposing them to a distribution that flips this relationship, essentially introducing a reversed artificial accent. It is possible to measure the impact of this shift in short-term speech input on perceptual weights in the two Exposure blocks by examining speech categorization

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.* ©2022 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

across tokens for which one dimension is ambiguous and the other varies. Interestingly, the perceptual weight of the primary, VOT, dimension for *beer* – *pier* remains stable, whereas F0 becomes a less diagnostic signal for speech category identity in the reverse block. The mapping of acoustics to speech categories rapidly adjusts to short-term regularities in speech input. Like lexically guided perceptual learning, this influence of short-term speech regularities, coined *dimension-based statistical learning* (DBSL), has been robustly observed across consonants and vowels (Idemaru & Holt, 2011; 2014, Liu & Holt, 2015; Zhang, Wu & Holt, 2021).

Age-Related Differences in Lexically Guided and Acoustically Guided Perceptual Learning

The majority of studies investigating short-term adaptive plasticity in speech perception have examined young, normal hearing adults. However, age-related changes in cognitive ability, such as declines in working memory, inhibitory control, and processing speed, affect speech comprehension in older adults (Hasher et al., 1991; Salthouse, 1994; Salthouse & Meinz, 1995; McCabe et al., 2010). Additionally, older adults exhibit greater levels of cognitive effort with increases in speech comprehension task difficulty and working memory load. Even when older adults comprehend speech well enough to support daily conversation, it may come at the cost of increased cognitive effort, which can negatively impact other aspects of language comprehension (Wingfield, 1996; Pichora-Fuller et al., 2016). The increased effort associated with speech processing also has been found to result in older adults adapting a "good enough" approach to language comprehension whereby they rely more on context and plausibility, or world knowledge (Ferreira & Patson, 2007). Although mild to moderate hearing loss, a common condition among adults over 65 years, exacerbates the detrimental effects of cognitive decline in language comprehension, even older adults without hearing impairment demonstrate increased reliance on heuristic-based strategies, suggesting that changes in cognitive, rather than solely perceptual, ability drive age-related differences (Amichetti, White, & Wingfield, 2016).

Given these findings, age-related changes in cognition might also yield group differences between normal-hearing younger and older adults' ability to engage perceptual learning in pre-lexical linguistic processing. At present, very few studies have investigated age-related differences in perceptual learning. One study examined whether phonemic category representations in older adults are more resistant to change compared to those of younger adults (Scharenborg & Janse, 2013). Consistent with this possibility, younger adults showed greater lexically driven perceptual learning than older adults. Yet, this single study cannot resolve whether stronger, more stable category representations in older adults drive the group difference, or if other possibilities – like less efficient perceptual learning in older adults – drive the age difference.

The Present Study: Age-Related Differences in Acoustically Guided DBSL

Here, we examine whether older adults exhibit perceptual learning driven by bottom-up regularities using DBSL and, if so, whether their learning differs from that of younger adults. Older adults, even those with no known hearing impairment, may demonstrate reduced DBSL due to a general decrease in reliance on the detailed acoustic properties conveying the statistical information needed to resolve signal ambiguity and robustly activate phonemic category representations. Additionally, older adults may demonstrate more rigid or stable category representations compared to younger adults, leading to reduced down-weighting of the secondary dimension (Scharenborg & Janse, 2013). In the present study, we aimed to address whether, and to what degree, older adults exhibit DBSL when the short-term statistical distributions in speech sounds are presented across overt identification decisions (Experiment 1) and passive exposure (Experiment 2).

In the DBSL paradigm, perceptual learning plays out pre-lexically; there is no semantic or lexical context to guide interpretation of ambiguous phonemes. Cognitive load has been shown to affect pre-lexical processing of phonemes in a graded manner, suggesting that increased load reduces the fidelity of acoustic processing (Mattys, Barden, & Samuel, 2014). If this is so, increasing cognitive load through working memory demands may render DBSL less effective such that changes in the VOT-F0 relationship are less salient. This may result in a reduction of any subsequent reweighting. Ambiguity in speech category membership, even over the brief span of a single word, may induce a measurable burden to working memory; the listener must resolve acoustic ambiguity to make a categorization decision even as the memory trace of the utterance fades. Even in statistical learning, domain-general resources, such as working memory, are thought to be recruited (Palmer & Mattys, 2016). In a DBSL paradigm, knowledge of the statistics describing the relationships among acoustic dimensions build over the course of a block as the listener hears speech tokens that convey the canonical and reversed distributions. It is possible that DBSL relieves working memory load over time as perceptual weights adjust to short-term regularities to aid the acoustic mapping. However, any effect of perceptual learning on working memory is likely negligible in the usual populations of younger adults, who typically do not have cognitive deficits; testing the performance of older adults in a DBSL paradigm could reveal whether adaption to changes in speech statistics impacts working memory.

Experiment 1 follows the format of the traditional DBSL paradigm in which the listener must make a category decision after each stimulus, restricting the build-up of statistics to one stimulus at a time before the listener must use the acquired statistical knowledge to inform a behavioral decision. Experiment 2 increases the amount of statistical information the listener must take in before making a category decision and forces the listener to accrue these statistics over a longer time period. This is done by having participants passively listen to a string of Exposure stimuli conveying the short-term VOTxF0 distribution and requiring a categorization response only to a VOT- or F0-neutral stimulus at the end of the string. We hypothesized that passive exposure would present an even greater burden to working memory as statistical information must be built up and maintained over a longer duration. This increased burden may impact adjustments to perceptual weights and therefore the magnitude of DBSL, especially in older adults.

Experiment 1

Experiment 1 compares younger and older adults' performance in an overt DBSL paradigm in which participants make a forced-choice categorization decision (*beer* or *pier*) after each stimulus. Unbeknownst to listeners, the second half of the experiment introduces a subtle "artificial accent" that flips the standard American English VOTxF0 relationship. Test stimuli ambiguous along the VOT dimension with a "high" vs. "low" F0 are used to evaluate perceptual learning, with prior literature demonstrating that a reversal in short-term speech regularities leads listeners to down-weight reliance on F0 in *beer-pier* decisions (Idemaru & Holt, 2011; 2014; 2020).

Participants A group of 30 younger adults (21 female; mean age = 19.6, SD = 1.25 years) and a group of 31 older adults (16 female; mean age = 65.5, SD = 4.12 years) completed Experiment 1 online using a Chrome browser on a laptop or desktop computer (no smartphone or tablet) with no operating system restriction. All were native speakers of American English with self-reported normal hearing recruited through Prolific (www.prolific.sc) and tested using the Gorilla Experiment Builder (www.gorilla.sc). Informed consent was obtained for all participants in compliance with a protocol approved by the Carnegie Mellon University Institutional Review Board.

Stimuli Stimuli were identical to those used in Idemaru and Holt (2011). Naturally spoken /bir/ and /pir/ sounds were manipulated along a continuum of VOT and F0 using Praat (Boersma & Weenik, 2010). VOT was manipulated in 10-ms steps from 0 ms (most like /bir/) to 30 ms (most like /pir/), and F0 was manipulated in 20-Hz steps from 200 Hz (most like /bir/) to 320 Hz (most like /pir/). This created a 7x7 two-dimensional acoustic space across VOT and F0 (Figure 1).

Design and Procedure Each participant completed a brief headphone check and volume calibration before starting the task to insure that each heard the same sounds at a clear and comfortable volume (Milne et al., 2020). Upon passing the check, participants categorized /bir/ and /pir/ sounds sampled from the Figure 1 acoustic space across two blocks with distinct short-term regularities across VOT and F0.

The first set of 10 blocks conveyed a short-term VOTxF0 regularity consistent with long-term norms of English (Canonical condition). A second set of 10 blocks conveyed an 'artificial accent' that reversed this correlation (Reverse condition). Figure 1 shows gray squares illustrating the Exposure stimuli conveying the regularity for each condition. In each block, participants heard one repetition of each Exposure stimulus and "test" stimulus - those with either a high (300 Hz) or low (220) F0 value and an ambiguous VOT value (15 ms) - in random order, for a total of 12 trials per block. Participants categorized each sound as *beer* or *pier* by clicking a box labeled with each word and were not made aware of any other study details. A self-timed break was offered every 60 trials. The 240 trials took about 25 minutes to complete.

A linear mixed-effect regression model was performed in R using *lme4* (Bates et al., 2015) to examine the effects of age group (default = young adults), condition (default = Canonical), and test stimulus (default = Low F0), as well as the interaction between those variables, on *pier* responses. The repeated measures model additionally included a random intercept effect of participant and random slope of block x test stimulus to account for inter-subject differences.

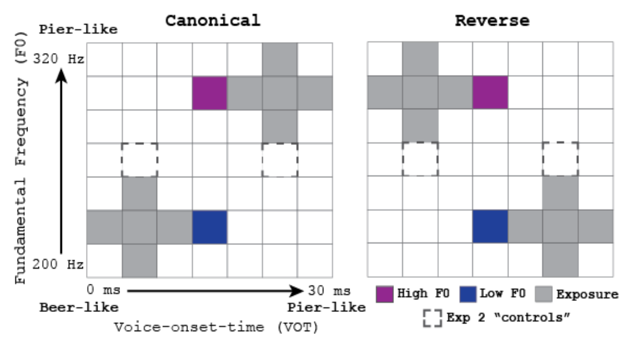

Figure 1. Schematic of acoustic stimuli VOTxF0 distribution. Gray indicates Exposure stimuli. Test stimuli, with ambiguous VOT and differing only in F0, are shown in color. Control stimuli for Experiment 2 are shown with dashed outlines.

Results and Discussion The top panel of Figure 2 shows the results, with main effects of F0 test stimulus (High F0 or Low F0; $\beta = 75.0$, SE = 5.01, $p < 0.001$) and condition (Canonical or Reverse; $\beta = 47.33$, SE = 5.89, $p < 0.001$), as well as a significant interaction between test stimulus and condition ($\beta = -75.33$, SE = 8.15, $p < 0.001$). These results replicate the finding that short-term input regularities affect perceptual weighting, as is characteristic of DBSL. In the context of Canonical speech distributions that mirror English, participants use F0 to categorize *beer-pier*. Introduction of the artificial accent leads to a rapid down-weighting of F0 such that it is much less effective in signaling category identity.

The analysis also revealed a marginal main effect of age group ($\beta = 11.97$, SE = 6.86, $p = 0.086$), likely due to the greater percent of *pier* responses across conditions observed in the older adults relative to younger adults. There was also a three-way interaction between F0, condition, and age group ($\beta = 25.98$, SE = 11.43, $p = .027$). Older adults adapt to short-term variability in the statistics among speech sounds. Yet, the magnitude of F0 down-weighting was slightly, but significantly, smaller for the older adults compared to young adults. The significant three-way interaction among age-group, block, and test stimulus provides support for our initial prediction that older adults would exhibit less down-weighting compared to younger adults. However, our sample size was small, and individual differences in responses among both age groups may have led to a difference in the magnitude of down-weighting F0. Further examination with greater numbers of both young and older adults will be required to investigate the magnitude of any potential age group differences in DBSL.

Experiment 2

Experiment 2 investigated whether older adults adapt to short-term speech input regularities under conditions that may be more taxing to working memory – listening passively to 10 Exposure stimuli prior to overt categorization of a final, Test, stimulus in the sequence. In this paradigm, F0 down-weighting in the Reverse block would be evoked only if listeners can accumulate the VOTxF0 regularity across passive exposure on a trial-by-trial basis lasting about 8 seconds, placing higher demands on working memory than in Experiment 1.

Methods

Participants A group of 30 older adults (22 female; *M* = 64.2, *SD* = 3.92 years) and 29 younger adults (20 female, 6 male, 2 non-binary, 1 chose not to disclose; *M* = 22.9, *SD* = 3.10) completed Experiment 2. Again, all participants reported normal hearing. Four older adults and three younger adult subjects were excluded from further analysis as they did not meet the criteria of at least 70 percent accuracy across total Control trials (see below). Recruitment and inclusion criteria followed the Experiment 1 strategy.

Stimuli. The Experiment 1 distribution of Exposure stimuli for Canonical and Reverse conditions was used in Experiment 2. The two Test stimuli (High F0 and Low F0) were used to quantify perceptual learning as F0 down-weighting for each subject. Additionally, two "Control" stimuli (the dashed-outlined squares in Figure 1) provided an attention-check for online participants' performance. These two stimuli were ambiguous along the F0 dimension with unambiguous VOT (5 ms (/bir/-like) and 25 ms (/pir/-like)).

Design and Procedure. On each trial, a randomized sequence of the five *beer* and five *pier* Exposure stimuli (300 ms silence between each stimulus) was drawn from either the Canonical (Block 1) or Reverse (Block 2) distributions and was followed by a 600 ms silent pause. Then, a final Test stimulus (ambiguous VOT, Low or High F0) was accompanied by a black question mark on the screen. Participants categorized the Test stimulus as *beer* or *pier* by pressing the F or J key, respectively. Presentation of Exposure sounds in the sequence was accompanied by a visual "progress bar" incrementing at the onset of each stimulus. The total duration of each trial was approximately 8 seconds.

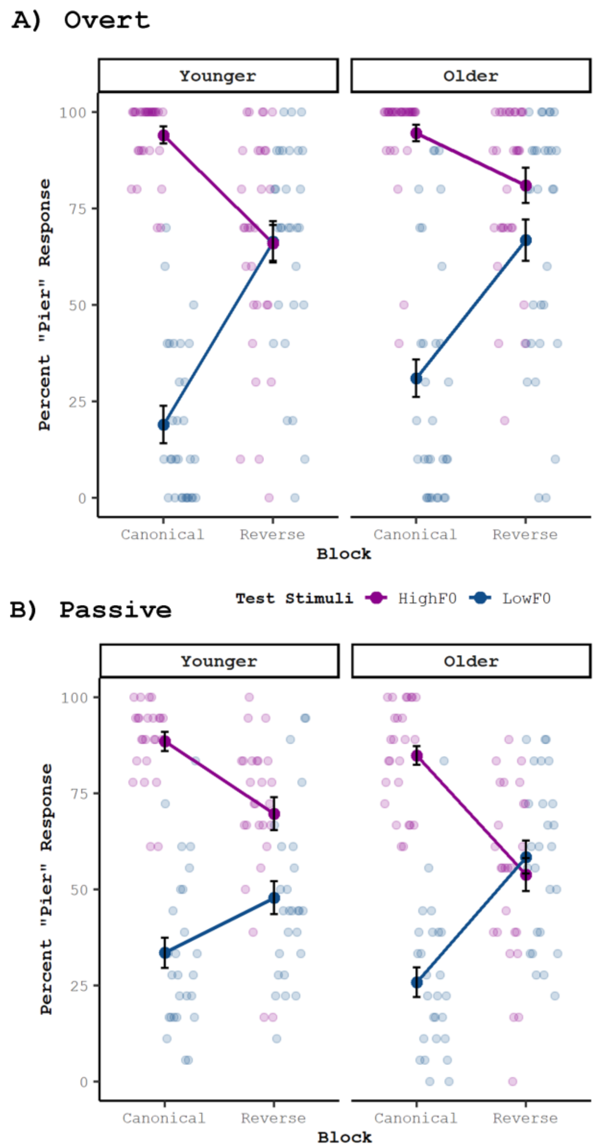

Figure 2. Mean percent of pier responses for the test stimuli by older and young adults in the A) overt response paradigm (Experiment 1) and B) passive listening paradigm (Experiment 2). Individual data points are plotted behind the group means for each test stimulus and per block. Error bars represent -/+ 1 SE.

Each block included 12 trials that ended with a Control stimulus (6 *beer*-like, 6 *pier*-like). The remaining 36 trials ended with the ambiguous VOT test stimuli (18 High F0, 18 Low F0), which served as indicators of F0 down-weighting. Participants took a self-timed break every 24 trials. The 96 trials total took about 20 minutes to complete.

As in Experiment 1, a linear mixed effects regression model assessed the percentage of test stimuli categorized as "pier" as a function of condition (default = Canonical), age group (default = young adults), and test stimulus type (default = Low F0). Random intercept of subject and random slope of condition x test stimulus that varied by subject were included in the model.

Results and Discussion. Results are shown in the bottom panels of Figure 2. There were significant main effects of test stimulus ($\beta = 54.92$, SE = 4.97, $p < 0.001$) and condition ($\beta = 14.32$, SE = 4.88, $p = .005$), and no effect of age group ($\beta = -7.69$, SE = 5.54, $p = .171$). There were significant interactions between test stimulus and condition ($\beta = -33.12$, SE = 7.63, $p < 0.001$), and age group and condition ($\beta = 18.16$, SE = 6.89, $p = .011$). Finally, there was significant three-way interaction among test stimulus, condition, and age group ($\beta = -30.34$, SE = 10.79, $p = .007$). As can be seen in Figure 2 (bottom panel), this three-way interaction came about because the magnitude of F0 down-weighting was slightly larger for the older adults compared to younger adults.

These results show that both younger and older adults' speech perception is impacted by passive exposure to short-term input regularities, with introduction of an accent leading F0 to become a *less* diagnostic indicator of speech category membership.

Less clear is whether there are true age group differences. In fact, the significant three-way interaction in Experiment 2 – for which greater working memory demands would predict greater challenges for older listeners -- was opposite this prediction, and opposite of Experiment 1. Thus, it seems most likely that the small magnitude differences across age groups highlight individual variability across a modest sample rather than true age effects. The small to medium effect sizes for these condition x test stimulus x age group interactions (Exp 1, $\eta^2_p = .08$; Exp 2, $\eta^2_p = .14$) support this possibility. Yet, the results of both experiments are very clear in that older, like younger adults, track regularities in short-term speech input and that the effectiveness of acoustic dimensions in supporting subsequent speech categorization is weighted accordingly.

General Discussion

Listeners rapidly adjust weighting of auditory dimensions in response to variation in the short-term input of speech sounds. This dimension-based statistical learning has been consistently observed in younger adults with normal hearing and no reported cognitive deficits (e.g., Idemaru & Holt 2011; 2014). Of interest in the present study was whether *older* adults with normal hearing adapt to altered distributional statistics of speech sounds, given the increased cognitive constraints in this population relative to younger adults. Results from our experiments indicate that older adults demonstrate acoustically guided perceptual learning in a DBSL paradigm, even when there are higher demands on working memory.

The finding that older adults engage in speech-related perceptual adaptation is consistent with previous research. For example, Peelle & Wingfield (2005) found that older adults improved their recognition of time-compressed and noise-vocoded sentences with practice. Scharenborg & Janse (2013) observed that older adults demonstrated lexically guided perceptual learning, altering their phonetic category boundaries upon exposure to nonword stimuli. However, unlike degraded sentences and lexically guided perceptual learning, DBSL is driven *acoustically* and thus occurs entirely pre-lexically. The present study demonstrates that older adults can rely solely on an acoustic signal to adjust perceptual mapping of speech.

Results from the present study suggest that DBSL may not be sensitive to age-related working memory decline—older adults effectively down-weighted F0 in both Experiments 1 and 2. Still, we did not assess participants' working memory capacity in this study. Although decreased working memory performance is typical among older adult populations, it is nevertheless possible that the older adults in these experiments had excellent working memory, resulting in better ability to track speech statistics over time compared to other individuals in their age group. It may also be the case that the amount of statistical information acquired or the increased trial duration in Experiment 2 was not a sufficient enough burden to working memory to affect performance.

Alternatively, it may be the case that accrual of statistical regularities in speech acoustics does not burden working memory or cognitive resources due to the bottom-up, automatic nature of the learning mechanism. However, studies modulating cognitive load on statistical learning of speech stimuli have found that increased cognitive load results in worse statistical learning performance, suggesting that even a learning mechanism that can facilitate changes in perceptual weights passively (Experiment 2), may still be supported by domain-general working memory processes (Palmer & Mattys, 2016; Mitterer & Mattys, 2017). Whether or not passive learning of speech statistics is effortful in a DBSL paradigm may be more objectively understood using pupillometry to measure changes in pupil size as a trial unfolds.

Prior literature shows that older adults have difficulty with syntactically complex sentence comprehension, even when hearing acuity is normal (Wingfield, Peelle, & Grossman, 2003; Tun, McCoy, & Wingfield, 2009). Findings from the present study suggest that older adults' perceptual learning from acoustic cues is largely spared. Challenges with speech perception within this population may instead come from downstream language processes, in which cognitive effort is increased and/or cognitive effects begin to challenge comprehension. Future research could investigate whether acoustically driven perceptual learning is affected when a greater number of statistics must be tracked or if learning must be applied across different stimuli of a similar voicing class (such as generalizing to other /b/-/p/ words).

We observed significant differences in the magnitude of F0 down-weighting between age groups that should be interpreted with caution. In Experiment 1, our cohort of older adults demonstrated a bias toward responding *pier* across all test stimuli, as well as seemingly down-weighting F0 to a lesser degree, relative to our group of younger adults. On the other hand, the statistics in Experiment 2 show even more down-weighting by the older adults in a paradigm in which one might expect that the increased working memory demands would make older adults *less* sensitive to evolving statistical regularities. Yet, as can be observed in the wide spread of individual data points in Figure 2, individual participant results indicate considerable variability in the degree of down-weighting across listeners, even among younger adults. While the findings from these experiments offer insight into pre-lexical perceptual learning abilities of older adults, interpreting age-related changes in DBSL will require further study.

In summary, despite cognitive decline with age, older adults remain able to track rapidly evolving regularities in acoustic input and use those regularities to dynamically adjust reliance on different acoustic dimensions. This should serve older adults well in coping with acoustic variability in speech sounds. Future research might address whether acoustically guided perceptual learning is more cognitively demanding for older adults – a phenomena not addressed by the paradigms we describe here, but which may affect perceptual learning during the extended time periods over which real-world conversations occur.

Acknowledgments

MD was supported by NIH NIDCD F32 DC018460. AJH was supported by and T32GM081760 to LLH. LLH was supported by R21DC019217. NSF BCS 1941357 also supported the work.

References

Amichetti, N. M., White, A. G., & Wingfield, A. (2016). Multiple solutions to the same problem: utilization of plausibility and syntax in sentence comprehension by older adults with impaired hearing. *Frontiers in psychology*, *7*, 789.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Boersma, P., & Weenink, D. (2010). *Praat: Doing phonetics by computer* (Version 5.0) [Computer program]. http://www.praat.org/

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General*, *134*(2), 222.

Drouin, J. R., Theodore, R. M., & Myers, E. B. (2016). Lexically guided perceptual tuning of internal phonetic category structure. *The Journal of the Acoustical Society of America*, *140*(4), EL307-EL313.

Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, *1*(1-2), 71-83.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of experimental psychology: Learning, memory, and cognition*, *17*(1), 163.

Idemaru, K., & Holt, L.L. (2011). Word recognition reflects dimension-based statistical learning. *Journal of Experimental Psychology: Human Perception and Performance, 37(6)*, 1939-1956.

Idemaru, K., & Holt, L.L. (2014). Specificity of dimension-based statistical learning in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 40(3)*, 1009-1021.

Idemaru, K., & Holt, L.L. (2020). Generalization of dimension-based statistical learning. *Attention, Perception, & Psychophysics*, *82(4)*, 1744-1762.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal?. *Cognitive psychology*, *51*(2), 141-178.

Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic bulletin & review*, *13*(2), 262-268.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review*, *74*(6), 431.

Liu, R., & Holt, L. L. (2015). Dimension-based statistical learning of vowels. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1783.

Luthra, S., Correia, J. M., Kleinschmidt, D. F., Mesite, L., & Myers, E. B. (2020). Lexical information guides retuning of neural patterns in perceptual learning for speech. *Journal of Cognitive Neuroscience*, *32*(10), 2001-2012.

Mattys, S. L., Barden, K., & Samuel, A. G. (2014). Extrinsic cognitive load impairs low-level speech perception. *Psychonomic bulletin & review*, *21*(3), 748-754.

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222.

McQueen, J. M., & Mitterer, H. (2005). Lexically-driven perceptual adjustments of vowel categories. ISCA.

Milne, A. E., Bianco, R., Poole, K. C., Zhao, S., Oxenham, A. J., Billig, A. J., & Chait, M. (2021). An online headphone screening test based on dichotic pitch. *Behavior Research Methods*, *53*(4), 1551-1562.

Mitterer, H., & Mattys, S. L. (2017). How does cognitive load influence speech perception? An encoding hypothesis. *Attention, Perception, & Psychophysics*, *79*(1), 344-351.

Norris, D., McQueen, J.M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47(2)*, 204-238.

Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *Quarterly Journal of Experimental Psychology*, *69*(12), 2390-2401.

Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of*

Experimental Psychology: Human Perception and Performance, *31*(6), 1315.
Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E. & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and hearing*, *37*, 5S-27S.
Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology*, *8*(4), 535.
Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory, and speed. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *50*(6), P297-P306.
Scharenborg O., & Janse, (2013). Comparing lexically guided perceptual learning in younger and older listeners. *Attention, Perception, & Psychophysics 75(3)*, 525-536.
Scharenborg, O., Mitterer, H., & McQueen, J. M. (2011). Perceptual learning of liquids. In *Interspeech 2011: 12th Annual Conference of the International Speech Communication Association* (pp. 149-152).
Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and aging*, *24*(3), 761.
Wingfield, A. (1996). Cognitive factors in auditory performance: context, speed of processing, and constraints of memory. *Journal of the American Academy of Audiology*, *7*(3), 175-182.
Wingfield, A., Peelle, J. E., & Grossman, M. (2003). Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. *Aging, Neuropsychology, and Cognition*, *10*(4), 310-322.
Winn, M.B., Chatterjee, M., & Idsardi, W.J. (2013). Roles of voice onset time and F0 in stop consonant voicing perception: Effects of masking noise and low-pass filtering. *Journal of Speech, Language, and Hearing Research*, *56*, 1097-1107.
Zhang, X., Wu, Y. C., & Holt, L. L. (2021). The Learning Signal in Perceptual Tuning of Speech: Bottom Up Versus Top-Down Information. *Cognitive Science*, *45*(3), e12947.