# Chapter 55 How Early Aging and Environment Interact in Everyday Listening: From Brainstem to Behavior Through Modeling

Barbara Shinn-Cunningham, Dorea R. Ruggles, and Hari Bharadwaj

**Abstract** We recently showed that listeners with normal hearing thresholds vary 6 in their ability to direct spatial attention and that ability is related to the fidelity of 7 temporal coding in the brainstem. Here, we recruited additional middle-aged lis-8 teners and extended our analysis of the brainstem response, measured using the 9 frequency-following response (FFR). We found that even though age does not pre-10 dict overall selective attention ability, middle-aged listeners are more susceptible to 11 the detrimental effects of reverberant energy than young adults. We separated the 12 overall FFR into orthogonal envelope and carrier components and used an existing 13 model to predict which auditory channels drive each component. We find that 14 responses in mid- to high-frequency auditory channels dominate envelope FFR, 15 while lower-frequency channels dominate the carrier FFR. Importantly, we find 16 that which component of the FFR predicts selective attention performance changes 17 with age. We suggest that early aging degrades peripheral temporal coding in mid-18 to-high frequencies, interfering with the coding of envelope interaural time differ-19 ences. We argue that, compared to young adults, middle-aged listeners, who do not 20 have strong temporal envelope coding, have more trouble following a conversation 21 in a reverberant room because they are forced to rely on fragile carrier ITDs that 22 are susceptible to the degrading effects of reverberation. 23

1

2

3

4

<sup>B. Shinn-Cunningham, Ph.D. (⊠) • D.R. Ruggles, Ph.D. • H. Bharadwaj, M.S.
Department of Biomedical Engineering, Boston University Center for Computational Neuroscience and Neural Technology,
677 Beacon St., Boston, MA 02215, USA
e-mail: shinn@bu.edu</sup> 

Author's Proof

B. Shinn-Cunningham et al.

### 24 **1 Introduction**

The cacophony of voices, noises, and other sounds that bombards our ears in 25 many social settings makes it challenging to focus selective auditory attention. 26 Various acoustic cues allow us to group sound components into perceptual objects 27 to which we can direct attention (Darwin 1997; Shinn-Cunningham 2008; Shamma 28 and Micheyl 2010; Shamma et al. 2011). In most common settings, reflected 29 sound energy intensifies the problem of separating sound sources and selecting 30 the source of interest by blurring the sound features that support source segrega-31 tion and selection. 32

Many listeners report difficulties in everyday situations demanding selective 33 attention, especially as they age (Leigh-Paffenroth and Elangovan 2011; Noble 34 et al. 2012). We wondered if these problems are most evident when reverberant 35 energy challenges the auditory system. We designed a task in which listeners 36 had to focus spatial attention on a center, target speech stream in a mixture of 37 three otherwise identical streams of spoken digits, and then varied the level of 38 reverberation (Ruggles et al. 2011; Ruggles and Shinn-Cunningham 2011). By 39 design, listeners are likely to rely on interaural timing differences (ITDs) to 40 perform this task (Ruggles et al. 2011). Since reverberant energy causes inter-41 aural decorrelation, we found, as expected, that selective attention performance 42 got worse with reverberation. We also found that individual ability on our task 43 was correlated both with perceptual sensitivity to frequency modulation (FM) 44 and overall strength of the frequency-following response (FFR; see also Strelcyk 45 and Dau 2009). However, we had too few middle-aged listeners to explore age 46 effects. 47

Here, we recruited additional middle-aged listeners so that we could look for aging effects. We extended our analysis of the FFR by separating the response into the portion phase locking to the stimulus envelope ( $FFR_{ENV}$ ) and that phase locking to the stimulus carrier ( $FFR_{CAR}$ ; similar to approaches described in Aiken and Picton 2008; Gockel et al. 2011). We used existing brainstem response models (Dau 2003; Harte et al. 2010) to investigate which acoustic frequencies contribute to  $FFR_{ENV}$  and  $FFR_{CAR}$ .

# 54 2 Methods

#### 55 **2.1** Subjects

A total of 22 listeners ranging in age from 20.9 to 54.7 years participated in the experiments. All listeners had average audiometric hearing thresholds of 20-dB HL or better for frequencies from 250 to 8,000 Hz and left-right ear asymmetry of 15 dB or less at all frequencies. Of the 22 listeners, 17 were participants in earlier studies; the newly recruited five all were over 40 years of age. All gave informed consent and were paid for their participation.

### 2.2 FFR Measurement

FFRs were measured in response to a /dah/syllable presented in positive polarity for 63 2,000 trials and in inverted polarity for 2,000 trials (Ruggles et al. 2011). Trials con-64 taining eyeblinks or other artifacts were removed, leaving at least 1,800 clean trials 65 for each subject, condition, and stimulus polarity. The time series from each trial was 66 windowed with a first-order Slepian taper (Thomson 1982) and the Fourier transform 67 was computed. We generated distributions of phase-locking values (PLV) for differ-68 ent conditions using a bootstrapping procedure to produce 200 independent PLVs, 69 each computed from a draw (with replacement) of 800 trials (Ruggles et al. 2011). 70 We broke the PLV into orthogonal envelope and carrier components (FFR<sub>ENV</sub> and 71  $FFR_{CAR}$ ) at every frequency from 30 to 3,000 Hz.  $FFR_{FNV}$  was calculated with equal 72 draws from responses to each polarity, treating positive- and negative-polarity trials 73 identically. FFR<sub>CAR</sub> was determined with equal draws from responses to each polarity 74 but inverting the phase of negative-polarity trials (see also Aiken and Picton 2008; 75 Gockel et al. 2011). For each harmonic of 100 Hz, we computed the proportion of the 76 total FFR in FFR<sub>ENV</sub> and in FFR<sub>CAR</sub>. 77

### 2.3 FFR Modeling

We used an existing model of brainstem responses (Dau 2003; Harte et al. 2010) to 79 analyze the sources of the different components of the FFR. We presented the model 80 with our /dah/ syllable, then calculated the FFR by summing model outputs across 81 peripheral channels with CFs spanning the range from 100 up to 10,000 Hz. At each 82 harmonic (multiple of 100 Hz), we then computed the proportion of the total FFR 83 phase locked to the envelope and the proportion phase locked to the carrier (FFR<sub>ENV</sub> 84 and  $FFR_{CAP}$ ). We then considered the output of each peripheral channel to explore 85 which acoustic frequencies contributed to which components of the FFR. Finally, 86 we analyzed the relative strength of the contribution of each peripheral channel to 87  $FFR_{FNV}$  at the fundamental frequency (100 Hz). 88

# 2.4 Spatial Attention Task

Subjects were asked to report a sequence of four digits appearing to come from in 90 front while ignoring two competing digit streams, spoken by the same talker, from 91  $+15^{\circ}$  to  $-15^{\circ}$  azimuth (Ruggles and Shinn-Cunningham 2011). Spatial cues were 92 simulated using a rectangular-room model with three different wall characteristics 93 (Ruggles and Shinn-Cunningham 2011). Prior to statistical analyses, percent cor-94 rect scores were transformed using a rationalized arcsine unit (RAU; Studebaker 95 1985). In the task, listeners report one of the three presented words nearly 95 % of 96



the time; errors arise because of failures of selective attention, rather than memory limitations (Ruggles and Shinn-Cunningham 2011). Therefore, percent scores in the range 0.33-1.0 were linearly transformed to 0-1.0 (scores < 0.33 set to 0) prior to applying the transform.

### 101 2.5 FM Detection Task

Listeners indicated which of three 750-Hz tones (interstimulus interval 750 ms) contained 2-Hz frequency modulation (Strelcyk and Dau 2009). A two-down, one-up adaptive procedure (step size 1 Hz) estimated the 70.7 % correct FM threshold. Individual thresholds were computed by averaging the last 12 reversals per run, then averaging across six runs.

#### 107 **3 Results**

# 108 **3.1** Generators of $FFR_{ENV}$ and $FFR_{CAR}$

Figure 55.1 compares measurements and model predictions of the relative strengths of  $FFR_{ENV}$  and  $FFR_{CAR}$  at harmonics of a periodic input (F0=100 Hz). The lowest frequencies of the FFR are dominated by  $FFR_{ENV}$  and the higher harmonics are dominated by  $FFR_{CAR}$ . Both FFR components approach the noise floor in the empirical measurements by 800–900 Hz, which may help explain why the percentages of FFR\_{ENV} and FFR\_{CAR} in the total FFR both asymptote to 0.5 as frequency increases



**Fig. 55.1** Proportion of total FFR contained in  $\text{FFR}_{ENV}$  and in  $\text{FFR}_{CAR}$  at each harmonic of 100 Hz from (a) experimental measures and (b) model predictions



**Fig. 55.2** (a) Relative strength of  $\text{FFR}_{\text{ENV}}$  and  $\text{FFR}_{\text{CAR}}$  generated by each peripheral channel as a function of characteristic frequency (*CF*). (b) Relative contribution of each CF channel to strength of  $\text{FFR}_{\text{ENV}}$  at stimulus F0 of 100 Hz

and why the measured  $FFR_{ENV}$  does not drop as completely or as rapidly as the modeled  $FFR_{ENV}$  as frequency increases. 116

Modeling results also suggest that different acoustic frequencies contribute to  $FFR_{ENV}$  117 and  $FFR_{CAR}$ . In the model, peripheral channels with the lowest characteristic frequencies (CFs) tend to contribute to  $FFR_{CAR}$  and peripheral channels with the highest CFs contribute to  $FFR_{ENV}$  with a crossover point of about 2 kHz (Fig. 55.2a). The model also predicts that the channels that contribute the most to the 100-Hz  $FFR_{ENV}$  for our /dah/ syllable have CFs in the mid-to-high frequency range, around 1 kHz (Fig. 55.2a). 122

### 3.2 Effects of Reverberation and Age on Selective Attention 123

Selective attention performance decreases as reverberant energy increases, reaching chance levels for all but five listeners in the highest reverberation level 125 (Fig. 55.3; chance performance is one-third; modeling performance as a binomial 126 distribution of 600 independent trials, we computed the 95 % confidence interval 127 around this level). 128

We quantified the fidelity of envelope temporal structure encoding for each listener as the  $\text{FFR}_{\text{ENV}}$  at 100 Hz. To quantify coding of the temporal fine structure in the input stimulus, we took the average of  $\text{FFR}_{\text{CAR}}$  for four harmonics (600–900 Hz, henceforth denoted  $\text{FFR}_{\text{CAR-AV}}$ ). Importantly, these two statistics are not significantly correlated (r=0.03, p=0.905, N=22), supporting the modeling prediction that each component reflects different aspects of temporal coding precision driven by different tonotopic portions of the auditory pathway.



**Fig. 55.3** Percentage of target digits correctly reported as a function of individual listener age for the three room conditions. *Open symbols* show subjects not in Ruggles et al. (2011)

We performed a multi-way, repeated-measures ANOVA on the selective atten-136 tion results with factors of reverberation, age, FFR<sub>ENV-100</sub>, and FFR<sub>CAR-AV</sub> (treating 137 reverberation as categorical and all other factors as continuous). Although there is 138 no statistically significant effect of age on selective attention performance 139 (Fig. 55.1a; F(1, 16) = 1.42, p = 0.251), there is a significant interaction between 140 age and reverberation (F(1, 16) = 5.88, p = 0.025) and a significant main effect of 141 reverberation (F(1, 16) = 155.17,  $p = 7.01 \times 10^{-11}$ ). Although age does not predict 142 how well an individual performs overall, the toll that reverberation takes increases 143 with age. 144

# 145 3.3 Relationship Between FFR Components and Performance

Consistent with previous results showing that the total FFR strength at 100 Hz (a 146 measure dominated by envelope phase locking; see Fig. 55.1) predicted selective 147 attention ability (Ruggles et al. 2011), we find a significant main effect of  $FFR_{FNV,100}$ 148 on performance (F(1, 16) = 5.03, p = 0.040). Importantly, however, there is a 149 significant interaction between age and FFR<sub>ENV-100</sub> (F(1, 16) = 4.64, p = 0.048). There 150 is also a significant interaction between age and  $FFR_{CAR-AVE}$  (F(1, 16)=4.64, 151 p=0.047), with no main effect of FFR<sub>CAR-AVE</sub> (F(1, 16)=0.216, p=0.649). The 152 regression coefficients of the ANOVA analysis reveal that the younger a listener is, 153 the better  $FFR_{ENV-100}$  is in predicting selective attention, whereas  $FFR_{CAR-AVE}$  is a bet-154 ter predictor the older the listener. These results suggest that FFR<sub>ENV.100</sub> and 155 FFR<sub>CAR-AVE</sub> reflect different perceptual cues that each aid in selective auditory 156 attention but that are weighted differently as listeners age. 157

506

Author's Proof

Author's Proof

55 How Early Aging and Environment Interact in Everyday Listening



**Fig. 55.4** (a) FFR<sub>ENV-100</sub> as a function of age. (b)  $FFR_{CAR-AVE}$  as a function of age. *Open symbols* show subjects not in Ruggles et al. (2011)

# 3.4 Individual Differences in FFR

Figure 55.4 plots  $FFR_{ENV-100}$  and  $FFR_{CAR-AVE}$  as a function of age. While both components tend to decrease as age increases, age is not significantly correlated with either 160  $FFR_{ENV-100}$  or with  $FFR_{CAR-AVE}$ . Notably, a good percentage of the younger adult listeners have strong FFRs (particularly for  $FFR_{ENV-100}$ ), whereas nearly all the older 162 listeners have weak FFRs. Thus, most of the variance in the FFRs is from the 163 younger listeners and cannot be explained by age alone. 164

# 3.5 Relationship Between FM Detection Threshold 165 and Performance 166

We previously found that FM detection threshold, a measure thought to reflect coding fidelity of temporal fine structure (Moore and Sek 1996), was also related to attention performance (Ruggles et al. 2011). This relationship remains significant with our additional subjects, as shown in Fig. 55.5.

#### 4 Discussion

Some previous studies have found that aging reduces FFR strength (Clinard et al. 172 2010); however, not all studies have found group age effects (Vander Werff and 173 Burns 2011). Moreover, even studies that find age-related group differences have 174 not consistently found corresponding age-related differences in auditory perceptual 175 abilities (Clinard et al. 2010). The current study helps explain these discrepant 176

158





Level of Reverberation

**Fig. 55.5** Percentage of target digits correctly reported as a function of FM threshold for the two levels of reverberation where performance is above chance. *Open symbols* show subjects not in Ruggles et al. (2011)

findings, in that there is a large variation in brainstem responses even among young
adults. By looking at individual subjects and considering different components of
the FFR, we find reliable interactions between aging, perceptual ability, and specific
components of the FFR.

Our results suggest that the FFR envelope component at the fundamental fre-181 quency of the stimulus tends to become weak as listeners reach middle age, 182 possibly because the neural response to suprathreshold sound at acoustic fre-183 quencies in the mid-to-high frequency range (e.g., around 1,000 Hz) is reduced 184 in overall strength. Physiological results show that noise exposure can reduce 185 the magnitude of neural responses that are suprathreshold, even when thresh-186 olds are "normal" (Kujawa and Liberman 2009). These changes may come about 187 because low-spontaneous-rate nerve fibers are particularly vulnerable to damage 188 (Schmiedt et al. 1996). 189

In our task, performance is primarily limited by the ability to successfully direct spatial auditory attention, which may help explain why performance depends on the fidelity of envelope temporal coding. Envelope ITD cues in high-frequency sounds are known to carry spatial information; however, a number of classic laboratory experiments establish that for wideband, anechoic sounds, low-frequency carrier

- Author's Proof
  - 55 How Early Aging and Environment Interact in Everyday Listening

ITDs perceptually dominate over high-frequency spatial cues (Wightman and<br/>Kistler 1992; Macpherson and Middlebrooks 2002). The current results suggest that<br/>in reverberant settings, high-frequency ITD cues, encoded in signal envelopes, may<br/>be more important for spatial perception of wideband sounds than past laboratory<br/>studies suggest.195198<br/>199199

In anechoic conditions, temporal fine structure cues and temporal envelope cues 200 both provide reliable information for directing selective spatial auditory attention. 201 However, in reverberant settings, interaural decorrelation of temporal fine structure 202 is more severe than interaural decorrelation of envelope structure; thus, high-203 frequency envelope ITD cues may be crucial to spatial perception in everyday set-204 tings. This possibility points to the importance of providing high-frequency 205 amplification in assistive listening devices, which have typically focused on audibil-206 ity of frequencies below 8 kHz. 207

Our results hint that middle-aged listeners, who have generally weak encoding of 208 mid- to high-frequency temporal cues, rely on temporal fine structure cues to direct 209 selective spatial auditory attention. This reliance on carrier ITD cues, which are 210 relatively fragile in ordinary listening environments, may explain why middle-aged 211 listeners report difficulty when trying to converse in everyday social settings. In 212 contrast, younger listeners appear to give great perceptual weight to envelope ITD 213 cues when directing selective attention, providing them with a more reliable cue for 214 selective spatial auditory attention. 215

AcknowledgmentsThis work was sponsored by the National Institutes of Health (NIDCD R01216DC009477 to BGSC and NIDCD F31DC011463 to DR) and the National Security Science and<br/>Engineering Faculty Fellowship to BGSC.217

### References

Aiken SJ, Picton TW (2008) Envelope and spectral frequency-following responses to vowel	220
sounds. Hear Res 245:35–47	221
Clinard CG, Tremblay KL, Krishnan AR (2010) Aging alters the perception and physiological	222
representation of frequency: evidence from human frequency-following response recordings.	223
Hear Res 264:48–55	224
Darwin CJ (1997) Auditory grouping. Trends Cogn Sci 1:327-333	225
Dau T (2003) The importance of cochlear processing for the formation of auditory brainstem and	226
frequency following responses. J Acoust Soc Am 113:936-950	227
Gockel HE, Carlyon RP, Mehta A, Plack CJ (2011) The frequency following response (FFR) may	228
reflect pitch-bearing information but is not a direct representation of pitch. J Assoc Res	229
Otolaryngol 12:767–782	230
Harte JM, Ronne F, Dau T (2010) Modeling human auditory evoked brainstem responses based on	231
nonlinear cochlear processing. In: Proceedings of the 20th international congress on acoustics,	232
Sydney, 2010	233
Kujawa SG, Liberman MC (2009) Adding insult to injury: cochlear nerve degeneration after	234
"temporary" noise-induced hearing loss. J Neurosci 29:14077–14085	235
Leigh-Paffenroth ED, Elangovan S (2011) Temporal processing in low-frequency channels: effects	236
of age and hearing loss in middle-aged listeners. J Am Acad Audiol 22:393–404	237



510

- Macpherson EA, Middlebrooks JC (2002) Listener weighting of cues for lateral angle: the duplex
   theory of sound localization revisited. J Acoust Soc Am 111:2219–2236
- Moore BC, Sek A (1996) Detection of frequency modulation at low modulation rates: evidence for
   a mechanism based on phase locking. J Acoust Soc Am 100:2320–2331
- Noble W, Naylor G, Bhullar N, Akeroyd MA (2012) Self-assessed hearing abilities in middle- and
   older-age adults: a stratified sampling approach. Int J Audiol 51:174–180
- Ruggles D, Shinn-Cunningham B (2011) Spatial selective auditory attention in the presence of
   reverberant energy: individual differences in normal-hearing listeners. J Assoc Res Otolaryngol
   12:395–405
- Ruggles D, Bharadwaj H, Shinn-Cunningham BG (2011) Normal hearing is not enough to guaran tee robust encoding of suprathreshold features important in everyday communication. Proc
   Natl Acad Sci 108:15516–15521
- Schmiedt RA, Mills JH, Boettcher FA (1996) Age-related loss of activity of auditory-nerve fibers.
   J Neurophysiol 76:2799–2803
- Shamma SA, Micheyl C (2010) Behind the scenes of auditory perception. Curr Opin Neurobiol 20:
   361–366
- Shamma SA, Elhilali M, Micheyl C (2011) Temporal coherence and attention in auditory scene
   analysis. Trends Neurosci 34:114–123
- Shinn-Cunningham BG (2008) Object-based auditory and visual attention. Trends Cogn Sci 12:
   182–186
- Strelcyk O, Dau T (2009) Relations between frequency selectivity, temporal fine-structure process ing, and speech reception in impaired hearing. J Acoust Soc Am 125:3328–3345
- 260 Studebaker GA (1985) A "rationalized" arcsine transform. J Speech Hear Res 28:455-462
- 261 Thomson DJ (1982) Spectrum estimation and harmonic-analysis. Proc IEEE 70:1055–1096
- Vander Werff KR, Burns KS (2011) Brain stem responses to speech in younger and older adults.
   Ear Hear 32:168–180
- Wightman FL, Kistler DJ (1992) The dominant role of low-frequency interaural time differences
   in sound localization. J Acoust Soc Am 91:1648–1661

Corre