Relationship between speech-evoked neural responses and perception of speech in noise in older adults

Authors: Guangting Mai[1], Jyrki Tuomainen[2], Peter Howell[1]

[1] Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, London, WC1H 0AP

[2] Department of Speech, Hearing and Phonetic Sciences, Division of Psychology and Language Sciences, University College London, London, WC1N 1PF

Corresponding author: Peter Howell, Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, London, WC1H 0AP (p.howell@ucl.ac.uk)

ABSTRACT

Speech-in-noise (SPIN) perception involves neural encoding of temporal acoustic cues. Cues include temporal fine structure (TFS) and envelopes that modulate at syllable (Slow-rate ENV) and fundamental frequency (F_0-rate ENV) rates. Here the relationship between speech-evoked neural responses to these cues and SPIN perception was investigated in older adults. Theta-band phase-locking values (PLV) that reflect cortical sensitivity to Slow-rate ENV and peripheral/brainstem frequency-following responses phase-locked to F_0-rate ENV (FFR_{ENV_F0}) and TFS (FFR_{TFS}) were measured from scalp-EEG responses to a repeated speech syllable in steady-state speech-shaped (SpN) and 16-speaker babble (BbN) noises. The results showed that: 1) SPIN performance and PLV were significantly higher under SpN than BbN, implying differential cortical encoding may serve as the neural mechanism of SPIN performance that varies as a function of noise types; 2) PLV and FFR_{TFS} at resolved harmonics were significantly related to good SPIN performance, supporting the importance of phase-locked neural encoding of Slow-rate ENV and TFS of resolved harmonics during SPIN perception; 3) FFR_{ENV_F0} was not associated to SPIN performance until audiometric threshold was controlled for, indicating that hearing loss should be carefully controlled when studying the role of neural encoding of F_0-rate ENV. Implications are drawn with respect to fitting auditory prostheses.

I. INTRODUCTION

Perception of speech-in-noise (SPIN) requires the target speaker's voice to be separated from background noises. This can be achieved by processing temporal acoustic cues (Moore, 2008). There are three temporal components of speech signals that provide different information: 1) slowly-varying envelope (Slow-rate ENV) that modulates at rates of < 10 Hz and represents syllable structure (Rosen, 1992; Greenberg et al., 2003). Slow-rate ENV plays an essential role in speech perception both in quiet (Shannon et al., 1995; Arai et al., 1999) and in noise (Drullman et al., 1994); 2) The envelope that modulates at rates around F_0 (F_0-rate ENV) and provides information about voicing and pitch (Rosen, 1992). Tracking the pitch of the target speaker is critical for segregating the target voice from background sounds (Bregman, 1994; Bird and Darwin, 1998; Binns and Culling, 2005); 3) Temporal fine structure (TFS) that fluctuates at rates above F_0 (Rosen, 1992; Moore, 2008). TFS is important for perception of formant structure (Moore, 2014) as well as voicing and pitch (Rosen, 1992; Smith et al., 2002). SPIN perception is significantly better when TFS is present than when it is absent (Zeng et al., 2005; Stickney et al., 2007; Eaves et al., 2011). Neurophysiological tracking of these cues is also related to speech perception. For example, the degree of neural phase-locking to Slow-rate ENV is associated with speech intelligibility (Ahissar et al., 2001; Peelle et al., 2012; Doelling et al., 2014). Also, magnitudes of peripheral/subcortical speech-evoked frequency-following responses (FFRs) that are phase-locked to F_0-rate ENV (FFR_{ENV_F0}) (Song et al., 2010) and TFS (FFR_{TFS}) (Fujihira and Shiraishi, 2015) correlate with SPIN performance.

Older adults experience more difficulties in SPIN perception than normal-hearing

younger adults (Moore, 2014). Behavioral studies suggest that this may be attributable to age-related deterioration in processing of TFS (Lorenzi et al., 2006; Hopkins, Moore and Stone, 2008; Hopkins and Moore, 2011). Neurophysiological studies show that FFR_{ENV_F0} and FFR_{TFS} have smaller magnitudes in older, than in younger, adults (Anderson et al., 2012; Presacco et al., 2016). This suggests that poorer neuro-temporal encoding of F_0-rate ENV and TFS by older adults may account for impaired SPIN performance. Despite these findings, the relative importance of neural processing of F_0-rate ENV and TFS remains obscure. While Anderson et al. (2011) showed that FFR_{ENV_F0} was associated with SPIN ability in older adults, this result did not replicate (Schoof and Rosen, 2016). Recent studies report that in older adults FFR_{TFS}, but not FFR_{ENV_F0}, is associated with speech perception in environments where there is reverberation (Fujihira and Shiraishi, 2015). This highlights the importance of neural encoding of TFS, compared to that of F_0-rate ENV, during SPIN perception (Marsh and Campbell, 2016). Furthermore, since these studies (Anderson et al., 2011, 2012; Fujihira and Shiraishi, 2015; Presacco et al., 2016; Schoof and Rosen, 2016) were conducted in older adults with relatively normal hearing (hearing thresholds ≤ 30 dB HL at frequencies ≤ 4 kHz), they do not reflect the significant variability in audiometric hearing in aging populations (Gopinath et al., 2009; Humes et al., 2010).

The present hypothesis is that SPIN performance is associated more with phase-locked neural encoding of Slow-rate ENV and TFS compared to that of F_0-rate ENV in older adults who vary in audiometric profiles. This is based on the observations that Slow-rate ENV is critical for speech perception both in quiet and in noise (Drullman et al., 1994; Shannon et al., 1995; Arai et al., 1999) and impaired sensitivity to TFS, rather than F_0-rate ENV, significantly

4

degrades SPIN performance (Stickney et al., 2007). To test the hypothesis, behavioral and neural assessments were conducted on healthy older adult participants in a sample with significant inter-individual variability in their audiometric profiles. Participants completed a SPIN task in which speech reception thresholds were measured under two types of noise: steady-state speech-shaped noise (SpN) and 16-speaker babble noise (BbN), where background speakers in BbN had similar voices characteristics to the target voice. SpN serves as a steady-state energetic masker, whilst BbN constitutes an envelope-modulated energetic masker as well as an informational masker that leads to phonetic interference. Rosen et al. (2013) showed that speech perception is better under SpN than BbN and that BbN is a particularly challenging form of noise for older adults (Helfer and Freyman, 2008; Schoof and Rosen, 2014). Participants listened to a repeated syllable in quiet, SpN and BbN, whilst neural activity was recorded concurrently over scalp-EEG electrodes. Speech-evoked FFRs, which reflect peripheral/brainstem phase-locked encoding of F_0-rate ENV and TFS (Aiken and Picton, 2008), were obtained. Cortical activity was represented by theta-band phase-locking values (PLVs) that reflect neural sensitivities to Slow-rate ENV (Luo and Poeppel, 2007; Howard and Poeppel, 2010). Signal-to-noise ratios (SNRs) of the acoustic stimuli in SpN/BbN during the neural recording were set at either 7 or -1 dB, with the former (7 dB) being the SNR at which speech is reasonably recognizable while the latter (-1 dB) is the SNR that is relatively more challenging for SPIN perception (see *Methods*). The relationship between the neural measures and SPIN performance was then examined. It was predicted that the neural measures, especially those recorded under the challenging SNR, would correlate with the behavioral differences across noise types and individuals, and that the

5

neural-behavioral relationship would further elucidate the underlying neural mechanisms of SPIN perception in older adults.

II. METHODS

A. Participants

Participants comprised 47 older healthy adults (aged 53 to 76 years, 17 males and 30 females). All were native English speakers who had lived in Britain for at least 25 years. No clinically-diagnosed neurological diseases, language-related or psychiatric problems were reported. All participants identified 100% of the words in the speech perception task correctly when SNR was 8 dB. Twelve participants used hearing aids (HAs, either monaural or binaural). They wore their HAs with their usual settings and these were kept the same throughout the experiment. **Fig. 1** shows the pure-tone audiograms (PTA) for frequencies between 0.25 and 8 kHz of the non-HA participants measured via a MAICO MA41 Audiometer (MAICO Diagnostics, Germany). Inter-individual variability in hearing thresholds occurred particularly at high frequencies ($\geq$ 3 kHz), consistent with the distributions of audiometric profiles in aging populations (Gopinath et al., 2009; Humes et al., 2010). **Fig. 1C** shows the free-field PTAs of the HA users. They were measured by presenting pure tones via a Fostex 6301B loudspeaker (Canford Group Ltd, UK) at 0 degree azimuth positioned 1 meter away from the participants' ears. The sounds were controlled by the experimenter using a ThinkPad T430i laptop (Lenovo Group Ltd). Sound pressure levels (SPL) at 1 meter distance were measured and were then converted to hearing level (dB HL) for each frequency according to the ISO calibration guideline for free-field PTAs (ISO

389-7-2005, British Society of Audiology, 2008). The free-field PTA confirmed that the HAs were functioning. However PTA does not precisely predict audibility of input stimuli with the current procedure because non-linear intensity gains are common in contemporary HAs (Levitt, 2007). Hence, free-field PTAs were not used in subsequent analyses.

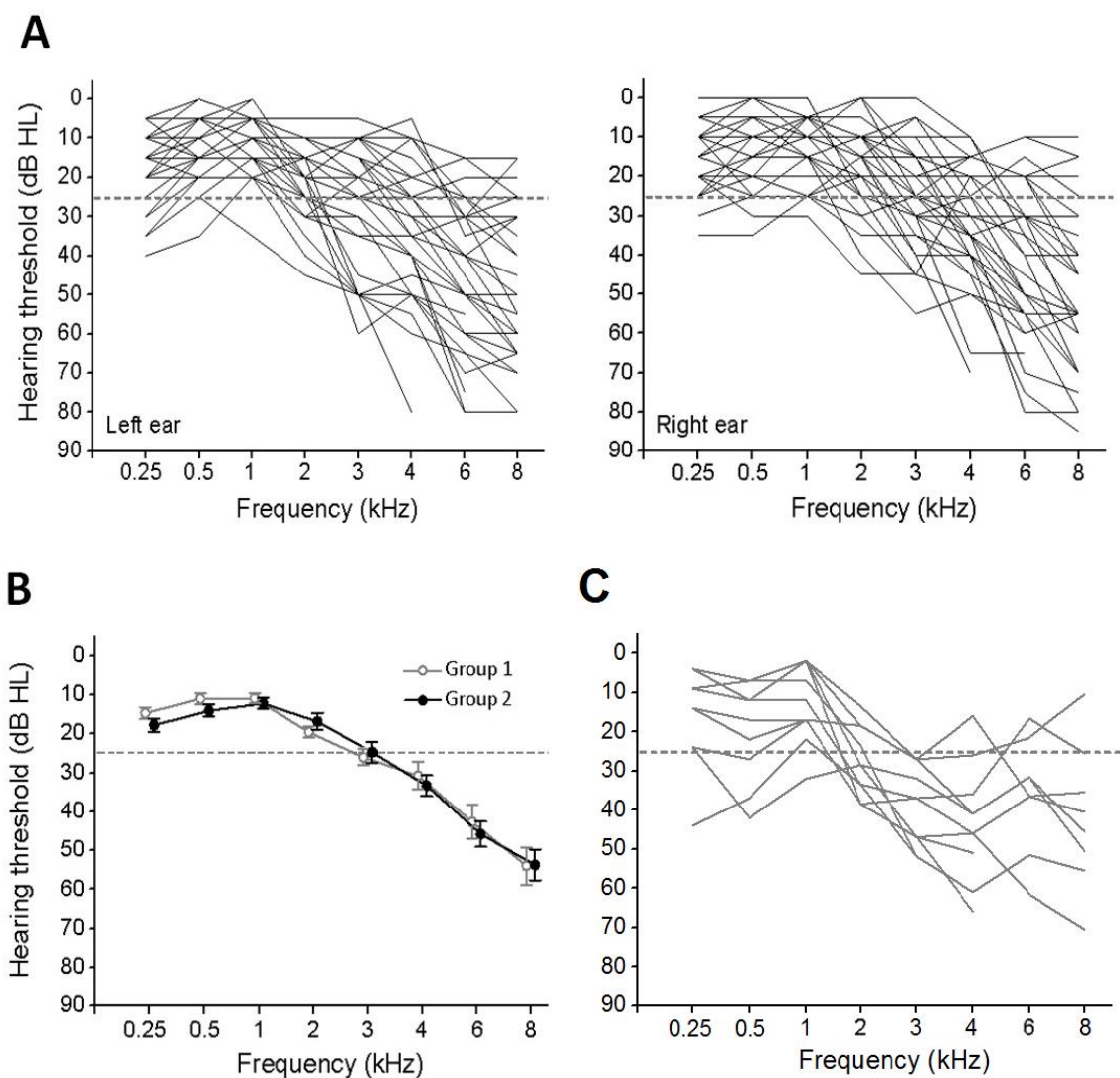**Figure 1.** Pure-tone audiograms (PTAs) of the 35 non-HA participants. **(A)** Individual PTAs (left: left ear; right: right ear) of all non-HA participants. There was one participant's data at 6 kHz (on both ears) and six participants' data at 8 kHz (two at both ears and four at either left

or right ear) that were not plotted, because their hearing thresholds were > 85 dB HL and not measurable at these frequencies. **(B)** PTAs (collapsing the left and right ears) averaged across all non-HA participants (open and filled circles are data from for Group 1 and Group 2, respectively). When points were not measurable, data values were set at the highest possible value (85 dB) when average values and standard errors (SEs) were calculated. Error bars denote the SEs. **(C)** Free-field PTAs for HA users. There were three participants' data at 6 and 8 kHz that were not plotted, because their hearing thresholds were > 85 dB. The dashed lines indicate the normal hearing threshold level (25 dB HL).

Participants were randomly assigned to one of two groups, with the proviso that each group had six HA participants. The syllable stimuli were presented at different SNRs to the two participant groups when neural measurements were made (see **TABLE I**). There were 23 participants (59 to 76 years old; mean $\pm$ SE = 68.22 $\pm$ 1.01; 8 males) in Group 1 and 24 participants (53 to 75 years old; mean $\pm$ SE = 67.33 $\pm$ 1.03; 9 males) in Group 2. There were no significant differences in ages (p = 0.542) nor hearing thresholds in non-HA participants (averaged across 0.25 to 4 kHz and across both ears) (p = 0.693) between the two groups. The neural data and neural-behavioral relationship of the two groups were analyzed separately.

TABLE I. Background noise types and SNR levels for Group 1 and 2 participants during the EEG recordings. During EEG recording, the stimuli in quiet background (NoN) were presented at the same level to both groups (all 47 participants), but the SNRs for noisy backgrounds (SpN and BbN) differed in level across Groups 1 and 2 (7 dB and -1dB

respectively).

SNR	+∞ dB	7 dB	-1 dB
	(NoN)	(SpN and BbN)	(SpN and BbN)
		Group 1	Group 2
Participants	All 47 participants	(23 participants)	(24 participants)

B. Speech-in-noise (SPIN) perception task

Speech Reception Thresholds (SRTs) (SNR level for 50% word correct, as in Plomp and Mimpen, 1979) were measured when participants recognized words within sentences under two different types of background noise, steady-state speech-shaped noise (SpN) and 16-speaker babble noise (BbN). Target sentences were pre-recorded BKB sentences (Bench et al., 1979) spoken by a male British English speaker whose F_0 ranged between 80 and 200 Hz. Each sentence was meaningful and included three key words (content words), e.g., for 'The clown has a funny face', the key words were 'clown', 'funny' and 'face'. The BKB sentences were used rather than more complex ones such as IEEE sentences (Rothauser et al., 1969) in order to minimize effects of individual variability in working memory on speech perception performance. The BbN was a mixture of 16 different utterances spoken by 16 male British English speakers who had similar F_0s and voice quality to the target speaker. The SpN was created by randomizing the phases in the long-term spectrum of BbN before the spectrum was transformed back to the time domain. Consequently SpN had the same long-term spectral power distribution as BbN and relatively stable temporal properties (Rosen et al., 2013). The

background noise always started about 500 ms prior to the onset of the target sentence and continued until the sentence ended.

Participants were seated in a sound-treated booth facing a Fostex 6301B loudspeaker (Canford Group Ltd, UK) at 0 degree azimuth positioned one-meter away from the their ears. The sentences were all presented at an intensity of 70 dB SPL at the 1 meter distance. Participants listened to different sets of 30 sentences in SpN and in BbN and repeated orally as many words as they could after each sentence was heard. Eight practice sentences were given prior to each noise condition. The sentences were delivered by a Matlab 2010a (Mathworks, USA) program and the experimenter scored participants' responses using a graphical user interface (GUI) not visible to participants. The SNR was varied adaptively to track the threshold for 50% words correct. The first sentence was always presented at a high SNR (8 dB and 10 dB for sentences in SpN and BbN, respectively), and SNR was decreased by 4 dB on successive sentences until fewer than two words (i.e., < 50% correct) were recognized correctly. Then the SNR was increased or decreased by 2 dB on the next sentence when fewer (< 50%) or more than (> 50%) two words were recognized correctly. Participants were not given feedback about their performances. SRT was obtained by linear interpolation when two SNRs with the minimal distance (i.e., 2 dB) had word scores that were higher and lower than 50%.

C. EEG experiment

1. Acoustic stimulus

The stimulus was a 120-ms-long /i/ syllable (**Fig 2A**) with a falling F_0 contour (160 to

110 Hz) (**Fig 2B**). This corresponded to the 4th lexical tone in Mandarin. It was pronounced by a male native Mandarin speaker. Previous studies have used flat F_0 contours (e.g., Aiken and Picton, 2008; Song et al., 2010). The syllable with a natural falling F_0 contour controls for exposure, as such contours are not encountered in English monosyllables (Krishnan et al., 2005; Wong et al., 2007). Moreover, the F_0 of the stimulus corresponds in range and direction (generally falling) to the F_0s of the target speakers in the BKB sentences employed in the SPIN task. The syllable had a relatively constant amplitude envelope except that 5 ms rising and falling cosine windows were applied at onset and offset to avoid transients. The formants occurred in frequency regions where the harmonics are and are not resolved (F1, F2 and F3 frequencies were approximately 280, 2400 and 3100 Hz, see **Fig 2A**).
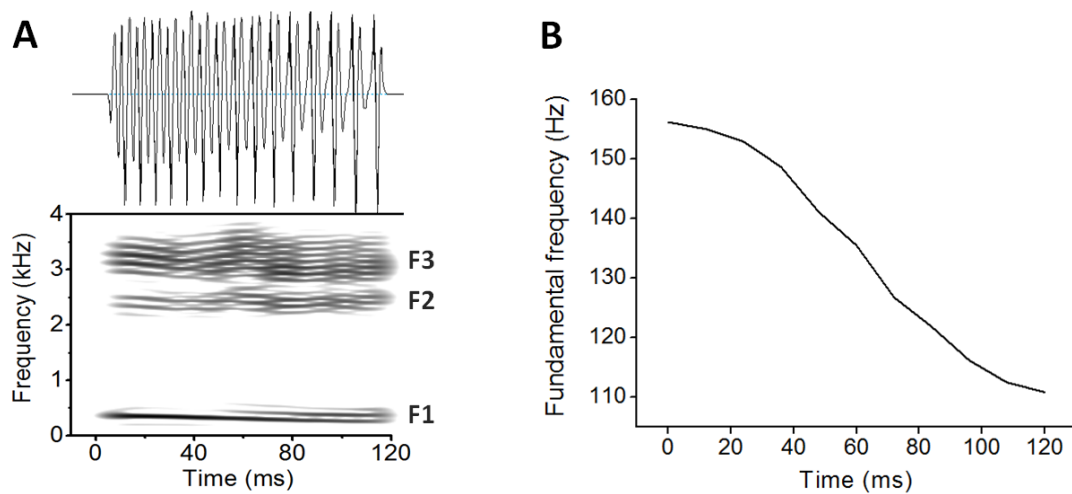

Figure 2. The syllable stimulus /i/ used during EEG recording. **(A)** The temporal waveform (top) and spectrogram (bottom) of the stimulus. F1, F2 and F3 frequencies are approximately 280, 2400 and 3100 Hz, respectively. **(B)** F_0 contour of the stimulus, which is a falling tone ranging between 160 and 110 Hz, obtained by autocorrelation. The waveforms, spectrograms and F_0 data were generated using PRAAT (Boersma and Weenink, 2013).

The syllable was presented at the original polarity or negative polarity where the syllable was inverted. The stimuli were presented in three backgrounds: quiet ("no noise" or NoN); steady-state speech-shaped noise (SpN); and 16-speaker babble noise (BbN). The last two corresponded to the types of noise in the SPIN task. The stimuli for NoN were presented at the same level to both groups of participants (see **TABLE I**) but the SNRs for SpN and BbN differed in level across Groups 1 and 2 (7 dB and -1dB respectively). There were 3200 sweeps at each polarity for each type of noise background (hence 6400 sweeps altogether). The 6400 sweeps were split into 16 segments for each background type and polarities were presented in a random order. Inter-stimulus intervals (ISI) varied randomly between 60 and 120 ms and syllables repetition rate was approximately five syllables per second.

2. EEG recording procedure

EEGs were recorded using an ActiveTwo system (Biosemi, The Netherlands) at a 16384 Hz sampling rate at the Cz (vertex), C3 and C4 position. The bilateral earlobes were used as references. Electrode impedance was always kept below 35 mV. Participants reclined comfortably in an adjustable chair located within an electromagnetic-shielded and sound-treated booth. The stimuli were presented via a Rogers LS3/5A loudspeaker (Falcon Acoustics, UK) positioned one meter from their ears at 0 degree horizontal azimuth directly in front of participants. The stimulus level (measured across time including inter-stimulus intervals) was 74.5 dB SPL at the 1 meter position before background noise was added in all the conditions. The stimulus levels after adding noise (SpN or BbN) were 75.8 dB in Group 1 (SNR of 7 dB) and 79.5 dB in Group 2 (SNR of -1 dB). Participants were instructed to close

their eyes, relax and keep as still as possible to avoid movement artifacts. They were

instructed not to attend actively to the stimuli and they were allowed to fall asleep to

minimize discomfort because of protracted exposure to the repeated sounds. The entire

in-booth test period lasted for approximately 70 minutes. They were monitored by the

webcam throughout the test and no significant changes in head or body positions were

observed. Algorithms that classified the sleep status in different periods were applied to the

EEG data. These algorithms separated wakeful (conscious) from sleep (unconscious) states.

Only the data from wakeful states were used in the current analyses (see *Arousal states*).

3. Arousal states

Wilf et al. (2016) reported cortical, but not subcortical, suppression whereas Portas et al.

(2000) found suppression at both cortical and subcortical levels. Whilst this discrepancy needs

resolving, both studios confirm that cortical activity is suppressed in sleep states. EEG data

were only used from wakeful periods in the analyses in the present paper because significant

correlations between SPIN performances and EEG parameters only occurred during wakeful

periods (see *Results*). The detailed effects of sleep status on speech-evoked EEG responses is

to be reported in the near future.

To separate wakeful from sleep periods, sleep spindles at the EEG sigma frequency (12

~ 16 Hz) were located. These provide signatures about sleep (Warby et al., 2014) in Martin et

al.'s (2013) method. Here EEGs at Cz were bandpass filtered into alpha (8 ~ 11 Hz), sigma

(12 ~ 16 Hz) and beta (17 ~ 20 Hz) bands using a 2nd-order zero-phase Butterworth filter.

The filtered signals were then divided into 250-ms-long successive segments. A spindle was

labeled when all three of the following conditions were met: 1) sigma root-mean-square

(RMS) voltage in a given segment exceeded the threshold of the 95th percentile of the sigma RMS of all segments; 2) sigma RMS was higher than both alpha and beta RMS in the current segment; 3) two successive segments met both 1) and 2). 1) and 2) were invoked because sigma-band spectral domination is the major characteristic of sleep spindles (Martin et al., 2013; Warby et al., 2014). Extending requirements across two segments was included because sleep spindles usually last for at least 500 ms (De Gennaro and Ferrara, 2003).

The entire EEG recording was divided into 192 21-second-long epochs. Epochs were classified into three categories; wakefulness, sleep, and transition between wakefulness and sleep. Sleep epochs were those that contained at least one sleep spindle. Wakeful epochs were those that contained no spindles and were not adjacent to an epoch with a speech spindle. The remaining epochs were designated as transition epochs. Participants also gave a subjective rating of how much they slept after the test was completed (scale points from 1 to 7, each of which had a textual description). There was a significant correlation between the subjective ratings and percentage of epochs classified algorithmically as 'sleep' ($p = 0.002$). This validated the classification algorithm.

4. Frequency Following Responses (FFRs)

FFRs were obtained from Cz (Skoe and Kraus, 2010). EEG signals were re-referenced to the average at the bilateral earlobes, bandpass filtered (70 ~ 4000 Hz) using a 2nd-order zero-phase Butterworth filter, and baseline corrected (based on the 50 ms pre-stimulus period in each sweep). To avoid movement artifacts, sweeps that had amplitudes exceeding ± 25 µV were rejected (c.f., Song, Nicol and Kraus, 2011; Schoof and Rosen, 2016). FFRs with the original (FFR_{orig}) and negative (FFR_{neg}) polarities were obtained by grand-averaging across

sweeps with the respective polarities. FFRs that represent ENV (FFR$_{ENV}$) and TFS (FFR$_{TFS}$) were obtained by adding and subtracting FFR$_{orig}$ and FFR$_{neg}$, respectively (Aiken and Picton, 2008). Three FFR magnitudes were calculated: FFR$_{ENV\_F0}$; FFR$_{TFS\_H2}$ and FFR$_{TFS\_F2F3}$. The FFR$_{ENV\_F0}$ magnitude represents the neural magnitude of FFR$_{ENV}$ along the F0 trajectory of the stimulus syllable (**Fig 3A**). Magnitudes of FFR$_{TFS\_H2}$ and FFR$_{TFS\_F2F3}$ represent the neural magnitudes of FFR$_{TFS}$ along the resolved harmonics H2 (2nd harmonics at 220 ~ 320 Hz representing F1) and unresolved F2 and F3 trajectories of the syllable respectively (**Fig 3B**). Magnitudes were calculated as follows.
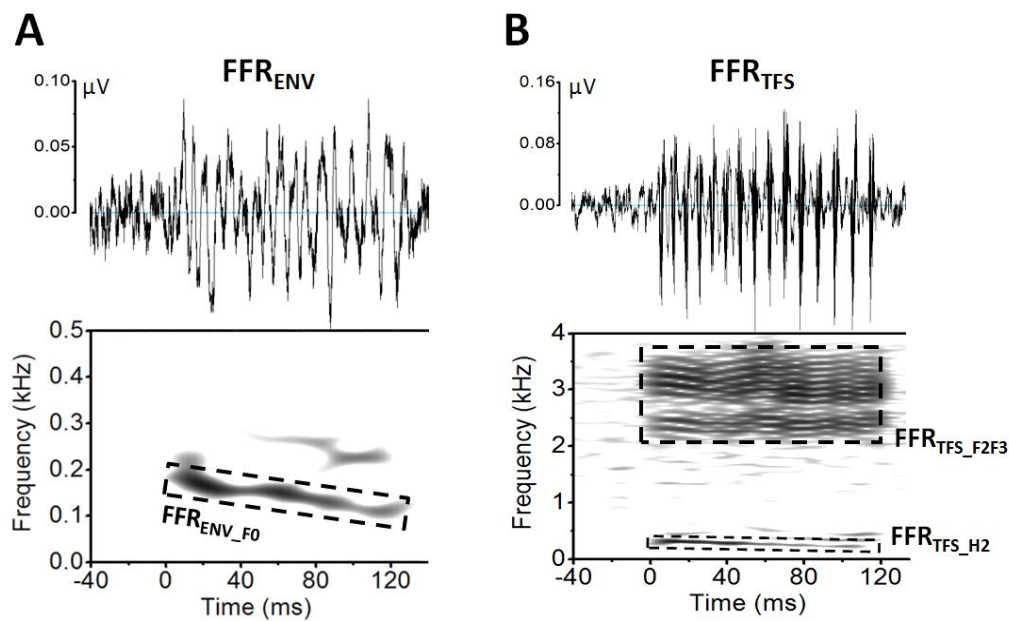

Figure 3. FFR$_{ENV}$ (**A**) and FFR$_{TFS}$ (**B**) of waveforms (top) and spectrograms (bottom) averaged across 35 selected participants during wakeful periods in the quiet background (NoN) condition. The 35 participants were those whose numbers of sweeps were no fewer than 1,450 after artifact rejection during wakeful periods in NoN. For each participant, the waveforms were obtained based on randomly selecting wakeful epochs which contained 1,450 to 1,550 sweeps (see *Normalization of sweep numbers*). FFR$_{ENV\_F0}$ (at F0 range between 160 ~ 110 Hz)

(**A**), FFR_{TFS_H2} (at H2 range between 220 ~ 320 Hz) and FFR_{TFS_F2F3} (at F2-F3 range between 2000 ~ 4000 Hz) (**B**) are indicated by the dashed boxes. X-axes give time in ms and '0' is the stimulus onset.

First, $ENV\text{-}F_0$, H2, F2 and F3 trajectories of the stimulus syllable were calculated. To obtain the $ENV\text{-}F_0$ trajectory, a set of 40 ms sliding windows (1-ms per step) was applied to the Hilbert envelope (Hilbert, 1912) of the syllable. Each 40 ms segment was Hanning-windowed, zero-padded to 1 second (to achieve 1 Hz frequency resolution) and Fourier-transformed. The frequency with the highest magnitude in the Fourier spectrum between 110 and 160 Hz (F_0 range of the syllable) was chosen as the F_0 value at each step. H2, F2 and F3 trajectories were obtained in the same way, except that: 1) the sliding windows were applied to the syllable rather than the Hilbert envelope of the syllable; 2) the H2 value at each step was selected in the range of H2 (220 ~ 320 Hz); 3) instead of zero-padding and applying a Fourier transform, F2 and F3 values at each step were chosen based on the magnitudes of the spectral profile (via cepstral smoothing, Proakis and Manolakis, 2007) in the F2 and F3 ranges respectively (2200 ~ 2600 Hz and 2800 ~ 3500 Hz).

Second, to calculate the FFR_{ENV_F0} magnitude, the set of 40 ms sliding windows (1 ms per step) was applied to FFR_{ENV}. At each step, the 40 ms segment was Hanning-windowed, zero-padded to 1 second and Fourier-transformed. The mean log-magnitude ($\log_e[power]$) was measured across a 20 Hz bandwidth centered at the frequency of the $ENV\text{-}F_0$ trajectory of the syllable stimulus at that step. The magnitudes were then averaged across all steps covering the

16

120 ms syllable length. In addition, as FFR_{ENV} occurs at the auditory brainstem after a neural transmission time of between 5 and 10 ms (Chandrasekaran and Kraus, 2010; Skoe and Kraus, 2010), FFR_{ENV} should have a delay of 8 to 13 ms relative to the syllable onset (allowing an additional 3 ms for air-transmission from loudspeaker to cochlea). The final FFR_{ENV_F0} magnitude was the maximum magnitude for a time lag between 8 and 13 ms (1 ms steps). Magnitudes of FFR_{TFS_H2} and FFR_{TFS_F2F3} were measured in the same way as FFR_{ENV_F0}, except that: 1) the procedure was applied on the FFR_{TFS} along the H2 (for FFR_{TFS_H2}) and the F2 and F3 (for FFR_{TFS_F2F3}) trajectories of the syllable stimulus; 2) for FFR_{TFS_F2F3}, instead of zero-padding and applying a Fourier transform, the magnitude at each step was the summed magnitude of the spectral profile (via cepstral smoothing) across a 150 Hz and 300 Hz bandwidth respectively centered at F2 and F3 of the syllable at that step; 3) as FFR_{TFS} occur at early stages of auditory processing in the periphery (Aiken and Picton, 2008), the neural transmission time for FFR_{TFS_H2} and FFR_{TFS_F2F3} was set to 5 ms at maximum, resulting in time lags between 3 and 8 ms with the air-transmission time included.

5. Cortical responses

Electrodes C3 and C4 were used to represent activity in the auditory cortex (Carpenter and Shahin, 2013; Noguchi, Fujiwara and Hamano, 2015). Theta-band (4 ~ 6 Hz, corresponding to stimulus repetition rate of ~ 5 syllables per second) phase-locking values (PLVs) were measured. EEGs were decimated to 1024 Hz, re-referenced to the average of bilateral earlobes and bandpass filtered (4 ~ 6 Hz) using a 2nd-order zero-phase Butterworth filter. Sweeps where amplitudes exceeded ± 15 µV on either electrode were rejected. Due to the relatively narrow frequency range, the theta-band signal normally does not have

excessively high amplitude and > 90% of the sweeps were retained in almost all participants (except for one participant who retained > 80%). PLV time series ($PLV(t)$) were calculated using the following formula (Morillon et al., 2012):

$$PLV(t) = \left| \frac{1}{n} \sum_{i=1}^{n} e^{j\phi_i(t)} \right|$$

where n denotes the total number of sweeps, $\phi_i(t)$ denotes the Hilbert phase series of the filtered EEG of the ith sweep time-locked to the syllable onset. Hilbert phase was used as it locks to stimuli even when EEG amplitude variation occurs because of eye closure and relaxation (Thatcher, 2012). $PLV(t)$ were then averaged across the 120 ms syllable length. As neural transmission from cochlea to auditory cortex takes 10 to 15 ms in primates (Lakatos et al., 2007), the final PLV was taken as the maximum among values with time lags between 13 and 18 ms (1 ms steps) with the air-transmission time included.

Measurements of all EEG parameters (FFRs and PLV) were obtained using Matlab 2010a.

D. Statistical analyses

1. Within-subject ANOVAs and linear regression analyses

Within-subject Repeated Measures ANOVAs were conducted for the four EEG parameters (i.e., FFR_{ENV_F0}, FFR_{TFS_H2}, FFR_{TFS_F2F3} and PLV) across the three background types (i.e., NoN, SpN and BbN) for Groups 1 and 2, respectively. Post-hoc within-subject pairwise t-tests were conducted if the main effects of background type (with Greenhouse-Geisser correction) were significant in the ANOVA.

Linear regression analyses were conducted for each group of participants. In each group, the analyses were conducted for all participants (the 'HA + non-HA' analysis) and for non-HA participants only (the 'non-HA only' analysis). In the 'HA + non-HA' analysis, SRT in SpN was predicted by the four EEG parameters (FFR_{ENV_F0}, FFR_{TFS_H2}, FFR_{TFS_F2F3} and PLV) obtained from the EEG recording whilst syllables were presented in SpN. Similarly, SRT in BbN was predicted by the four EEG parameters obtained from the EEG recording whilst syllables were presented in BbN. A second regression analysis was further performed that excluded HA users (the 'non-HA only' analysis) to assess the effect of neural responses after controlling for audibility and hearing loss, as stimulus audibility information was not available for HA users in the present study. PTA (averaged across 0.25 ~ 4 kHz and across left and right ears) was included as a predictor in addition to the EEG parameters to assess the effect of neural responses. A non-zero intercept was always included as an IV in each regression equation. Best-Subsets Regression was used, which chose predictors that generated the lowest Akaike information criterion (AIC). This provided the optimal model with best goodness of fit and the least chance of overfitting (Burnham and Anderson 2003). Multicollinearity was diagnosed if the variance inflation factors (VIFs) of any predictor were above 10 or if the condition index was above 30 (Hair. et al., 2010). The assumption that residuals were normally-distributed was violated if the Shapiro-Wilk test on the residuals had a *p*-value below 0.05 (Shapiro and Wilk, 1965). A subset was discarded when there was multicollinearity or normality of residuals was violated. Furthermore, outliers which had Cook's distances above 1 (Cook, 1979) were removed.

2. Normalization of sweep numbers

FFRs have small amplitudes, and averages of large numbers of stimulus repetitions are needed for reliable responses (Skoe and Kraus, 2010). Around 1,500 sweeps are needed after artifact rejection (c.f., Dajani et al., 2005; Wong et al., 2007). Participants' data on any particular background type were not included in subsequent analyses if the number of retained wakeful sweeps was below 1,450 for that background type.

Furthermore, since the current EEG parameters are based on phase-locked activities whose magnitudes are sensitive to the number of sweeps (Aviyente et al., 2011), problems can arise during statistical analyses if numbers of sweeps differ significantly across participants and background types. To reduce the risk of this happening, the number of sweeps was normalized to around 1,500 for each background type and participant. The normalization was conducted by randomly selecting the wakeful epochs which contained 1,450 to 1,550 artifact-free sweeps (summed across both polarities). The EEG parameters were then obtained from the selected epochs. This random selection procedure was repeated 100 times for each background type, giving 100 estimates of each EEG parameter. Measures averaged over the 100 estimates were used in the final statistical analyses. The process thus ensured that measurements of the EEG parameters were always based on around 1,500 sweeps regardless of the rejection rate. The normalization procedures for FFRs and PLV were conducted separately.

3. Controlling for background EEG noise in FFRs

As FFR magnitudes are small, their robustness can be influenced by the background EEG noise levels across participants (Skoe and Kraus, 2010). Therefore, the magnitudes of background EEG noises were obtained, and it was confirmed that magnitudes of FFR

parameters were statistically above the noise level. We also tested whether EEG noise influenced the relationship between FFRs and the behavioral data by conducting partial correlation analyses between FFR parameters and SRTs after noise magnitudes were controlled for. The 50 ms pre-stimulus period was used as the background period. A set of 40 ms sliding windows (1 ms steps) was applied on the background period as described in the procedures for calculating FFR_{ENV_F0}, FFR_{TFS_H2} and FFR_{TFS_F2F3} magnitudes. Each 40 ms segment was Hanning-windowed, zero-padded to 1 second and Fourier-transformed. For $FFR_{ENV\_F0\_Background}$ and $FFR_{TFS\_H2\_Background}$, magnitude at each step was the average spectral log-magnitude of FFR_{ENV} across 110 and ~ 160 Hz (range of F_0) and that of FFR_{TFS} across 220 ~ 320 Hz (range of H2), respectively. For $FFR_{TFS\_F2F3\_Background}$, instead of zero-padding and applying the Fourier transform, its magnitude at each step was calculated as the summed magnitude of the FFR_{TFS} spectral profile (obtained via cepstral smoothing) across 2325 ~ 2475 Hz (150 Hz bandwidth centered at 2400 Hz (F2)) and 2950 ~ 3250 Hz (300 Hz bandwidth centered at 3100 Hz (F3)). The final magnitudes were the averages across all steps of the background period. The procedure for normalization of sweep numbers for the noise magnitudes was the same as that done when the EEG parameters were calculated.

ANOVAs, t-tests, linear regressions and partial correlations were all conducted using SPSS 13.0 (SPSS Inc., USA).

III. RESULTS

A. Behavioral results

Fig 4 shows the SRT results across all 47 participants. Consistent with Rosen et al. (2013), SRT was significantly lower (i.e., better speech performance) in SpN (mean ± SE, -4.13 ± 0.21 dB) than in BbN (mean ± SE, 0.54 ± 0.22 dB) (two-tailed *t*-tests, $t_{(46)}$ = -30.127, $p < 10^{-31}$).
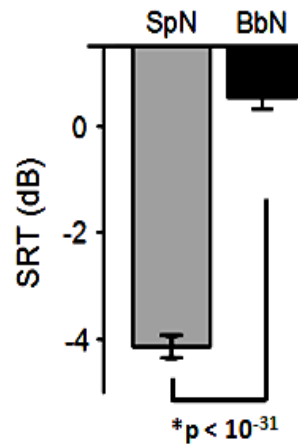

Figure 4. SRT results under SpN (grey) and BbN (black) across all 47 participants. The asterisk denotes $p < 0.05$. Error bars denote SEs.

B. Robustness of FFRs

First, the magnitudes of the three FFR parameters (FFR_{ENV_F0}, FFR_{TFS_H2}, FFR_{TFS_F2F3}) were compared with the magnitudes of their respective EEG background noise ($FFR_{ENV\_F0\_Background}$, $FFR_{TFS\_H2\_Background}$ and $FFR_{TFS\_F2F3\_Background}$). The magnitudes were significantly higher for all three FFR parameters than their respective EEG background noise for all three background types (NoN, SpN and BbN) in both groups of participants (two-tailed *t*-tests, all $p < 0.001$). This confirmed the robustness of FFRs at the group level.

C. Neural results across acoustic background types

Next, within-subjects Repeated Measures ANOVAs were conducted for the four EEG parameters (three FFR parameters and PLV) across the background types separately for Groups 1 (SNR of 7 dB) and 2 (SNR of -1 dB) (**Fig 5**). Only data from participants whose numbers of sweeps were above 1,450 for all three background types during wakeful periods were used. This resulted in 15 and 17 participants being retained in Groups 1 and 2, respectively, with 4 HA users in each group. Age (HA plus non-HA participants; Group 1 vs. 2: $t_{(30)} = 0.666$, $p = 0.511$) and audiometric threshold in non-HA participants (averaged across frequencies from 0.25 to 4 kHz and across both ears; Group 1 vs. 2: $t_{(22)} = -0.161$, $p = 0.873$) did not differ between the two groups. Post-hoc pairwise t-tests (two-tailed, all with Bonferroni corrections) were conducted when a significant main effect (with Greenhouse-Geisser correction) of background type occurred in each analysis. A significant main effect was found for FFR_{ENV_F0} in Group 2 ($F_{(2, 14)} = 15.615$; $p < 0.001$), with a significantly higher FFR_{ENV_F0} magnitude in quiet background (NoN) than in either noise backgrounds ($FFR_{ENV\_F0\_NoN} > FFR_{ENV\_F0\_SpN}$, $p = 0.001$; $FFR_{ENV\_F0\_NoN} > FFR_{ENV\_F0\_BbN}$, $p = 0.001$). The difference in FFR_{ENV_F0} magnitude between SpN and BbN was not significant ($p = 1.000$). Significant main effects of background type were found for PLV in both groups (Group 1: $F_{(2, 12)} = 15.112$, $p < 0.001$; Group 2: $F_{(2, 14)} = 19.548$, $p < 10^{-4}$). Post-hoc comparisons showed that, for Group 1, PLV was significantly higher in NoN and SpN than in BbN ($PLV_{NoN} > PLV_{BbN}$, $p < 0.001$; $PLV_{SpN} > PLV_{BbN}$, $p < 0.001$); for Group 2, PLV was significantly higher in NoN than in either noise background ($PLV_{NoN} > PLV_{SpN}$, $p = 0.003$; $PLV_{NoN} > PLV_{BbN}$, $p < 0.001$), and PLV was significantly higher in SpN than in BbN

$(\text{PLV}_{\text{SpN}} > \text{PLV}_{\text{BbN}}, p = 0.023)$.
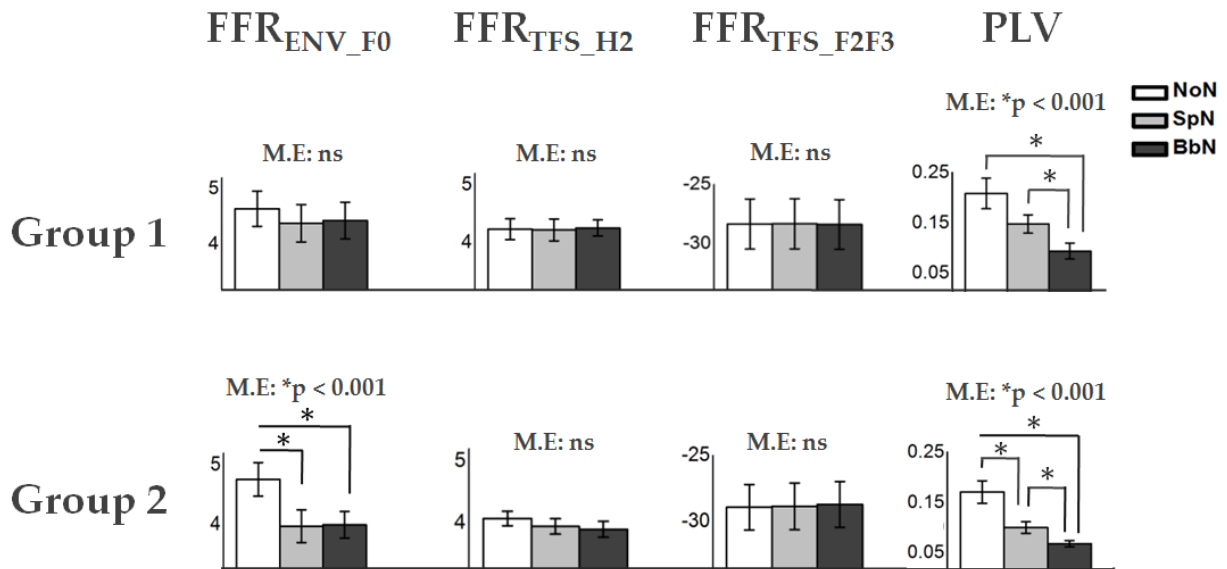
Figure 5. The four EEG parameters ($\text{FFR}_{\text{ENV\_F0}}$, $\text{FFR}_{\text{TFS\_H2}}$, $\text{FFR}_{\text{TFS\_F2F3}}$ and PLV) under three background types (NoN, SpN and BbN). Sections at the top and bottom give data from Groups 1 and 2, respectively. For FFRs, y-axes denote spectral log-magnitude (in dB). M.E.: main effects. ns: non-significant. Asterisks for the post-hoc pairwise t-tests denote $p < 0.05$. Error bars denote SEs.

D. Neural-behavioral relationship

Linear regression was used to evaluate neural-behavioral relationships. For each group of participants, the four EEG measures were used to predict SRT in SpN/BbN ('HA + non-HA'). PTA was also included as a predictor in a second analysis that excluded the HA users ('non-HA only'). According to the data-inclusion criterion (number of sweeps $\geq 1,450$ during wakeful periods), 18 (6) and 20 (6) participants were retained respectively in SpN and

BbN for Group 1, and 19 (5) and 21 (5) participants were retained respectively in SpN and BbN for Group 2 (the number in brackets being the number of HA users).

Best-Subsets Regression results are summarized in **TABLE II** ('HA + non-HA') and **TABLE III** ('non-HA only'). For both 'HA + non-HA' and 'non-HA only' conditions, significant correlations between SRTs and EEG parameters were only found in Group 2 (SNR of -1 dB during EEG recording). Specifically, for Group 2; i) in the 'HA + non-HA' condition, PLV correlated negatively with SRT in SpN ($t_{(17)}$ = -2.265, p = 0.037, see **TABLE II** and **Fig 6A, B**), whilst FFR_{TFS_H2} magnitude correlated negatively with SRT in BbN ($t_{(19)}$ = -2.576, p = 0.019, see **TABLE II** and **Fig 6C, D**); ii) in the 'non-HA only' condition, after controlling for PTA, FFR_{ENV_F0} magnitude and PLV correlated negatively with SRT in SpN (FFR_{ENV_F0}: $t_{(10)}$ = -2.246, p = 0.049; PLV: $t_{(10)}$ = -2.450, p = 0.034; see **TABLE III**), whilst FFR_{TFS_H2} magnitude and PLV correlated negatively with SRT in BbN (FFR_{TFS_H2}: $t_{(12)}$ = -3.009, p = 0.011; PLV: $t_{(12)}$ = -8.256, $p < 10^{-5}$; see **TABLE III**).

TABLE II. Results of Best-Subsets linear regression for 'HA + non-HA' participants in Groups 1 and 2. SRTs (in SpN/BbN) were predicted by the four EEG parameters. *N* denotes the number of participants based on the data inclusion criterion that the number of sweeps $\geq$ 1,450 during wakeful periods; numbers shown in brackets denote the number of HA participants. R_a^2 denotes the adjusted R-square of the models. Significant p values are in bold and * = significant at $p < 0.05$.

Group	N	R_a^2	Noise type	Dependent variable	Predictor(s)	T-value(s)	p-value(s)

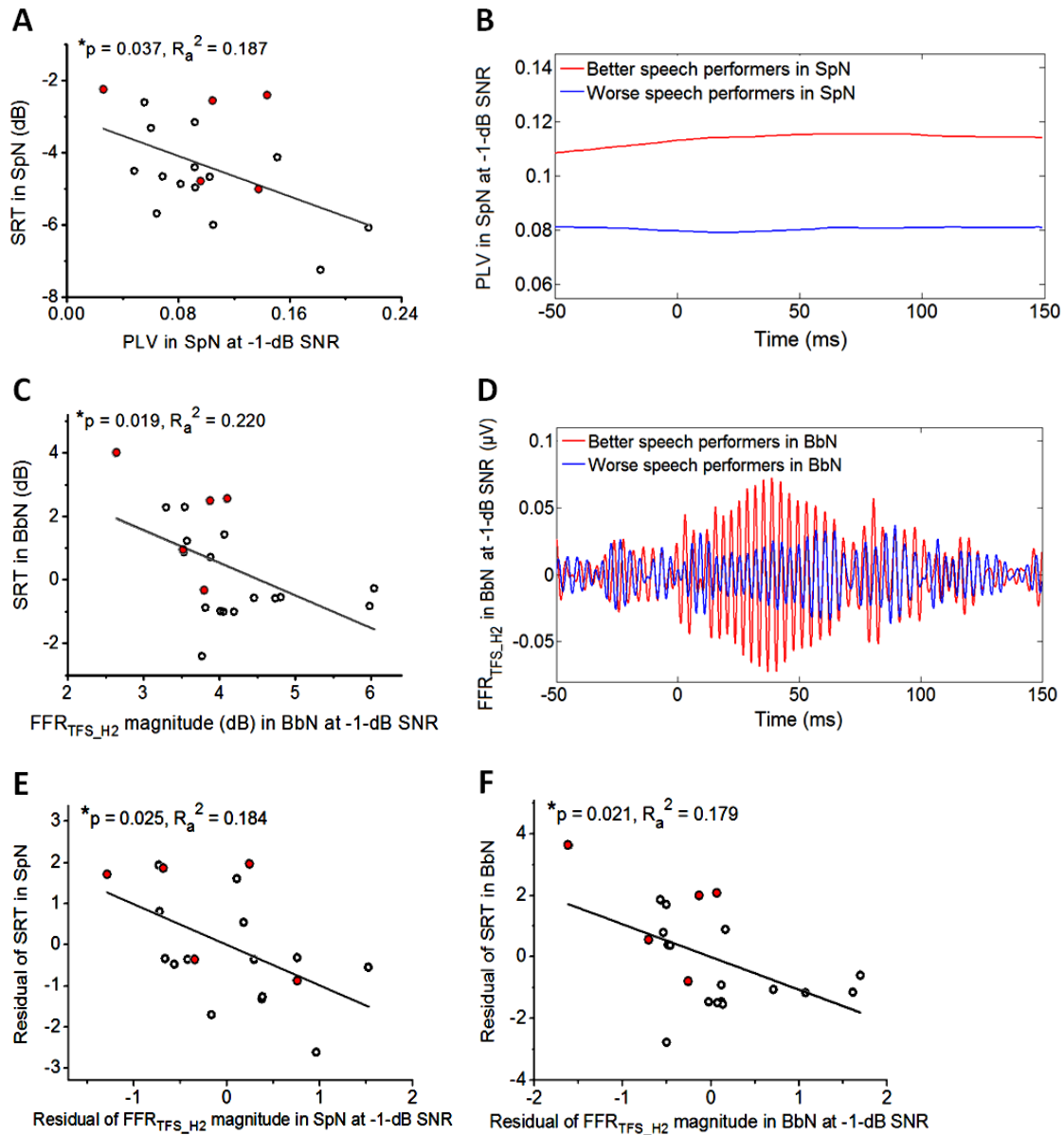|							
Group 1	18 (6)	0.010	SpN	SRT	FFR$_{TFS\_H2}$	-1.086	0.293
	20 (6)	0.015	BbN	SRT	FFR$_{TFS\_H2}$	-1.134	0.272
Group 2	19 (5)	0.187	SpN	SRT	PLV	-2.265	**0.037***
	21 (5)	0.220	BbN	SRT	FFR$_{TFS\_H2}$	-2.576	**0.019***

Figure 6. Neural-behavioral relation. Red dots represent HA participants. *p* values and R_a^2 are given. Asterisks denote $p < 0.05$. Sections **A** and **C** give significant correlations between theta PLV and SRT in SpN (**A**) and between FFR$_{TFS\_H2}$ magnitude and SRT in BbN (**C**) for Group 2

according to the linear regression analyses (see **TABLE II**). Sections **B** and **D** give time series

of theta PLV in SpN (**B**) and FFR$_{TFS\_H2}$ (FFR$_{TFS}$ bandpass filtered at H2 range of 220 ~ 320

Hz) in BbN (**D**) for Group 2. Red and blue lines represent the averages of better (SRTs lower

than the group median) and worse (SRTs higher than the group median) SPIN performers,

respectively. The time series were obtained from the sweeps with normalized numbers (1,450

~ 1,550) during the wakeful periods (see Normalization of sweep numbers). X-axes give time

in ms and '0' is the stimulus onset. Sections **E** and **F** give the partial correlations for Group 2

between FFR$_{TFS\_H2}$ magnitudes and SRTs after controlling for FFR$_{TFS\_H2\_Background}$ magnitudes

in SpN (**E**) and BbN (**F**). The FFR$_{TFS\_H2}$ residuals which were obtained from correlations

between the FFR$_{TFS\_H2}$ and FFR$_{TFS\_H2\_Background}$ magnitudes. SRT residuals were obtained from

correlations between SRT and the FFR$_{TFS\_H2\_Background}$ magnitude.

TABLE III. Results of Best-Subsets linear regression for 'non-HA only' participants in

Groups 1 and 2. SRTs (in SpN/BbN) were predicted by the four EEG parameters and PTA. *N*

denotes the number of participants based on the data inclusion criterion that the number of

sweeps $\geq$ 1,450 during wakeful periods. R_a^2 denotes the adjusted R-square of the models.

Significant *p* values ($p < 0.05$) are in bold and ^ = marginal significance ($p < 0.1$), * =

significant at $p < 0.05$ and ** = significant at $p < 0.01$.

Group	N	R_a^2	Noise type	Dependent variable	Predictor(s)	T-value(s)	P-value(s)
Group 1	12	0.114	SpN	SRT	PTA	1.555	0.151
	14	0.246	BbN	SRT	PTA	2.289	**0.041***

Group 2	14	0.538	SpN	SRT		FFR_{ENV_F0}	-2.246	**0.049***
						PLV	-2.450	**0.034***
						PTA	2.062	0.066^
	16	0.882	BbN	SRT		FFR_{TFS_H2}	-3.009	**0.011***
						PLV	-8.256	**$< 10^{-5}$****
						PTA	7.600	**$< 10^{-5}$****

Partial correlation analyses (two-tailed) were conducted next between the FFR parameters (FFR_{ENV_F0}, FFR_{TFS_H2} and FFR_{TFS_F2F3}) and SRTs while controlling for the magnitudes of background EEG noise. This partialled out the effects of EEG noise on the FFR-SRT relation mathematically. Results showed that:

1) For Group 1, there were no significant correlations between any of the FFRs parameters and SRTs in SpN or BbN after the respective EEG noise magnitudes were controlled for in the 'HA + non-HA' or 'non-HA only' condition (all $p > 0.2$).

2) For Group 2:

i) In the 'HA + non-HA' condition, there were no significant correlations between FFR_{ENV_F0}, FFR_{TFS_F2F3} magnitudes and SRTs after $FFR_{ENV\_F0\_Background}$, $FFR_{TFS\_F2F3\_Background}$ magnitudes were controlled for in either SpN or BbN (all $p > 0.7$). There were significant negative correlations between FFR_{TFS_H2} magnitude and SRT in SpN ($t_{(16)} = -2.464$, $p = 0.025$, **Fig 6E**), and between FFR_{TFS_H2} magnitude and SRT in BbN ($t_{(18)} = -2.525$, $p = 0.021$, **Fig 6F**) after $FFR_{TFS\_H2\_Background}$ magnitudes were controlled for. The latter replicated the previous linear regression result on Group 2 (see **TABLE II** and **Fig 6C**).

ii) In the 'non-HA only' condition, there were no significant correlations between any of the FFRs parameters and SRTs in SpN or BbN after the respective EEG noise magnitudes were controlled for (all $p > 0.09$). To further test the influence of EEG background noise on the previous Best-Subset Regression results in 'non-HA only' (see **TABLE III**), magnitudes of $FFR_{ENV\_F0\_Background}$ and $FFR_{TFS\_H2\_Background}$ were used as additional predictors in the Best-Subset equations in SpN and BbN respectively. After adding $FFR_{ENV\_F0\_Background}$ as a predictor in SpN, correlations between FFR_{ENV_F0} and SRT and between PLV and SRT were not significant (FFR_{ENV_F0} and SRT: $t_{(9)} = -1.361$, $p = 0.207$; PLV and SRT: $t_{(9)} = -2.240$, $p = 0.052$; see **TABLE IV**). Thus, it cannot be entirely ruled out that, while controlling for PTA, the significant correlations of FFR_{ENV_F0} and PLV (especially FFR_{ENV_F0}) with SRT in SpN may be due to the influence of EEG background noise. Lack of significance may have occurred because of the small number of participants ($N = 14$) and the increased number of predictors. On the other hand, after adding $FFR_{TFS\_H2\_Background}$ as a predictor in BbN, correlations between FFR_{TFS_H2} and SRT and between PLV and SRT were still significant (FFR_{TRS_H2} and SRT: $t_{(11)} = -2.548$, $p = 0.027$; PLV and SRT: $t_{(11)} = -7.274$, $p < 10^{-4}$; see **TABLE IV**), showing that the significant correlations were not influenced by the EEG background noise.

TABLE IV. Results of including magnitudes of EEG background noise as predictors for Group 2 ('non-HA only') participants based on the results of Best-subset linear regression. This analysis assesses the influence of the EEG background noise on the neural-behavioral relation. $FFR_{ENV\_F0\_Background}$ and $FFR_{TFS\_H2\_Background}$ were added to predict SRT in SpN (top

half of table) and in BbN, (bottom half of table). N denotes the number of participants based on the data inclusion criterion that the number of sweeps $\geq 1{,}450$ during wakeful periods. R_a^2 denotes the adjusted R-square of the models. Significant p values ($p < 0.05$) are in bold and ^ = marginal significance ($p < 0.1$), * = significant at $p < 0.05$ and ** = significant at $p < 0.01$.

Group	N	R_a^2	Noise type	Dependent variable	Predictor(s)	T-value(s)	P-value(s)
Group 2	14	0.487	SpN	SRT	FFR_{ENV_F0}	-1.361	0.207
					PLV	-2.240	0.052^
					PTA	1.948	0.083^
					$FFR_{ENV\_F0\_Background}$	-0.098	0.924
	16	0.871	BbN	SRT	FFR_{TFS_H2}	-2.548	**0.027***
					PLV	-7.274	$< 10^{-4}$**
					PTA	6.170	$< 10^{-4}$**
					$FFR_{TFS\_H2\_Background}$	0.197	0.847

IV. DISCUSSION

The present results support our hypothesis that SPIN perception is associated statistically with phase-locked neural encoding of Slow-rate ENV and TFS, rather than that of F_0-rate ENV, in older adults who presented with different audiometric profiles. Magnitude of FFR_{ENV_F0} that reflects encoding of F_0-rate ENV correlated significantly with behavioral performances in SpN when audiometric threshold was mathematically controlled for. The participants here showed wide individual variability of audiometric thresholds as is typical in aging populations (Gopinath et al., 2009; Humes et al., 2010) and hence should be more

representative of clinical populations than previous studies (Anderson et al., 2011; Presacco et al., 2016; Schoof and Rosen, 2016). Some HA users were also included in the present study and contributed to the current results (the 'HA + non-HA' condition). This extended individual variability in the regression analyses (see **Fig. 6**; HA users are shown as red dots). However, it should be noted that stimulus audibility was not quantified for this group, and the numbers are small. Future work should allow more informative hearing aid-related analyses. Significant correlations between EEG parameters and SRT performance were found in Group 2 but not in Group 1. This indicated that individual variability of SPIN abilities may be better explained using EEG parameters obtained under low SNR, which is relatively challenging for speech perception (-1 dB in Group 2), compared to high SNR at which speech is fairly recognizable (7 dB in Group 1). The neural-behavioral relationship is summarized schematically in **Fig 7** and is discussed further below.

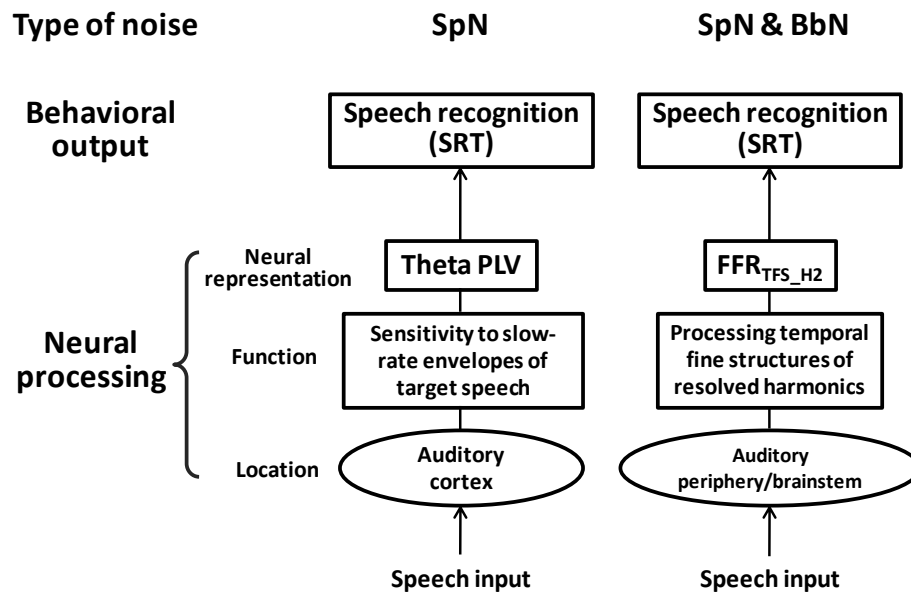| Type of noise	SpN	SpN & BbN
Behavioral output	Speech recognition (SRT)	Speech recognition (SRT)
Neural representation	Theta PLV	FFR$_{TFS\_H2}$
Function	Sensitivity to slow-rate envelopes of target speech	Processing temporal fine structures of resolved harmonics
Location	Auditory cortex	Auditory periphery/brainstem
	Speech input	Speech input

Figure 7. Summary of the neural-behavioral relationship in older adults who varied in audiometric hearing (the 'HA + non-HA' condition). Theta PLV, which reflects cortical

sensitivity to Slow-rate ENV of target speech, was associated with speech performances in SpN. FFR$_{TFS\_H2}$ magnitude, which reflects peripheral/brainstem processing of TFS in the resolved harmonics region, was associated with speech performances for both SpN and BbN. Also see **TABLE II** and **Figure 6**.

A. Cortical sensitivity to Slow-rate ENV

The SRT result showed that speech perception was significantly better in SpN than in BbN (**Fig 4**), as reported in previous work (Rosen et al., 2013). Correspondingly, the neural results showed that theta-band PLV time-locked to the syllable onset was significantly higher in SpN than in BbN for both groups of participants and there were no significant differences for other EEG parameters between the two noise types (**Fig 5**). Also, PLV was higher in quiet than in noisy backgrounds (particularly in Group 2, see **Fig 5**). Such results may explain how speech perception deteriorates in noisy environments and why it is better in SpN than in BbN. Previous research has reported that the theta PLV reflects neural sensitivity to the Slow-rate ENV in the auditory cortex (Luo and Poeppel, 2007; Howard and Poeppel, 2010) and that Slow-rate ENV is critical for speech perception (Drullman et al., 1994; Shannon et al., 1995; Arai et al., 1999). Furthermore, theta PLV obtained from scalp-EEG time-locked to input acoustic stimuli could reflect neural firing in the auditory cortex when decoding stimuli (Ng, Logothetis and Kayser, 2013). SpN is a steady-state energetic masker (EM) with no linguistic information, whilst BbN is an EM with envelope modulations more similar to speech compared to SpN. Also, BbN is an informational masker (IM) that can lead to phonetic

interference (Rosen et al., 2013). BbN had no lexical interference in the present study because of the high number of speakers (16) in the babble (Hoen et al., 2007). It is likely that the combined EM and IM properties of BbN caused lower cortical sensitivity to Slow-rate ENV of the target speech, leading to worse SPIN performance than that in SpN. This differential cortical sensitivity may serve as the neural mechanism of SPIN performance which varies as a function of noise type. Moreover, the linear regression results from Group 2 showing that higher theta PLV was associated with better speech perception in SpN (see **TABLE II** and **TABLE III**) and BbN (after controlling for PTA in non-HA participants, see **TABLE III**) further highlighted the important role of cortical sensitivity to Slow-rate ENV during SPIN perception.

B. Neural encoding of TFS in the resolved harmonics region

The neural-behavioral relationship observed in Group 2 showed that higher FFR_{TFS_H2} magnitude, which reflects phase-locked neural encoding of TFS in the resolved harmonics (H2) region, was associated with better speech perception in both SpN and BbN (**TABLE II** and **Fig 6**). FFR_{TFS_H2} magnitude correlated significantly with better speech performances in BbN after PTA was controlled for (**TABLE III and IV**). Previous studies using cochlear implant simulation methods have shown that encoding TFS is critical for SPIN perception (Zeng et al., 2005; Stickney et al., 2007; Eaves et al., 2011). Unprocessed (where TFSs were retained compared to those "processed" without TFS) low-frequency (< 500 Hz) speech components benefited speech perception for both normal-hearing people and cochlear implant users (Turner et al., 2004; Qin and Oxenham, 2006; Carroll, Tiaden, and Zeng, 2011). A harmonic sieve, in which pitch of the target voice is extracted and used to segregate this voice

from background noise, is a possible explanation (Bregman, 1994; Bird and Darwin, 1998; Binns and Culling, 2005). Pitch perception mainly relies on TFS (Smith et al., 2002; Zeng et al., 2004), especially TFS in the low-frequency resolved harmonics region (Oxenham, Micheyl and Keeble, 2009). In the present study, the syllable stimulus used during EEG recording had an F_0 range and contour (falling) similar to those of the target voices in the sentence stimuli during the SPIN perception task. Therefore, better FFR_{TFS_H2} representation reflects better processing of pitch information of the target speech, hence better voice segregation ability during SPIN perception.

C. Neural encoding of F_0-rate ENV

FFR_{ENV_F0} magnitudes were significantly lower in noise (at SNR of -1 dB) than in quiet, as reported in previous studies (Russo et al., 2004; Presacco, Simon and Anderson, 2016). This indicates how background noise affects speech perception. Anderson et al. (2011) showed that FFR_{ENV_F0} magnitude correlated significantly with better SPIN performances in older adults. While only normal-hearing older adults (hearing thresholds $\leq$ 25 dB HL at $\leq$ 4000 Hz) were tested in Anderson et al. (2011), the present study recruited participants who varied in audiometric profiles (including some who wore HAs), which gives a more appropriate audiometric distribution for aging populations (Gopinath et al., 2009; Humes et al., 2010). The present results did not show any significant correlation between FFR_{ENV_F0} magnitude and SPIN performance until PTA was mathematically controlled for in non-HA participants (for SpN in Group 2, see **TABLE III**). This significant correlation could also have resulted from the influence of EEG pre-stimulus magnitudes (**TABLE IV**), although such influence may simply reflected the variation in pre-stimulus brain activities and, hence,

may need to be interpreted with caution. The correlation between FFR_{ENV_F0} magnitude and SPIN performance indicates that phase-locked neural encoding of F_0-rate ENV may be associated with SRT only when participants have similar audiometric thresholds and such association is limited to speech perception in SpN. As FFR_{ENV_F0} magnitude depends on participants' sensitivity to the syllable, with poorer audibility leading to smaller FFR_{ENV_F0} magnitude (Ananthakrishnan et al., 2016), higher degrees of hearing loss should cause smaller FFR_{ENV_F0} magnitude. However, other evidence has shown that hearing loss can lead to greater encoding of F_0-rate ENV compared to normal-hearing in animal studies (Kale and Heinz, 2010; Henry et al., 2014; Zhong et al., 2014) as well as in older human participants (Anderson et al., 2013). It has been suggested that such exaggerated encoding of F_0-rate ENV distracts encoding of salient acoustic cues (e.g., TFS) in noisy environment and hence could be detrimental to SPIN perception (Henry et al., 2014; Zhong et al., 2014). Therefore, the extent of audiometric hearing loss should be controlled carefully when studying the relationship between SPIN performances and neural encoding of F_0-rate ENV in older adults.

D. Clinical implications

Older adults who varied in their audiometric thresholds were tested in the present study. Audiometric impairments in older adults often manifest as high-frequency hearing loss (Humes et al., 2010). However, high-frequency amplification in older adults has limited clinical benefits for speech perception (Turner and Cummings, 1999; Ching et al., 2001; Horwitz, Ahlstrom and Dubno, 2008). People using auditory prostheses, such as hearing aids (HAs) have difficulty in understanding conversations in noise partly because of limited improvements in SNR despite advanced contemporary noise reduction approaches (Bentler,

2005; Géléoc and Holt, 2014; Tremblay and Miller, 2014). Furthermore, even older adults with normal hearing experience difficulties in SPIN perception, especially in multi-speaker backgrounds, compared to normal-hearing young adults (Helfer and Freyman, 2008; Rajan and Cainer, 2008; Schoof and Rosen, 2014). Therefore, as well as developing strategies for restoring peripheral hearing and applying noise reduction approaches, it could be helpful to attend to neural biomarkers in the auditory pathway in order to determine optimal rehabilitation in older adults. In addition to decline in audiometric hearing, older adults have been shown to have diminished auditory processing of TFS that is related to poorer SPIN perception (Lorenzi et al., 2006; Grose and Mamo, 2010; Hopkins et al., 2008; Hopkins and Moore, 2011). This may be related to the loss of neural inhibition in the aging auditory system which is important for phase-locking to TFS information (Anderson et al., 2012). Furthermore, broadened frequency tuning at the auditory periphery due to hearing loss may decrease the phase-locking abilities to TFS causing SPIN perception to deteriorate further (Henry and Heinz, 2012). This echoes the present finding that peripheral/brainstem processing of TFS in the low-frequency resolved harmonics region is critical for SPIN perception. The present findings thus provide neural evidence to support signal processing strategies for cochlear implants that retain TFS information in the low-order harmonics region that render significant benefits during speech perception (Turner et al., 2004; Qin and Oxenham, 2006; Carroll et al., 2011).

It has been proposed that scalp-recorded electrophysiological measures, such as cortical auditory evoked potentials (cAEPs) (Tremblay and Miller, 2014) and speech-evoked FFRs (Dajani, Heffernan and Giguere, 2013; Bellier et al., 2015), could be used to evaluate the

effects of HA fitting. Components of cAEPs (such as the N1-P2 complex) have been reported to be good indicators for assessing HA effects both in quiet and in background noise (Tremblay and Miller, 2014; Kuruvilla-Mathew, Purdy and Welch, 2015). While low-frequency (delta, theta and alpha) PLVs can predict both amplitudes and latencies of N1 and P2 in cAEPs in noisy backgrounds (Koerner and Zhang, 2015), it is plausible that PLV can also serve as a marker that helps to assess the effects of HA fitting on cortical encoding during SPIN perception. Bellier et al. (2015) demonstrated that magnitudes of speech-evoked FFR_{ENV} and FFR_{TFS} can be modulated by amending HA settings which generate different auditory stimulations. FFRs could thus be used as potential peripheral/subcortical markers that would benefit HA users. Therefore, a further clinical implication of the present study is that it could provide cortical and peripheral/subcortical neural markers for evaluating and optimizing the settings in HAs in older adults. Compared to evaluation of these neural markers, evaluation of SPIN performances tends to be a more direct approach to assess HA benefit. However, improvements of SPIN perception based on different HA settings may also be influenced by higher-level cognitive factors, such as working memory capacity (see a review by Lunner, Rudner and Rönnberg, 2009). Evaluation of sensory-neural markers in HAs could thus dissociate its role from individual cognitive factors in assisting SPIN perception. To this end, future work is needed to assess the neural-behavioral relation with further control of cognitive factors during SPIN perception.

Acknowledgements

REFERENCES

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., and Merzenich, M.M. (2001). "Speech comprehension is correlated with temporal response patterns recorded from auditory cortex," Proc. Natl. Acad. Sci. 98, 13367-13372.

Aiken, S. J., and Picton, T. W. (2008). "Envelope and spectral frequency-following responses to vowel sounds," Hear. Res. 245, 35-47.

Ananthakrishnan, S., Krishnan, A., and Bartlett, E. (2016). "Human frequency following response: Neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss," Ear Hear. 37, 91-103.

Anderson, S., Parbery-Clark, A., Yi, H. G., and Kraus, N. (2011). "A neural basis of speech-in-noise perception in older adults," Ear Hear. 32, 750-757.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., Drehobl, S., and Kraus, N. (2013). "Effects of hearing loss on the subcortical representation of speech cues," J. Acoust. Soc. Am. 133, 3030-3038.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (2012). "Aging affects neural precision of speech encoding," J. Neurosci. 32, 14156-14164.

Arai, T., Pavel, M., Hermansky, H., and Avendano, C. (1999). "Syllable intelligibility for temporally filtered LPC cepstral trajectories," J. Acoust. Soc. Am. 105, 2783-2791.

Aviyente, S., Bernat, E. M., Evans, W. S., and Sponheim S. R. (2011). "A phase synchrony

measure for quantifying dynamic functional integration in the brain," Hum. Brain Mapp. 32, 80-93.

Bellier, L., Veuillet, E., Vesson, J. F., Bouchet, P., Caclin, A., and Thai-Van, H. (2015). "Speech Auditory Brainstem Response through hearing aid stimulation," Hear. Res. 325, 49-54.

Bench, J., Kowal, A., and Bamford, J. (1979). "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children," Brit. J. Audiol. 13, 108-112.

Bentler, R. A. (2005). "Effectiveness of directional microphones and noise reduction schemes in hearing aids: A systematic review of the evidence," J. Am. Acad. Audiol. 16, 473–484.

Binns, C., and Culling, J. F. (2005). "The role of fundamental frequency (F0) contours in the perception of speech against interfering speech," J. Acoust. Soc. Am. 117, 2606-2607.

Bird, J., and Darwin, C. J. (1998). "Effects of a difference in fundamental frequency in separating two sentences," in Psychophysical and Physiological Advances in Hearing, edited by A. R. Palmer, A. Rees, A. Q. Summerfield and R. Meddis (Whurr, London), pp. 263-269.

Boersma, P., and Weenink, D. (2013). *Praat: doing phonetics by computer [Computer program] Version 5.3.51,* Retrieved 2 June 2013.

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound* (MIT press).

Burnham, K. P., and Anderson, D. (2003). *Model selection and multi-model inference: A Practical informatio-theoretic approach* (Springer Science and Business Media).

Carpenter, A. L., and Shahin, A. J. (2013). "Development of the N1–P2 auditory evoked

response to amplitude rise time and rate of formant transition of speech sounds," Neurosci. Lett. 544, 56-61.

Carroll, J., Tiaden, S., and Zeng, F. G. (2011) "Fundamental frequency is critical to speech perception in noise in combined acoustic and electric hearing," J. Acoust. Soc. Am. 130, 2054-2062.

Chandrasekaran, B., and Kraus, N. (2010). "The scalp-recorded brainstem response to speech: Neural origins and plasticity," Psychophysiology 47, 236-246.

Ching, T. Y., Dillon, H., Katsch, R., and Byrne, D. (2001). "Maximizing effective audibility in hearing aid fitting," Ear Hear. 22, 212-224.

Cook, R. D. (1979). "Influential observations in linear regression," J. Am Stat. Assoc. 74, 169-174.

Dajani, H. R., Purcell, D., Wong, W., Kunov, H., and Picton, T. W. (2005). "Recording human evoked potentials that follow the pitch contour of a natural vowel," IEEE Trans. Biomed. Engineer. 52, 1614-1618.

Dajani, H. R., Heffernan, B. P., and Giguere, C. (2013). "Improving hearing aid fitting using the speech-evoked auditory brainstem response," 35th Ann. Int. Conf. IEEE pp.2812-2815.

De Gennaro, L., and Ferrara, M. (2003). "Sleep spindles: An overview," Sleep Med. Rev. 7, 423-440.

Doelling, K.B., Arnal, L.H., Ghitza, O., and Poeppel, D. (2014). "Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing," Neuroimage, 85, 761-768.

Drullman, R., Festen, J. M., and Plomp, R. (1994). "Effect of temporal envelope smearing on

speech reception," J. Acoust. Soc. Am. 95, 1053-1064.

Eaves, J. M., Summerfield, A. Q., and Kitterick, P. T. (2011). "Benefit of temporal fine structure to speech perception in noise measured with controlled temporal envelopes," J. Acoust. Soc. Am. 130, 501-507.

Faulkner, A., Rosen, S., and Smith, C. (2000). "Effects of the salience of pitch and periodicity information on the intelligibility of four-channel vocoded speech: Implications for cochlear implants," J. Acoust. Soc. Am. 108, 1877–1887.

Fujihira, H., and Shiraishi, K. (2015). "Correlations between word intelligibility under reverberation and speech auditory brainstem responses in elderly listeners," Clin. Neurophysiol. 126, 96-102.

Géléoc, G. S., and Holt, J. R. (2014). "Sound strategies for hearing restoration," Science 344, 1241062.

Gopinath, B., Rochtchina, E., Wang, J. J., Schneider, J., Leeder, S. R., and Mitchell, P. (2009). "Prevalence of age-related hearing loss in older adults: Blue Mountains Study," Arch. Intern. Med. 169, 415-418.

Greenberg, S., Carvey, H., Hitchcock, L., and Chang, S. (2003). "Temporal properties of spontaneous speech - A syllable-centric perspective," J. Phonet. 31, 465-485.

Grose, J. H., and Mamo, S. K. (2010). "Processing of temporal fine structure as a function of age," Ear Hear. 31, 755.

Hair, J. F., Anderson, R. E., Babin, B. J., and Black, W. C. (2010). *Multivariate data analysis: A global perspective* (Upper Saddle River, NJ Pearson).

Helfer, K. S., and Freyman, R. L. (2008). "Aging and speech-on-speech masking," Ear Hear.

29, 87-98.

Henry, K.S., and Heinz, M.G. (2012). "Diminished temporal coding with sensorineural hearing loss emerges in background noise," Nat. Neurosci. 15, 1362-1364.

Henry, K.S., Kale, S., and Heinz, M.G. (2014). "Noise-induced hearing loss increases the temporal precision of complex envelope coding by auditory-nerve fibers," Front. Syst. Neurosci. 8:20.

Hilbert, D. (1912). "Begründung der kinetischen Gastheorie," Mathematische Annalen 72, 562-577.

Hoen, M., Meunier, F., Grataloup, C. L., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X., and Collet, L. (2007). "Phonetic and lexical interferences in informational masking during speech-in-speech comprehension," Speech Comm. 49, 905-916.

Hopkins, K., Moore, B. C., and Stone, M. A. (2008). "Effects of moderate cochlear and hearing loss on the ability to benefit from temporal fine structure information in speech," J. Acoust. Soc. Am. 123, 1140-53.

Hopkins, K., and Moore, B. C. (2011). "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise," J. Acoust. Soc. Am. 130, 334-49.

Horwitz, A. R., Ahlstrom, J. B., and Dubno, J. R. (2008). "Factors affecting the benefits of high-frequency amplification," J. Speech, Lang. Hear. Res. 51, 798-813.

Howard, M. F., and Poeppel, D. (2010). "Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension," J. Neurophysiol. 104, 2500-2511.

Humes, L. E., Kewley-Port, D., Fogerty, D., and Kinney, D. (2010). "Measures of hearing threshold and temporal processing across the adult lifespan," Hear. Res. 264, 30-40.

Kale, S., and Heinz, M.G. (2010). "Envelope coding in auditory nerve fibers following noise-induced hearing loss," J. Assoc. Res. Otolaryngol. 11, 657-673.

Koerner, T.K., and Zhang, Y., (2015). "Effects of background noise on inter-trial phase coherence and auditory N1-P2 responses to speech stimuli," Hear Res. 328, 113-119.

Krishnan, A., Xu, Y., Gandour, J., and Cariani, P. (2005). "Encoding of pitch in the human brainstem is sensitive to language experience," Cogn. Brain Res. 25, 161-168.

Kuruvilla-Mathew, A., Purdy, S.C., and Welch, D. (2015) "Cortical encoding of speech acoustics: Effects of noise and amplification," Intern. J. Audiol. 54, 852-864.

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., and Schroeder, C. E. (2007). "Neuronal oscillations and multisensory interaction in primary auditory cortex," Neuron 53, 279-292.

Lavitt, H. (2007). "A historical perspective on digital hearing aids: how digital technology has changed modern hearing aids." Trends Amp 11, 7-24.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. (2006). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," Pro. Nat. Acad. Sci. U.S.A. 103, 18866-18869.

Lunner, T., Rudner, M., and Rönnberg, J. (2009). "Cognition and hearing aids," Scand. J. Psychol. 50, 395-403.

Luo H., and Poeppel D. (2007). "Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex," Neuron 54, 1001-1010.

Marsh, J. E., and Campbell, T. A. (2016). "Processing complex sounds passing through the

rostral brainstem: The new early filter model," Front. Neurosci. 10, 136.

Martin, N., Lafortune, M., Godbout, J., Barakat, M., Robillard, R., Poirier, G., Bastien, C., and Carrier, J. (2013). "Topography of age-related changes in sleep spindles," Neurobiol. Aging 34, 468-476.

Moore, B. C. (2008). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," J. Assoc. Res. Otolaryngol. 9, 399-406.

Moore, B. C. (2014). *Auditory processing of temporal fine structure: Effects of age and hearing loss* (World Scientific).

Morillon, B., Liégeois-Chauvel, C., Arnal, L. H., Bénar, C. G., and Giraud, A. L. (2012). "Asymmetric function of theta and gamma activity in syllable processing: An intra-cortical study," Front. Psychol., 3, 248.

Ng, B. S., Logothetis, N. K., and Kayser, C. (2013). "EEG phase patterns reflect the selectivity of neural firing," Cereb. Cort. 23, 389-398.

Noguchi, Y., Fujiwara, M., and Hamano, S. (2015). "Temporal evolution of neural activity underlying auditory discrimination of frequency increase and decrease," Brain Topogr. 28, 437-444.

Oxenham, A. J., Micheyl, C., and Keebler, M. V. (2009). "Can temporal fine structure represent the fundamental frequency of unresolved harmonics?" J. Acoust. Soc. Am. 125, 2189-2199.

Peelle, J.E., Gross, J., and Davis, M.H., (2012). "Phase-locked responses to speech in human auditory cortex are enhanced during comprehension," Cereb. Cort., 23, 1378-1387.

Plomp, R., and Mimpen, A. M. (1979). "Speech reception threshold for sentences as a function of age and noise level," J. Acoust. Soc. Am. 66, 1333-1342

Portas, C. M., Krakow, K., Allen, P., Josephs, O., Armony, J. L., and Frith, C. D. (2000). "Auditory processing across the sleep-wake cycle: Simultaneous EEG and fMRI monitoring in humans," Neuron 28, 991-999.

Presacco, A., Simon, J. Z., and Anderson, S. (2016). "Evidence of degraded representation of speech in noise, in the aging midbrain and cortex," J. Neurophysiol. 116, 2346-2355.

Proakis, J. G., and Manolakis, D. G. (2007). *Digital Signal Processing, 4th Edition* (Pearson).

Qin, M. K., and Oxenham, A. J. (2006). "Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech," J. Acoust. Soc. Am. 119, 2417-26.

Rajan, R., and Cainer, K. E. (2008). "Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers," Neuroscience 154, 784-95.

Rosen, S. (1992). "Temporal information in speech: Acoustic, auditory and linguistic aspects," Phil. Trans. R. Soc. B: Bio. Sci. 336, 367-373.

Rosen, S., Souza, P., Ekelund, C., and Majeed, A. A. (2013) "Listening to speech in a background of other talkers: Effects of talker number and noise vocoding," J. Acoust. Soc. Am. 133, 2431-2443.

Rothauser, E. H., Chapman, W. D., Guttman, N., Nordby, K. S., Silbiger, H. R., Urbanek, G. E., and Weinstock, M. (1969). "IEEE recommended practice for speech quality measurements," IEEE Trans. Aud. Electroacoust. 17, 225-246.

Russo, N., Nicol, T., Musacchia, G., and Kraus, N. (2004). "Brainstem responses to speech syllables," Clin. Neurophysiol. 115, 2021-2030.

Schoof, T., and Rosen, S. (2014). "The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners," Front. Aging Neurosci. 6:307 10.3389/fnagi.2014.00307.

Schoof, T., and Rosen, S. (2016). "The role of age-related declines in subcortical auditory processing in speech perception in noise," J. Assoc. Res. Otolaryngol. 17, 441-460.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," Science. 270, 303-304.

Shapiro, S. S., and Wilk, M. B. (1965). "An analysis of variance test for normality (complete samples)," Biometrika 52, 591-611.

Skoe, E., and Kraus, N. (2010). "Auditory brainstem response to complex sounds: A tutorial," Ear Hear. 31, 302-324.

Smith, Z. M., Delgutte, B., and Oxenham, A. J. (2002). "Chimaeric sounds reveal dichotomies in auditory perception," Nature 416, 87-90.

Song, J.H., Nicol, T. and Kraus, N. (2011). "Test–retest reliability of the speech-evoked auditory brainstem response," Clin. Neurophysiol. 122, 346-355.

Song, J. H., Skoe, E., Banai, K., and Kraus, N. (2010). "Perception of speech in noise: Neural correlates," J. Cogn. Neurosci. 23, 2268-79.

Stickney, G. S., Assmann, P .F., Chang, J., and Zeng, F. G. (2007). "Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences," J. Acoust. Soc. Am. 122, 1069-1078.

Thatcher, R. W. (2012). "Coherence, phase differences, phase shift, and phase lock in EEG/ERP analyses," Dev. Neuropsychol. 37, 476-496.

Tremblay, K. L., and Miller, C. W. (2014). "How neuroscience relates to hearing aid amplification," Int. J. Otolaryngol. 2014, 1-7.

Turner, C. W., and Cummings, K. J. (1999). "Speech audibility for listeners with high-frequency hearing loss," Am. J. Audiol. 8, 47-56.

Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., and Henry, B. A. (2004). "Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing," J. Acoust. Soc. Am. 115, 1729-1735.

Warby, S. C., Wendt, S. L., Welinder, P., Munk, E. G., Carrillo, O., Sorensen, H. B., and Mignot, E. (2014). "Sleep-spindle detection: C rowdsourcing and evaluating performance of experts, non-experts and automated methods," Nat. Meth. 11, 385-392.

Wilf, M., Ramot, M., Furman-Haran, E., Arzi, A., Levkovitz, Y., and Malach, R. (2016). "Diminished auditory responses during NREM sleep correlate with the hierarchy of language processing," PloS One 11, e0157143.

Wong, P. C., Skoe, E., Russo, N.M., Dees, T., and Kraus, N. (2007). "Musical experience shapes human brainstem encoding of linguistic pitch patterns," Nat. Neurosci. 10, 420-422.

Zeng, F. G., Nie, K., Liu, S., Stickney, G., Del Rio, E., Kong, Y. Y., and Chen, H. (2004). "On the dichotomy in auditory perception between temporal envelope and fine structure cues," J. Acoust. Soc. Am. 116, 1351-1354.

Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A., Wei, C., and

Cao, K. (2005). "Speech recognition with amplitude and frequency modulations," Pro. Nat. Acad. Sci. U.S.A. 102, 2293-2298.

Zhong, Z., Henry, K.S., and Heinz, M.G. (2014). "Sensorineural hearing loss amplifies neural coding of envelope information in the central auditory system of chinchillas," Hear. Res. 309, 55-62.