

RESEARCH

Open Access



Using frequency-following responses (FFRs) to evaluate the auditory function of frequency-modulation (FM) discrimination

Zhen Fu, Xihong Wu and Jing Chen*

*Correspondence:
chenj@cis.pku.edu.cn
Department of Machine
Intelligence, Speech
and Hearing Research
Center, and Key Laboratory
of Machine Perception
(Ministry of Education),
Peking University,
Beijing 100871, China

Abstract

Precise neural encoding of varying pitch is crucial for speech perception, especially in Mandarin. A valid evaluation of the listeners' auditory function which accounts for the perception of pitch variation can facilitate the strategy of hearing compensation for hearing-impaired people. This auditory function has been evaluated by behavioral test in previous studies, but the objective measurement of auditory-evoked potentials, for example, is rarely studied. In this study, we investigated the scalp-recorded frequency-following responses (FFRs) evoked by frequency-modulated sweeps, and its correlation with behavioral performance on the just-noticeable differences (JNDs) of sweep slopes. The results showed that (1) the indices of FFRs varied significantly when the sweep slopes were manipulated; (2) the indices were all strongly negatively correlated with JNDs across listeners. The results suggested that the listener's subjective JND could be predicted by the objective index of FFRs to tonal sweeps.

Keywords: Frequency-following responses, Pitch perception, Just-noticeable differences, Tonal sweeps, Sweep slope

Introduction

Speech intelligibility varies greatly across hearing-impaired listeners and even across normal-hearing people, especially in noise, despite the similarity between their ages or audibility (Humes et al. 2009). Compensation effects introduced by hearing aids seem to vary greatly among individual listeners, thereby indicating the insufficiency of audibility as a predictor for speech reception and instructor for compensation strategy. Consequently, supra-threshold audiometry metrics were introduced to evaluate listener's auditory function, (Strelcyk and Dau 2009; Papakonstantinou et al. 2011). Among them, the frequency-modulation detection limen (FMDL), achieved by requiring listeners to select a manipulated FM signal among pure tones in test trials, reflecting listeners' auditory function of frequency discrimination, was shown to correlate with speech reception threshold (SRT) (Strelcyk and Dau 2009). However, it is a metric, dependent on listeners' subjective feedback, termed as subjective measurement, which is not practical for clinical application. In this study, we aimed to develop an objective measurement to reflect the FMDL.

Pitch is the psychological perception of fundamental frequency (f_0) and plays an essential role in speech perception. Precise encoding of voice pitch and its variation with time are crucial for listeners to perceive different intonational cues, especially for understanding the lexical meanings in tonal languages under masking conditions (Wu et al. 2011). Frequency-modulated glide, also referred to as tonal sweep, is regarded as a simplified version of the f_0 variation over time for voiced speech. Previous psychophysical tests using tonal sweeps suggest that the processing mechanism underlying the perception of slowly frequency-modulated tones is related to the temporal processing mechanisms of auditory system, e.g., the phase-locking firing of auditory nerves (Sek and Moore 1995).

Frequency-modulation detection limen is reported to be consistent with listeners' performance on pitch-change discrimination and is degraded for hearing-loss-affected listeners (Papakonstantinou et al. 2011), but the corresponding objective measurement is rarely studied. In the present study, the just-noticeable differences (JNDs) of the slopes of sweep signals were measured, and frequency-following responses (FFRs) served as the objective measurement for this function.

Frequency-following responses reflect subcortical phase-locking activity evoked by periodic sounds in brainstem (Skoe and Kraus 2010). The spectral peaks of FFRs locate at each harmonic of stimulus f_0 , with energy mainly concentrating at f_0 . The fidelity of FFRs has been shown to be correlated to pitch perception (Marmel et al. 2013) and sensitive to hearing situation (Plyler and Ananthanarayan 2001).

Frequency discrimination limen (FDL), measured with pure tones, was found to be negatively correlated with pitch strength reflected in FFRs (Marmel et al. 2013). However, it was suggested that the neural processing of repetitive FM sweeps in the human auditory cortex differs from that of pure tones (Okamoto and Kakigi 2017). Tonal sweeps and syllables were used to evoke FFRs to test the effect of signal phase on neural encoding of speechlike sounds (Jeng et al. 2011; Bidelman 2014). These studies mainly focused on developing new metrics for analysis of FFRs, or on factors impacting the neural representation of speechlike sounds. So far, the relation between the behavioral performances on FMDL and FFRs to tonal sweeps was not analyzed and discussed. The primary purpose of the present study was to make a reliable index with FFRs to predict listeners' FMDL in normal-hearing listeners.

Materials and methods

Experiment 1: behavioral JND measurement

The behavioral experiment measured the just-noticeable difference (JND) of the onset f_0 between standard stimulus (pure tone) and deviation stimulus (rising sweep).

Participants

Thirteen Mandarin-native adults (mean age = 22.8 year, SD = 0.8 year) participated in the experiment. All participants who had normal hearing with threshold no more than 20 dB HL at octave frequencies between 125 and 8000 Hz were paid, and all of them gave informed consent in compliance with a protocol approved by the Institutional Review Board at Peking University.

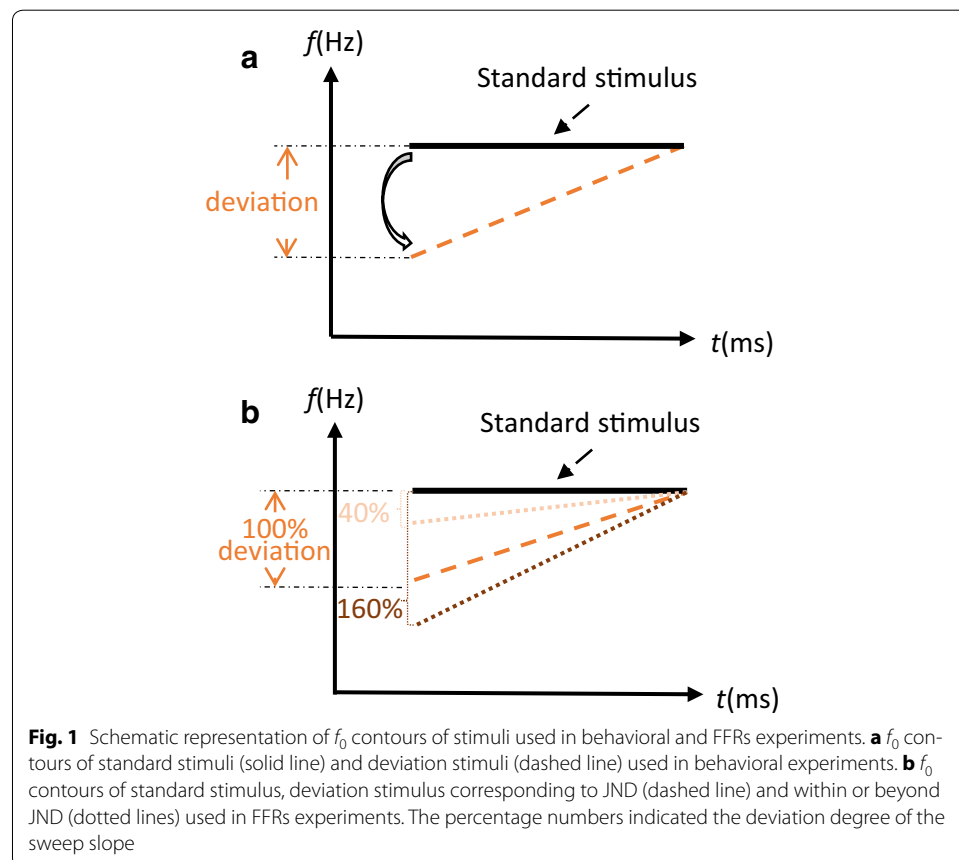
Stimuli

The standard stimulus was a 150 Hz pure tone, which roughly locates in the range of f_0 of human voice. The deviation stimulus was a rising tonal sweep obtained by manipulating the onset f_0 downward, while the offset f_0 was fixed to be equal to 150 Hz, as proposed in Liu (2013) and shown in Fig. 1a.

The signal duration was fixed at 200 ms, including 10-ms rise–fall times shaped by a cosine-squared window. All stimuli were generated by Matlab (Mathworks, Natick, MA) with 16 bits quantization and 44.1 kHz sampling rate. Signal presentation was controlled by a customized routine program written in Matlab. Digital stimuli were presented to the right ears of listeners, who were seated in a sound-attenuation booth, through a Sennheiser HD 265 headphone. The sound level was fixed at 77 dB SPL.

Procedure

Just-noticeable difference was measured through a 3-interval, 2-alternative forced-choice procedure, estimating 71% correct responses (Levitt 1971). The first interval always represented the standard stimulus, while the deviation stimulus and another standard stimulus were randomly assigned to the last two intervals. When subjects clicked a “play button,” the three intervals would be played successively with an inter-stimulus-interval (ISI) of 500 ms. Subjects were instructed to choose the deviation stimulus from the last two intervals with feedbacks. The deviation of onset f_0 was manipulated through



a 2-down 1-up algorithm, meaning that the deviation was decreased after two correct responses and increased after one wrong response. The deviation was 30 Hz initially and adjusted by a factor of 1.414 and following two reversals in the direction of f_0 change, the factor was reduced by its square root. JND for one run was determined by averaging the last 8 reversals after 12 reversals were obtained. Listeners need to finish two or three runs for getting a stable JND.

Experiment 2: FFRs to tonal sweeps

In this experiment, FFRs to standard stimulus and deviation stimulus within, at, and beyond the individual JNDs were recorded and analyzed to explore the neural representation patterns of sweep signals.

Participants

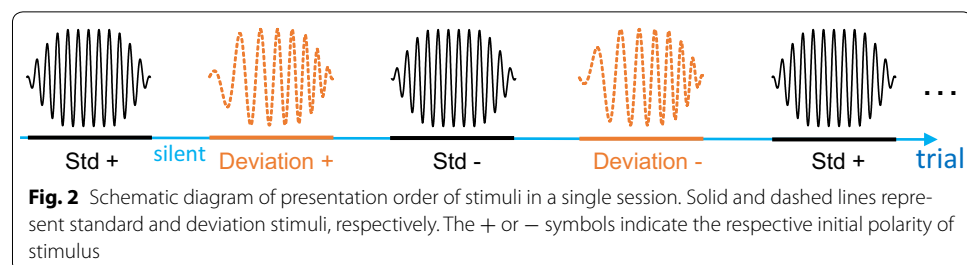
All subjects in experiment 1 also participated in this experiment.

Stimuli

The standard stimulus was a 150-Hz pure tone for all subjects, but the deviation stimulus was adjusted according to each subject's JND. Stimuli with the deviation less or greater than the JND (i.e., within or beyond threshold) were also used to evoke FFRs. In the present study, five different deviation percentages of JND were used: 20, 40, 80, 100, and 160%. Figure 1b shows the schematic representation of f_0 contours of stimuli used to evoke FFRs. All the stimuli parameters were kept the same as in experiment 1, except that the ISI was 100 ms. Stimuli were presented to the right ear through an ER-3A insert earphone. Subjects watched a silent, captioned movie to keep them awake and avoid ignoring the acoustic stimuli.

Experimental protocol

There were five sessions for each subject, corresponding to the five deviation conditions. In each session, the standard stimulus and the deviation stimulus were presented alternatively with alternative polarities of zero onset phase. Therefore, there were totally 10 (5 sessions \times 2 conditions) sets of FFR data to be analyzed offline. Please notice that subtraction should be conducted between the responses to two polarities to get the responses to f_0 (Aiken and Picton 2008). The orders of sessions were randomly assigned across subjects. Each subject completed all sessions in about 2 h. Figure 2 shows the presentation order of the stimuli in one session.



Recording system

Frequency-following responses recordings were conducted in a double-walled sound-attenuation booth (IAC acoustics, North Aurora, IL). A vertical electrode montage was adopted with three Ag–AgCl electrodes placed at Cz (noninverting), right earlobe (inverting), and forehead (ground). The inter-electrode impedances were maintained under 5 k Ohm. The EEG recordings were amplified on using a gain of 20,000 (NeuroScan SynAmps2 amplifier, 24-bit resolution and 0.15 nV/LSB accuracy), bandpass filtered (0.05–3000 Hz, 6 dB/octave) and digitized at a sampling rate of 20,000. Continuous EEG data were recorded through NeuroScan Acquire 4.3 software (Compumedics, Charlotte) and stored for offline analysis.

Response evaluation

Recordings were segmented into sweeps of 300 ms in length, including two 50 ms pre- and post-stimulus intervals. A single sweep would be rejected if it contained voltages exceeding $\pm 25 \mu\text{V}$. After artifact rejection, sweeps of the same polarity (positive or negative) and the same stimulus (standard or deviation) in one session were averaged to improve the signal-to-noise ratio (SNR). Subtraction was conducted between the averaged sweeps of two polarities for the same stimulus in each session to extract the responses to f_0 . Subtracted data were then passed through a posteriori Wiener filter to attenuate the stochastic noise and emphasize the deterministic evoked component. The data were then filtered through a 200th order FIR lowpass filter with cutoff frequency of 400 Hz, to extract the responses to f_0 more precisely.

A high order of filter would introduce a temporal delays. To identify the onset of FFRs signal, cross-correlation of waveforms between stimulus and filtered responses were calculated to identify and compensate this temporal delay. Figure 3 shows an example of these processing methods applied to a recording FFRs in this study.

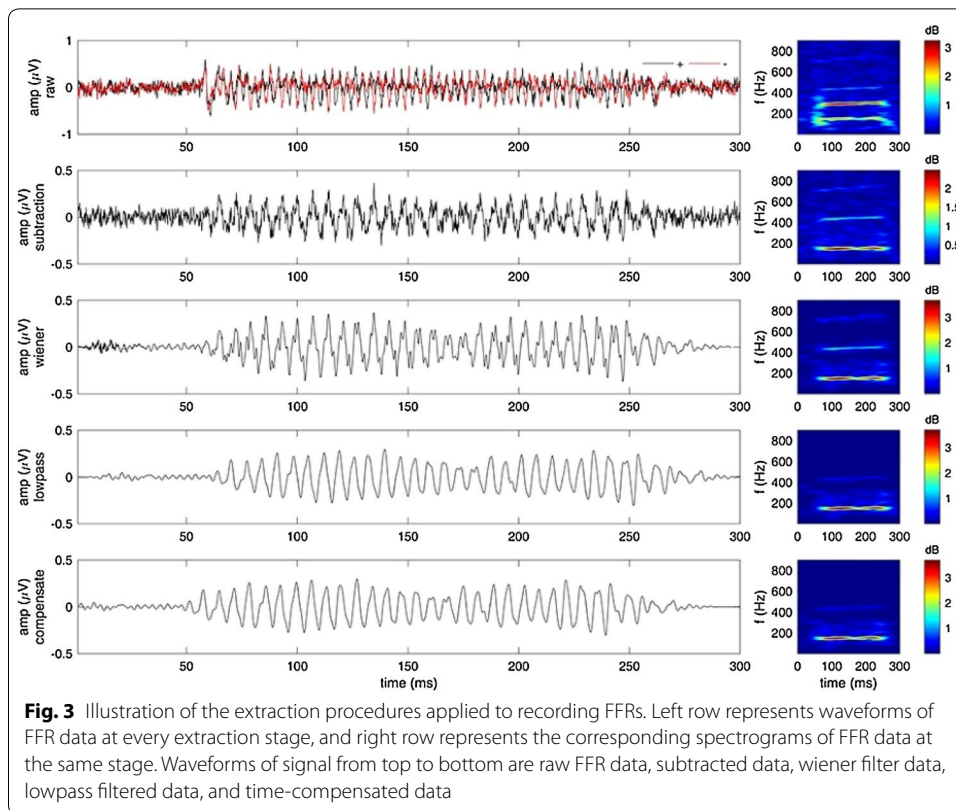
Objective indices

To obtain an evaluation of phase-locked activity in the evoked responses synchronized to stimuli, or say, to describe the fidelity of FFRs, the indices below were used to evaluate the FFRs with respect to standard stimulus and deviation stimulus.

Pitch strength Pitch strength (PS) measured the degree of neural phase locking to the f_0 contours of stimulus. This index was derived from a normalized autocorrelation function that measured the overall periodicity of a signal. Specifically, the responses part of the recording waveforms (i.e., 200 ms of the entire 300 ms segment) were first segmented into frames using a 20-ms Hanning window with 1-ms overlap. This resulted in 181 windowed frames to be analyzed. The function of autocorrelation values $r_i(m)$ versus time shift m for the i th frame could be obtained by Eq. (1):

$$r_i(m) = \frac{\sum_{N=1}^{N_s} s_i(n)s_i(n-m)}{\sum_{N=1}^{N_s} s_i^2(n)}, \quad m = 0, \dots, N_s - 1. \quad (1)$$

Pitch strength of each frame was calculated by finding the longitudinal distance between the first peak and the subsequent trough in the autocorrelation function output (Jeng et al. 2011). Because the f_0 contours of all stimuli used in this study fell within the



range of 120–160 Hz (with a certain amount of buffer for error measurement), the time shifts were limited to 6.25–8.33 ms when searching for the location of the peak. Finally, the general PS of the entire responses sweep was calculated by averaging the scores across frames. In addition, PS of stimuli were also calculated to be comparable to that of responses. Normalization was performed to all PS data.

Mutual information For arbitrary inputs, mutual information (MI) is a dimensionless quantity (in bits) that measures the degree of shared information (i.e., mutual dependence) between two random variables. In general, for two random variables A and B , mutual information is calculated according to Eq. (2):

$$MI(A, B) = \sum_{a \in A} \sum_{b \in B} p(a, b) \log \left(\frac{p(a, b)}{p(a)p(b)} \right), \quad (2)$$

where $p(a, b)$ is the joint probability of A and B , $p(a)$ and $p(b)$ are the marginal probabilities of A and B , respectively. In the present study, MI is used to compute the similarity between the two images (i.e., spectrograms) of stimulus and corresponding responses. Similar computational method was also adopted in previous study (Bidelman 2014).

Results

Experiment 1: behavioral JND measurement

For twelve subjects (one was excluded from counting and data analysis due to poor SNR of FFRs), the mean values of JNDs of about 6.5 Hz, with the standard deviation of 2.8 Hz,

were similar to the results (mean value of 5.5 Hz) reported in Liu (2013), confirming that the experimental setup was effective. JND value for each subject is specified in Table 1.

Experiment 2: FFRs to tonal sweeps

Pitch strength–JND function

Pitch strength curve as a function of JND is shown in Fig. 4. The left panel shows the PS of the stimuli and the responses averaged across subjects in each session. It is shown that the PS was generally greater for stimulus than for responses, and the strength of deviation stimulus decays linearly with the deviation increases. For responses, there was no significant main effect of deviation degree ($F(4,44) = 0.427, p = 0.788$), nor of stimulus condition ($F(1,11) = 1.972, p = 0.188$), while the interaction was significant ($F(4,44) = 3.544, p = 0.014$). Furthermore, a one-way ANOVA test on the PS of responses showed a significant difference between the two responses evoked by the standard stimuli and the deviation stimuli, only at 160% deviation degree ($F(1,22) = 5.589, p = 0.027$).

The detailed response pattern for the 100% deviation is shown in the right panel. The circles represent PS of deviation stimuli-evoked responses for each subject. These discrete data were fitted by the least square method, as indicated by the solid line. Pearson’s correlation analysis indicated that PS was significantly correlated with JNDs ($r = -0.587, p = 0.045$). This negative correlation between neural index and behavioral threshold of tonal sweeps is as expected and consistent with that of FDL (Marmel et al. 2013; Zhang and Gong 2017). The similar correlation analysis was conducted for the other four deviation conditions separately, and it was found that correlation was not significant for the fewer degrees ($r = -0.233, p = 0.465$ for 20% and $r = -0.431, p = 0.161$

Table 1 Individual JND value in behavioral measurement

Subject no.	1	2	3	4	5	6	7	8	9	10	11 ^a	12	13	Mean
JND (Hz)	3.9	10.4	7.6	5.1	7.3	6.8	10.5	4.1	2.3	3.2	3.9 ^a	11.3	5.5	6.5

^a Subject was excluded from counting and data analysis due to a poor SNR of FFRs

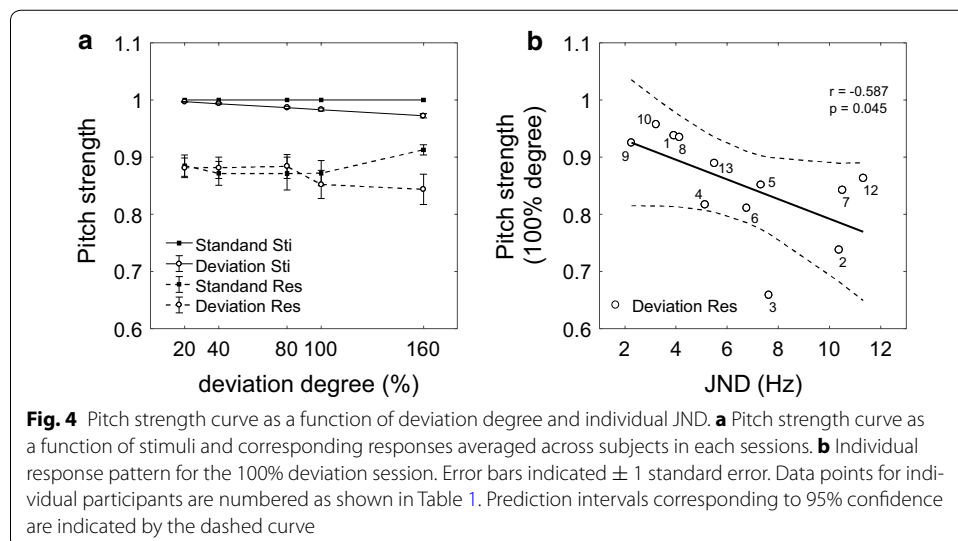


Fig. 4 Pitch strength curve as a function of deviation degree and individual JND. **a** Pitch strength curve as a function of stimuli and corresponding responses averaged across subjects in each sessions. **b** Individual response pattern for the 100% deviation session. Error bars indicated ± 1 standard error. Data points for individual participants are numbered as shown in Table 1. Prediction intervals corresponding to 95% confidence are indicated by the dashed curve

for 40%) but significant for degrees near and beyond threshold ($r = -0.578$, $p = 0.049$ for 80% and $r = -0.622$, $p = 0.031$ for 160%). Therefore, the results suggested that the objective neural PS of FFRs to sweep signals probably could be used to predict listener's subjective JND in this study, although more data are needed to build such a computational model.

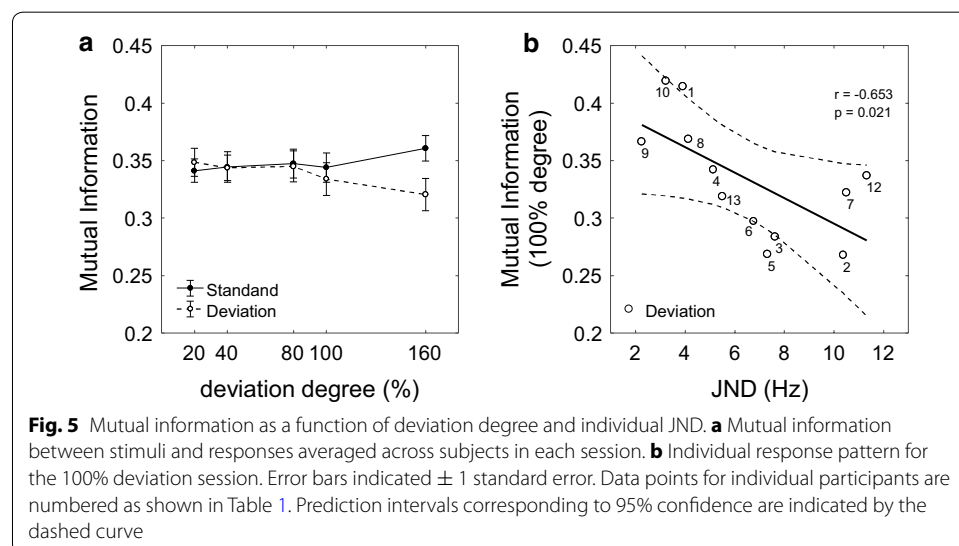
Mutual information–JND function

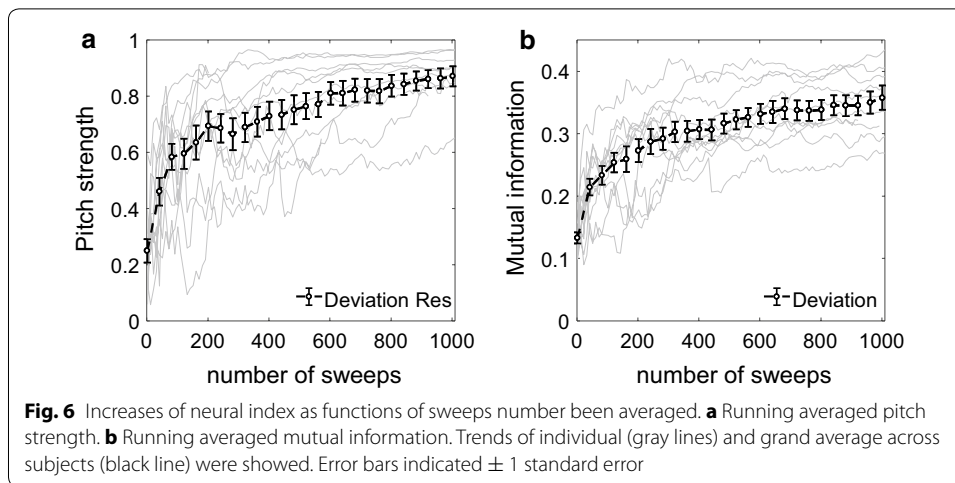
The mutual information as a function of JND is plotted in Fig. 5. Patterns of MI across different deviation degrees were similar to PS (as in the left panel of Fig. 4). Main effect of deviation degree ($F(4,44) = 0.317$, $p = 0.865$) was not significant, while main effects of stimulus types ($F(1,11) = 5.009$, $p = 0.047$) and the interaction were significant ($F(4,44) = 6.334$, $p < 0.001$). Similarly, a one-way ANOVA test on the MI showed a significant difference between two conditions only at 160% deviation ($F(1,16) = 4.577$, $p = 0.044$).

The individual MI for the 100% deviation is shown in the right panel. Pearson's correlation analysis indicated that MI is significantly correlated with JNDs ($r = -0.653$, $p = 0.021$). Likewise, the negative correlation was as expected and consistent with previous research. Correlation was not significant for fewer degrees ($r = -0.101$, $p = 0.756$ for 20% deviation and $r = -0.566$, $p = 0.055$ for 40% deviation), but significant for degrees near and beyond threshold ($r = -0.623$, $p = 0.03$ for 80% deviation and $r = -0.617$, $p = 0.033$ for 160% deviation). These results suggested that the objective index, MI of FFRs, could also be used to predict listener's subjective JND.

FFRs' running trends

Capability of online analysis for an evaluation metric is necessary for clinical application, which requires the metric to be sensitive and efficient as the sweeps are being averaged continuously. Besides, an organized running trend would be instructive to design criterion for audiometry. Figure 6 shows the running averages of PS and MI of FFRs as a





function of sweep numbers. Only FFRs to the 100% deviation condition is drawn due to its practical value.

Asymptotic trends were observed for both indices (PS and MI). The fidelity of FFRs to sweeps nearly reached saturation when the number of running-sweeps exceeded 600 (300 for each polarities), indicating the efficiency of using these two metrics as real-time monitors of the neural encoding for tonal sweeps, and the potential for clinical application.

Discussion

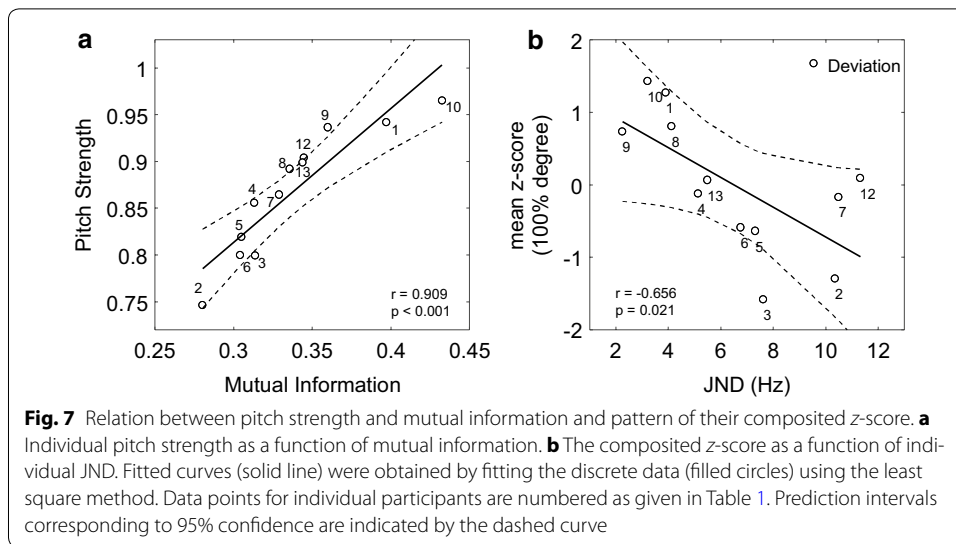
Consistency between pitch strength and mutual information

Pitch strength and mutual information had similar trends for predicting JNDs in this study. The correlations between these two metrics are illustrated in the left panel of Fig. 7, showing that the two metrics are significantly correlated with each other ($r = -0.909$, $p < 0.001$).

A comprehensive index could be obtained by calculating z -scores of PS and MI separately and then averaging them. In statistics, z -score is the multiple of SD by which the value of a data is above the mean of group. The composited z -scores for the 100% deviation are plotted against JNDs for each subject in the right panel of Fig. 7, which significantly correlated with individual JNDs ($r = -0.656$, $p = 0.021$), and the correlation was observed to be slightly stronger than only PS or MI as the index.

Pitfall of utilizing posteriori wiener filtering

Although Wiener filtering did perform well in the extraction procedures shown in Fig. 3, it did not perform well all the time. In the posteriori wiener filtering algorithm, if the noise dominates the recording data (e.g., due to a negative example of a NH subject or a data from severe SNHL subject), the filter would mistake the very weak evoking potential to be noise so that the filtered SNR would be worse. One participant had to be excluded from the data analysis in this study for this same reason. Similar incidences were also reported in (Gong et al. 2013). Therefore, during the application of wiener filter to FFRs for subjects with hearing loss, one should be extra cautious, because the responses would be smeared due to sensorineural hearing loss.



Conclusion

The present study found a strong correlation between pitch discrimination ability and neural synchrony to tonal sweeps for normal-hearing people. Pitch discrimination was measured by a behavioral JND paradigm for FMDL, and the neural synchrony was indexed by the fidelity of scalp-recorded FFRs to rising sweeps. These results indicated that the two objective indices of FFRs to tonal sweep could be used to predict listener's subjective FMDL. The proposed method of this study is worthy to apply on hearing-impaired listeners to testify its feasibility for clinical diagnosis studies in future.

Abbreviations

FMDL: frequency-modulation detection limen; FFRs: frequency-following responses; JNDs: just-noticeable differences; PS: pitch strength; MI: mutual information.

Authors' contributions

ZF, XW, and JC conceived and designed the experiments. ZF performed the experiments. ZF and JC analyzed the data and also wrote the paper. All the authors read and approved the final manuscript.

Acknowledgements

The authors thank Jiping Wu for his valuable instructions on FFRs data analysis, and Hongying Yang for her help with statistical analyses.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Funding

The study was supported by the National Natural Science Foundation of China (Nos. 61473008, 11590773, and 61771023), and a Newton alumni funding by the Royal Society, UK. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 12 September 2017 Accepted: 16 October 2017

Published online: 23 October 2017

References

- Aiken SJ, Picton TW (2008) Envelope and spectral frequency-following responses to vowel sounds. *Hear Res* 245:35–47. doi:[10.1016/j.heares.2008.08.004](https://doi.org/10.1016/j.heares.2008.08.004)
- Bidelman GM (2014) Objective information-theoretic algorithm for detecting brainstem-evoked responses to complex stimuli. *J Am Acad Audiol* 25:715–726. doi:[10.3766/jaaa.25.8.2](https://doi.org/10.3766/jaaa.25.8.2)
- Gong Q, Xu Q, Sun W (2013) Design and implementation of frequency-following response recording system. *Int J Audiol* 52:824–831. doi:[10.3109/14992027.2013.834537](https://doi.org/10.3109/14992027.2013.834537)
- Humes LE, Ahlstrom JB, Bratt GW, Peek BF (2009) Studies of hearing-aid outcome measures in older adults: a comparison of technologies and an examination of individual differences. *Semin Hear* 30:112–128. doi:[10.1055/s-0029-1215439](https://doi.org/10.1055/s-0029-1215439)
- Jeng F-C, Hu J, Dickman B et al (2011) Evaluation of two algorithms for detecting human frequency-following responses to voice pitch. *Int J Audiol* 50:14–26. doi:[10.3109/14992027.2010.515620](https://doi.org/10.3109/14992027.2010.515620)
- Levitt H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49:467–477. doi:[10.1121/1.1912375](https://doi.org/10.1121/1.1912375)
- Liu C (2013) Just noticeable difference of tone pitch contour change for English- and Chinese-native listeners. *J Acoust Soc Am* 134:3011–3020. doi:[10.1121/1.4820887](https://doi.org/10.1121/1.4820887)
- Marmel F, Linley D, Carlyon RP et al (2013) Subcortical neural synchrony and absolute thresholds predict frequency discrimination independently. *J Assoc Res Otolaryngol* 14:757–766. doi:[10.1007/s10162-013-0402-3](https://doi.org/10.1007/s10162-013-0402-3)
- Okamoto H, Kakigi R (2017) Modulation of auditory evoked magnetic fields elicited by successive frequency-modulated (FM) sweeps. *Front Hum Neurosci*. doi:[10.3389/fnhum.2017.00036](https://doi.org/10.3389/fnhum.2017.00036)
- Papakonstantinou A, Strelcyk O, Dau T (2011) Relations between perceptual measures of temporal processing, auditory-evoked brainstem responses and speech intelligibility in noise. *Hear Res* 280:30–37. doi:[10.1016/j.heares.2011.02.005](https://doi.org/10.1016/j.heares.2011.02.005)
- Plyler PN, Ananthanarayan AK (2001) Human frequency-following responses: representation of second formant transitions in normal-hearing and hearing-impaired listeners. *J Am Acad Audiol* 12:523–533
- Sek A, Moore BCJ (1995) Frequency discrimination as a function of frequency, measured in several ways. *J Acoust Soc Am* 97:2479–2486. doi:[10.1121/1.411968](https://doi.org/10.1121/1.411968)
- Skoe E, Kraus N (2010) Auditory brain stem response to complex sounds: a tutorial. *Ear Hear* 31:302–324. doi:[10.1097/AUD.0b013e3181cdb272](https://doi.org/10.1097/AUD.0b013e3181cdb272)
- Strelcyk O, Dau T (2009) Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J Acoust Soc Am* 125:3328–3345. doi:[10.1121/1.3097469](https://doi.org/10.1121/1.3097469)
- Wu X, Yang Z, Huang Y et al (2011) Cross-language differences in informational masking of speech by speech: English versus Mandarin Chinese. *J Speech Lang Hear Res* 54:1506–1524. doi:[10.1044/1092-4388\(2011/10-0282](https://doi.org/10.1044/1092-4388(2011/10-0282)
- Zhang X, Gong Q (2017) Correlation between the frequency difference limen and an index based on principal component analysis of the frequency-following response of normal hearing listeners. *Hear Res* 344:255–264. doi:[10.1016/j.heares.2016.12.004](https://doi.org/10.1016/j.heares.2016.12.004)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
