Separating Chimeric Novelty from Signal Manipulation in Lexical-Tone Chimeras: Behavioral and Electrophysiological Approaches

Fuh-Cherng Jeng[1,3], Breanna N. Hart[1], Ching-Hua Chen[2], Chia-Der Lin[2], Hui-Chi Tien[3]

[1]Communication Sciences and Disorders, Ohio University, Athens, Ohio, USA
[2]Department of Otolaryngology-HNS, China Medical University Hospital, Taichung, Taiwan
[3]Department of Audiology and Speech Pathology, Asia University, Taichung, Taiwan

INTRODUCTION

- Every acoustic stimulus can be mathematically divided into the product of two components: a slow-changing modulation (i.e., the envelope) and a fast-altering carrier (i.e., the fine structure). To evaluate the relative contributions of the envelope and fine structure of acoustic sounds on speech perception, Smith and colleagues (2002) synthesized acoustic stimuli by integrating the envelope of one sound with the fine structure of another.
- Due to the flexibility in manipulating and interchanging the envelopes and fine structures of acoustic stimuli, many studies (Smith et al., 2002; Kraus & Nicol, 2005; ; Jeng et al., 2016 Hou & Xu, 2018; Oxenham 2018; Hu et al., 2020; Warnecke et al., 2020) have utilized auditory chimeras to investigate the relative importance of the envelope and fine structure cues in speech and musical perceptions.
- The purpose of this study was to examine and separate the effects of chimeric novelty from the effects of signal manipulation, by using both behavioral and electrophysiological approaches.

STUDY DESIGN

- When synthesizing auditory chimeras, each input signal was first filtered through a number of filter banks that were equally spaced along the frequency map of human cochleae. The envelope and fine structure of each filtered signal were extracted in each filter bank by using a Hilbert transform. Within each filter bank, the envelope of one input signal was combined with the fine structure of another input signal to constitute a single-bank chimera. These single-bank chimeras were summed across all filter banks to construct multi-bank chimeras.
- Two sets of lexical-tone chimeras were created. The first set of lexical-tone chimeras consisted of the envelope of the Mandarin tone 4 (i.e., /yi^4/) with a falling pitch contour and the fine structure of the Mandarin tone 2 (i.e., /yi^2/) with a rising pitch contour. For simplicity, this set of lexical-tone chimeras was named env4-fts2 chimeras. The second set of lexical-tone chimeras consisted of the envelope of the Mandarin tone 4 and the fine structure of the same Mandarin tone 4. This set of chimeras was named env4-fts4 chimeras.
- By measuring the participant's behavioral and electrophysiological responses elicited by using the env4-fts2 and env4-fts4 chimeras, along with the participant's responses to the original /yi4/ Mandarin tone, the effects of chimeric novelty might be separated from the effects of signal manipulation.
- Specifically, any differences observed in the participant's responses between the env4-fts2 lexical-tone chimeras and the original /yi^4/ Mandarin tone would include both the chimeric novelty and signal manipulation effects. Similarly, any differences observed in the participant's responses between env4-fts2 and env4-fts4 lexical-tone chimeras would include only the chimeric novelty effects. Lastly, any differences observed in the participant's responses between the env4-fts4 lexical-tone chimeras and the original /yi^4/ Mandarin tone would include only the signal manipulation effects.

METHODS

Participants

- Twenty native speakers of Mandarin Chinese (13 females and 7 males, mean age ± standard deviation = 26.70 ± 5.38 years)

Stimuli

- Env4-fts2 lexical-tone chimeras, filtered at 1, 2, 3, 4, 6, 8, 16, 32, and 64 frequency banks
- Env4-fts4 lexical-tone chimeras, filtered at 1, 2, 3, 4, 6, 8, 16, 32, and 64 frequency banks
- Original /yi^4/ Mandarin tone

Behavioral Pitch-Perception Tasks

- Single-interval, two-alternative, forced-choice paradigm
- Familiarization: Each participant was asked to listen to the original /yi^2/ and /yi^4/ Mandarin tones for as many times as they wanted.
- Practice: 19 acoustic stimuli (nine env4-fts2 chimeras, nine env4-fts4 chimeras, and the original /yi^4/ Mandarin tone) were randomly presented through supra-aural headphones to the participant's ears at a self-adjusted, comfortable listening level. After a stimulus was presented, the participant would indicate whether a rising (tone 2) or falling (tone 4) pitch contour dominated their perception.
- Data Collection: Procedures were all the same as those in Practice, except that each stimulus was presented 40 times.

Electrophysiological Measurements

- 3 gold-plated surface recording electrodes
  - High forehead, mastoid, and low forehead
- Due to time constraints, only env4-fts2 filtered at 4 frequency banks, env4-fts4 filtered at 4 frequency banks, and the original /yi^4/ Mandarin tone were used to elicit frequency-following responses (FFRs).
- Stimuli intensity = 70 dB SPL
- 4000 accepted sweeps for each FFR recording

Data Analysis

- Behavioral Pitch-Perception Tasks
  - Reaction Time: A two-way analysis of variance (ANOVA) with repeated measures was used to determine the significance of the participants' reaction times in response to the chimeric stimuli within and across the filter bank (1, 2, 3, 4, 6, 8, 16, 32, and 64 filter banks) and effect type (chimeric novelty, signal manipulation, and total effects) factors.
- Electrophysiological Measurements
  - Two objective indices were applied to estimate the accuracy and magnitude of phase-locking in the brain
  - Frequency Error represents the accuracy of frequency-encoding during the stimulus presentation.
  - Pitch Strength measures the magnitude of the neural phase-locking to the f0 contour of the stimulus waveform.
  - A one-way ANOVA was conducted to determine the significance among the three effect types (chimeric novelty, signal manipulation, and total effects) for Frequency Error and Pitch Strength.

RESULTS

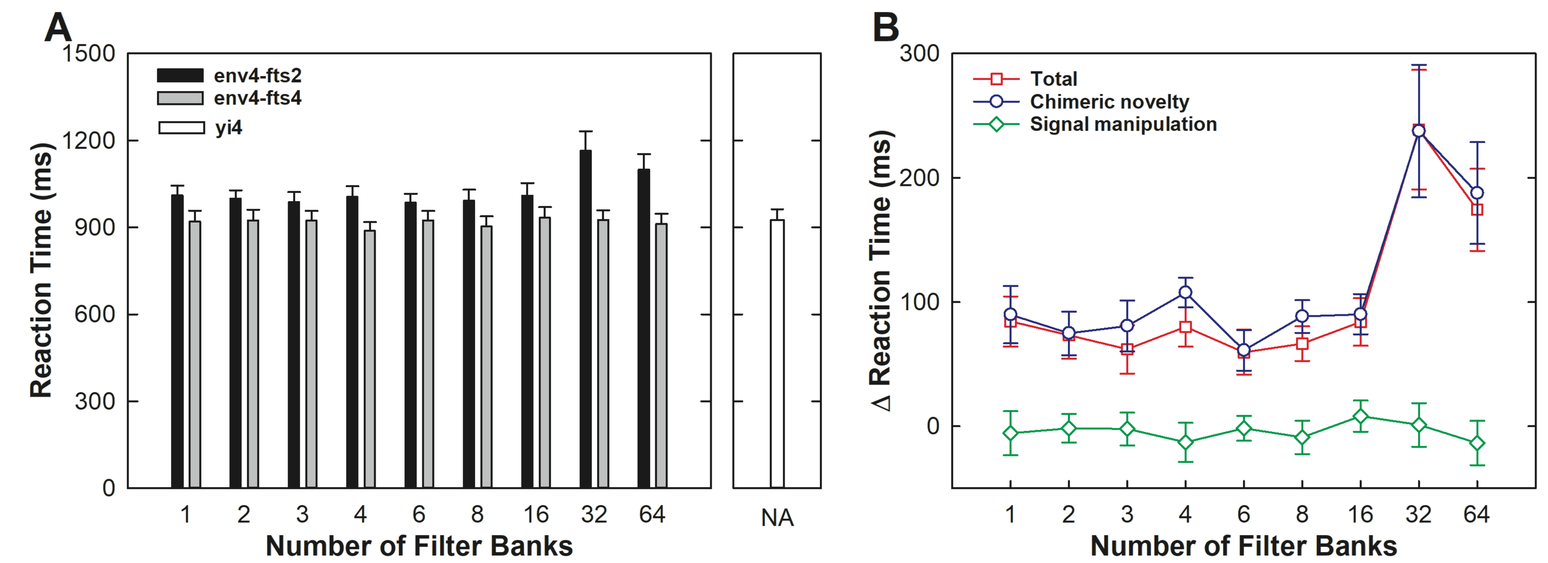

Figure 1 Group means ± 1 standard error of the reaction times (Fig. 1**A**) obtained from twenty native Mandarin speakers in response to the env4-fts2 chimeras (black vertical bars), env4-fts4 chimeras (gray vertical bars), and original /yi^4/ Mandarin tone (white vertical bar). The means and standard errors of the total (red squares), chimeric novelty (blue circles), and signal manipulation (green diamonds) effects are plotted in the same panel for comparison (Fig. 1**B**). Δ reaction time = difference in reaction times.

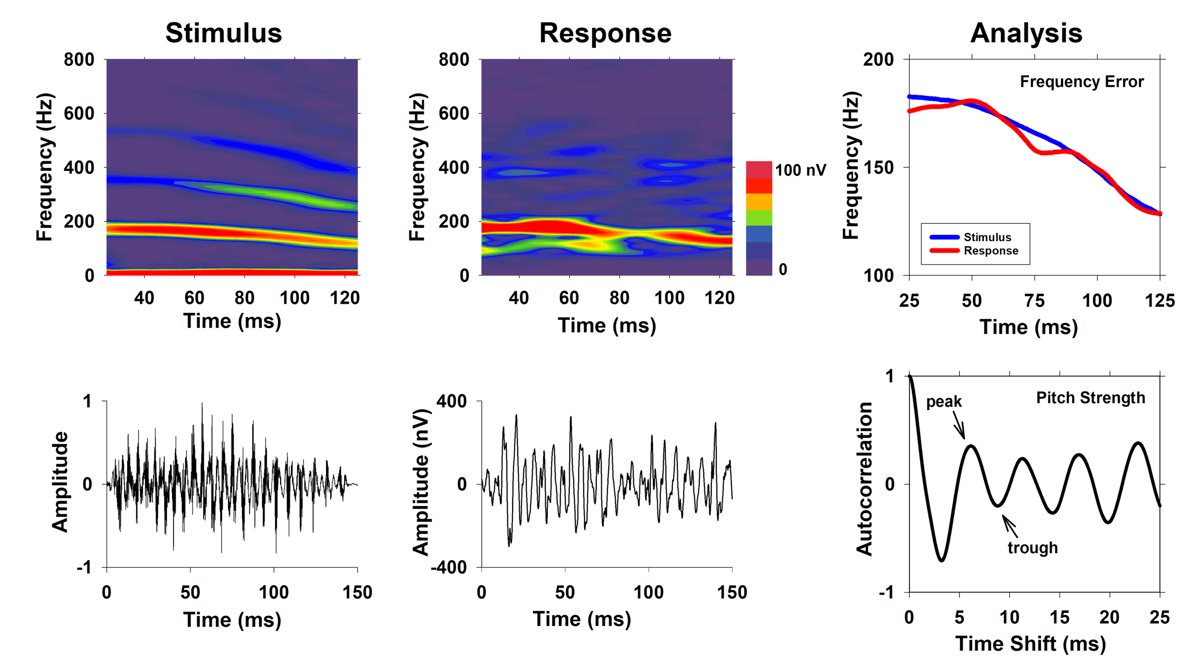

Figure 2 Representative amplitude spectrograms and time waveforms of a lexical-tone chimera (left column), an FFR recording (middle column), and data analysis. This lexical-tone chimera consisted of the envelope of the /yi^4/ Mandarin tone and the fine structure of the /yi^2/ Mandarin tone, both filtered at 4 filter banks.

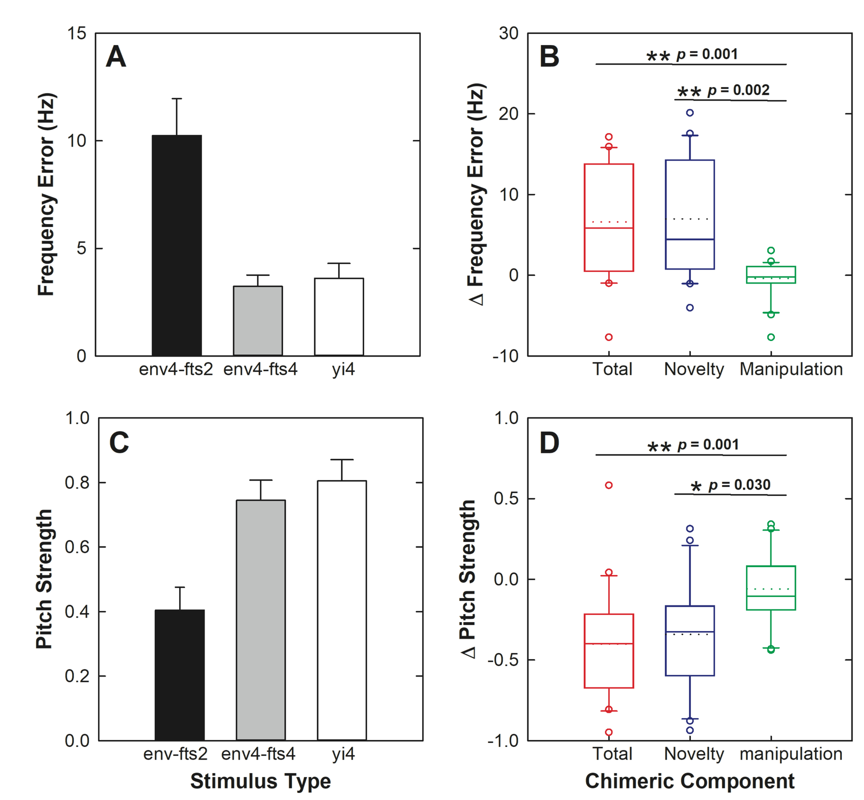

Figure 3 Group means ± 1 standard error of the *Frequency Errors* (Fig. 3**A**) and *Pitch Strengths* (Fig. 3**C**) derived from frequency-following responses (FFRs) evoked by the env4-fts2 chimera (black vertical bars), env4-fts4 chimera (gray vertical bars), and the original /yi^4/ Mandarin stimulus (white vertical bar). The total (red box and whiskers), chimeric novelty (blue box and whiskers), and signal manipulation (green box and whiskers) effects are plotted in the same panel for comparison (Fig. 3**B** *Frequency Error*, Fig. 3**D** *Pitch Strength*). Δ *Frequency Error* = difference in *Frequency Errors*. Δ *Pitch Strength* = difference in *Pitch Strengths*.

DISCUSSION

- For *Reaction Time*, a two-way ANOVA demonstrated statistical significances on the differences of the participants' reaction times in response to the env4-fts2 chimeras, env4-fts4 chimeras, and original /yi^4/ Mandarin tone for the filter bank ($p < 0.001$, $F = 9.349$, power = 0.977) and effect type ($p < 0.001$, $F = 54.242$, power = 1.000) factors as well as the interaction between the two factors ($p = 0.005$, $F = 6.950$, power = 0.898).
- For *Frequency Error*, a one-way ANOVA demonstrated a statistical significance for the effect type factor ($p < 0.001$, $F = 18.225$, power = 0.980).
- For *Pitch Strength*, a one-way ANOVA demonstrated a statistical significance for the effect type factor ($p = 0.001$, $F = 21.502$, power = 0.947).
- The analysis of variances (ANOVA) and post hoc Greenhouse-Geisser procedures demonstrated that the differences observed in the participants' reaction times and FFR measurements were attributed primarily to the chimeric novelty effects, but not the signal manipulation effects.
- Both the behavioral and electrophysiological measurements provide consistent and corroborative evidence, separating the effects of chimeric novelty from the effects of signal manipulation.
- These findings can be useful in assessing neuroplasticity of the human brain and improving speech-processing strategies for hearing aid and cochlear implant recipients.

ACKNOWLEDGMENTS

The authors thank the participants who enrolled in this study.

REFERENCES

Hou, L., & Xu, L. (2018). Role of short-time acoustic temporal fine structure cues in sentence recognition for normal-hearing listeners. The Journal of the Acoustical Society of America, 143(2), EL127. https://doi.org/10.1121/1.5024817

Hu, G., Determan, S. C., Dong, Y., Beeve, A. T., Collins, J. E., & Gai, Y. (2020). Spectral and Temporal Envelope Cues for Human and Automatic Speech Recognition in Noise. Journal of the Association for Research in Otolaryngology: JARO, 21(1), 73–87. https://doi.org/10.1007/s10162-019-00737-z

Jeng, F.-C., Lin, C.-D., Sabol, J. T., Hollister, G. R., Chou, M.-S., Chen, C.-H., Kenny, J. E., & Tsou, Y.-A. (2016). Pitch perception and frequency-following responses elicited by lexical-tone chimeras. International Journal of Audiology, 55(1), 53–63. https://doi.org/10.3109/14992027.2015.1072774

Kraus, N., & Nicol, T. (2015). Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. Trends in Neurosciences, 28(4), 176–181. https://doi.org/10.1016/j.tins.2005.02.003

Oxenham, A. J. (2018). How We Hear: The Perception and Neural Coding of Sound. Annual Review of Psychology, 69, 27–50. https://doi.org/10.1146/annurev-psych-122216-011635

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. Nature, 416(6876), 87–90. https://doi.org/10.1038/416087a

Warnecke, M., Peng, Z. E., & Litovsky, R. Y. (2020). The impact of temporal fine structure and signal envelope on auditory motion perception. PloS One, 15(8), e0238125. https://doi.org/10.1371/journal.pone.0238125