Neural processing of phonetic and talker information in a tone language: An fMRI study

Caicai Zhang[1, 2], *Kenneth R. Pugh*[3,4,5], *W. Einar Mencl*[3,5], *Peter J. Molfese*[3], *Stephen J. Frost*[3], *James S. Magnuson*[3 ,4], *Gang Peng*[1, 2], and *W. S-Y. Wang*[1, 2]

[1]CUHK-PKU-UST Joint Research Centre for Language and Human Complexity;
[2]Department of Linguistics and Modern Languages, The Chinese University of Hong Kong;
[3]Haskins Laboratories; [4]Department of Psychology, University of Connecticut; [5]Department of Linguistics, Yale University

Haskins Laboratories

INTRODUCTION

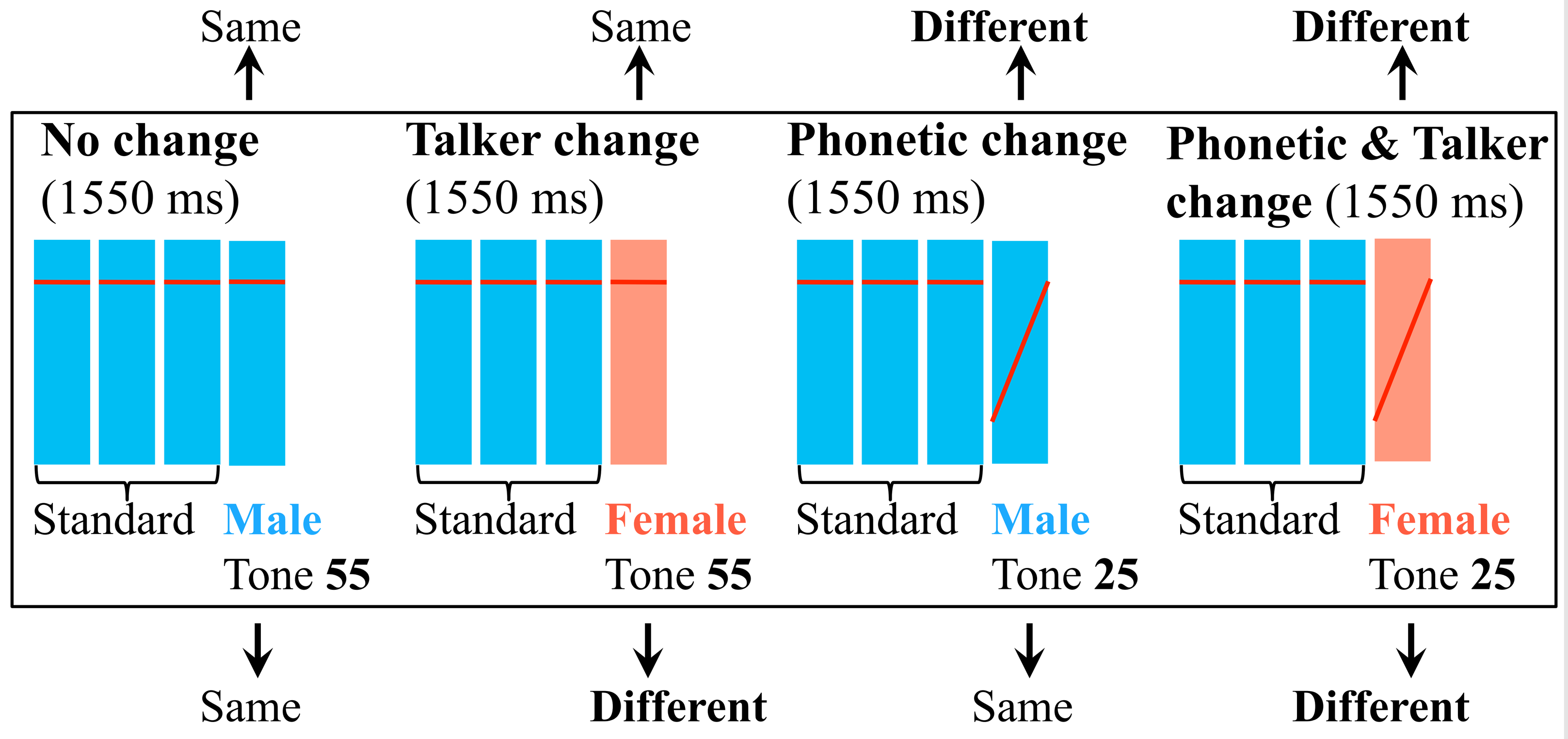
- **1. Interdependence of phonetic and talker processing** (Green, Tomiak & Kuhl, 1997; Kaganovich et al., 2006; Mullennix & Pisoni, 1990; Mullennix, Pisoni & Martin, 1989).
- Variability of talker identity (e.g. female/male) in speech stimuli interferes with phonetic classification (e.g. b/p);
- Variability in phonetic category (e.g. b/p) in speech stimuli interferes with talker classification (e.g. female/male), but to a less degree.
- Perceptual encoding of phonetic representations from acoustic signals depends on processing of talker information.
- **2. Neural locus of such interdependence**
- Integral ***perceptual processing*** of common acoustic parameters
  Posterior superior temporal gyrus/sulcus (STG/STS) that is activated in speech recognition tasks, is sensitive to vocal tract length change that differentiates talker identity/size (von Kriegstein et al., 2007; 2010; Kreitewolf, Gaudrain & von Kriegstein, 2014);
- Integral ***representation***
  Neural representation of real words stored in **left middle temporal gyrus** (MTG) could be talker-specific (Chandrasekaran, Chan & Wong, 2011; von Kriegstein et al., 2003);
- **3. This fMRI study**
- We investigated the integral perceptual processing of **fundamental frequency** (F0) in a tone language, where F0 distinguishes phonetic categories and correlates with talker identity.

METHODS

- ***Stimuli change*** (No change, Talker change, Phonetic change, Phonetic & Talker change) × ***Task*** (Phonetic & Talker judgment)
- **Stimuli:**

Male	Tone 55	Male	Tone 25
Female	Tone 55	Female	Tone 25

Task: Phonetic same/different judgment

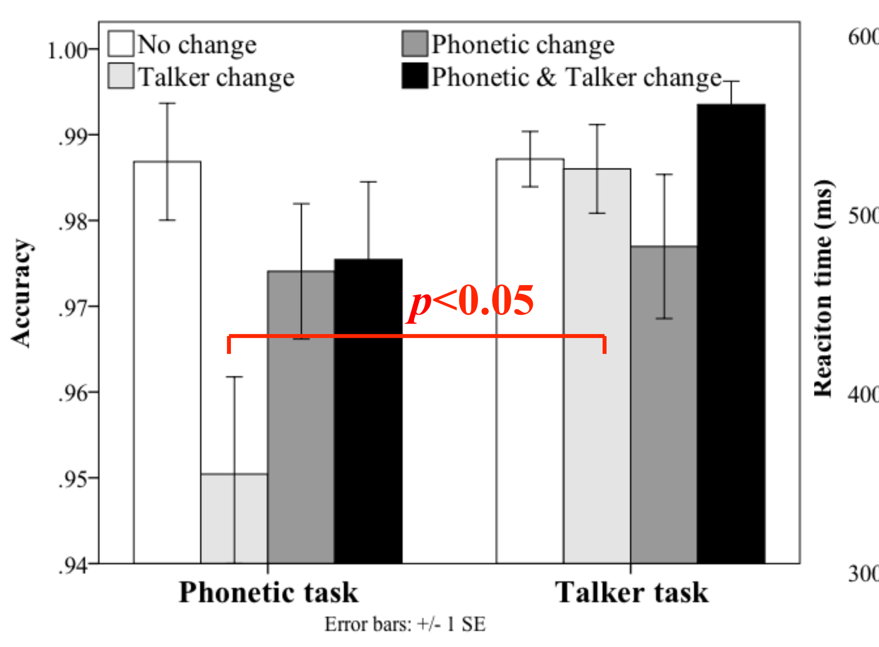


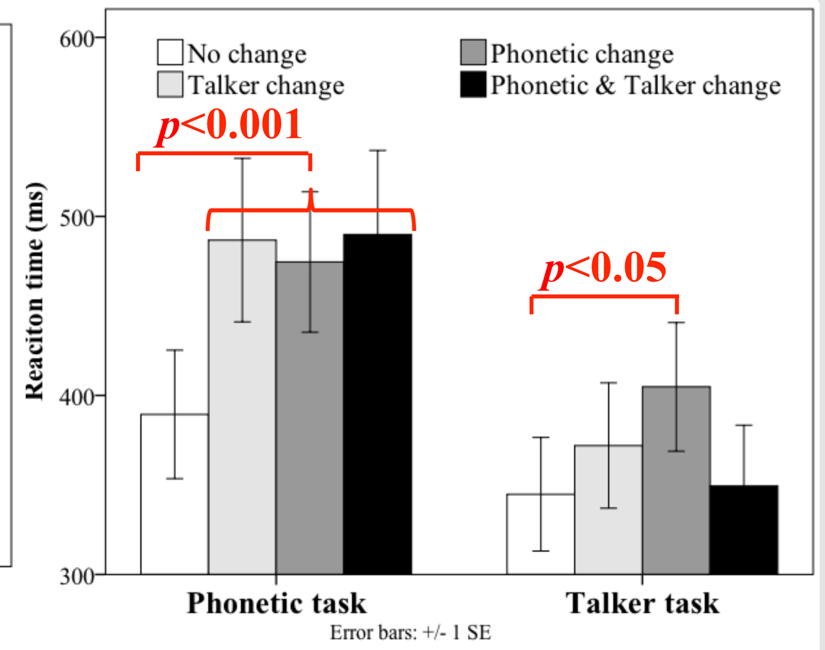
Task: Talker same/different judgment

- **Event-related fMRI design**
- Within a block, 4 types of trials pseudorandomized;
- Alternating blocks of phonetic task and talker task;
- Order of 8 blocks (4 blocks × 2 tasks) counterbalanced across subjects.
- **Subjects**: 18 right-handed Cantonese subjects (12 F; 21.4 yr ± 1.13)
- Siemens 3T scanner (TR=2s)

RESULTS

- **1. In-scanner behavioural performance**

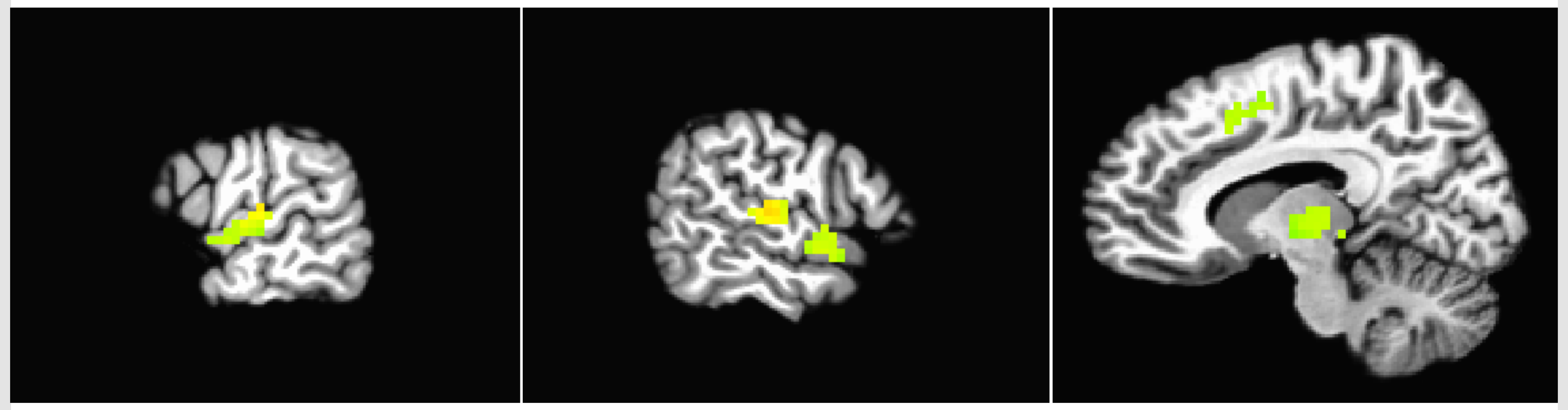




- **2. fMRI cluster analysis** (uncorrected $p<0.001$, FWE corrected $p<0.05$)

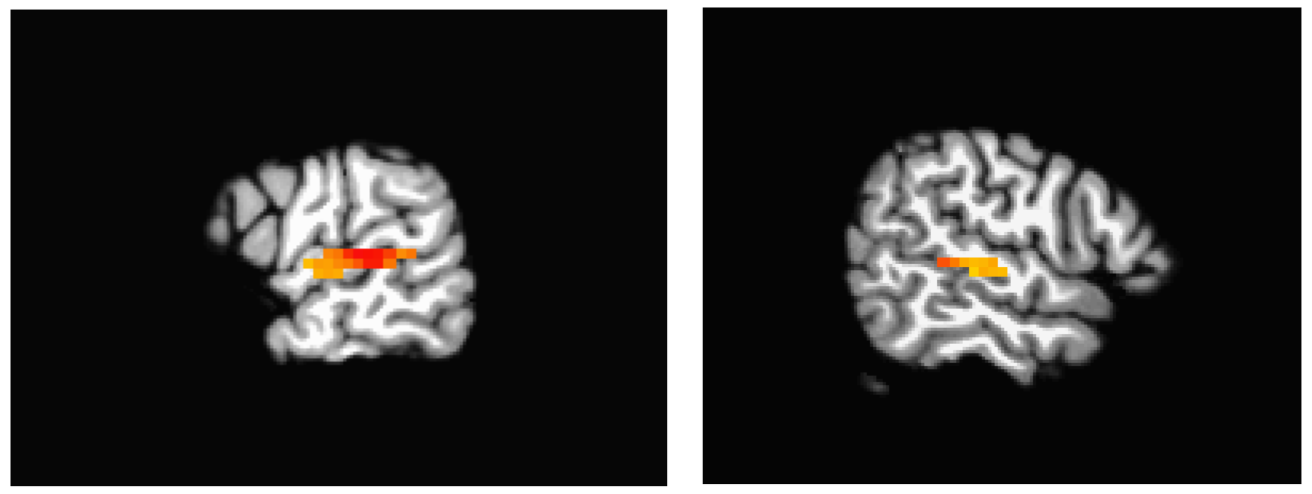
(1) *Phonetic change* vs. No change

- **Phonetic** task: No effect;
- **Talker** task: **Bilateral STG**, R thalamus, R cingulate gyrus, L inferior frontal gyrus, R middle frontal gyrus.

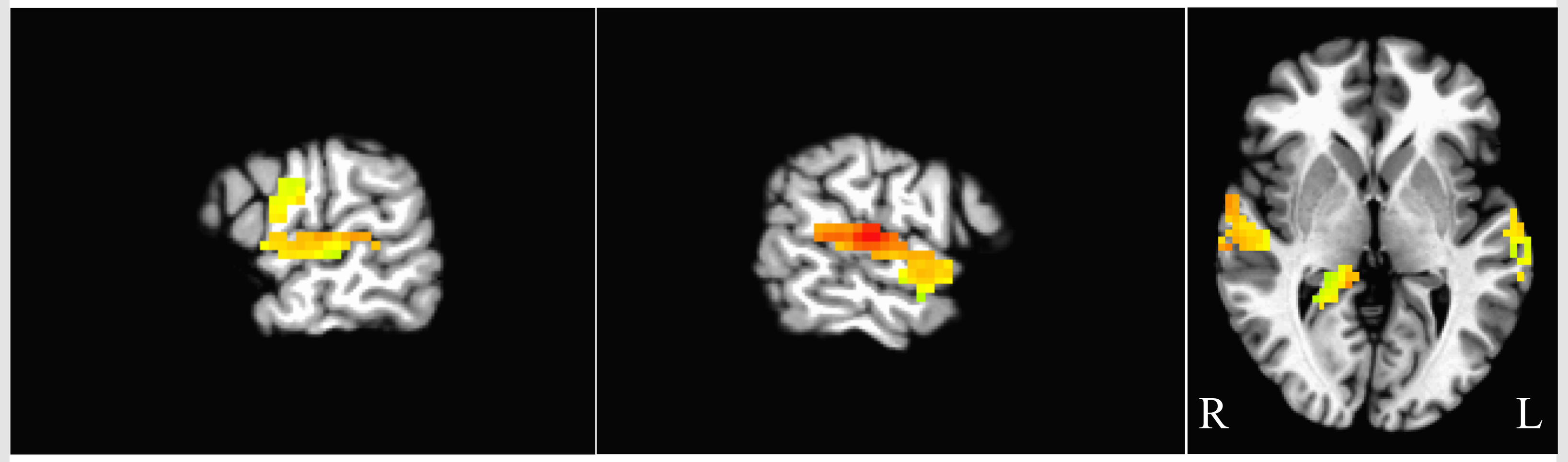


(2) *Talker change* vs. No change

- **Phonetic** task: **Bilateral STG**;



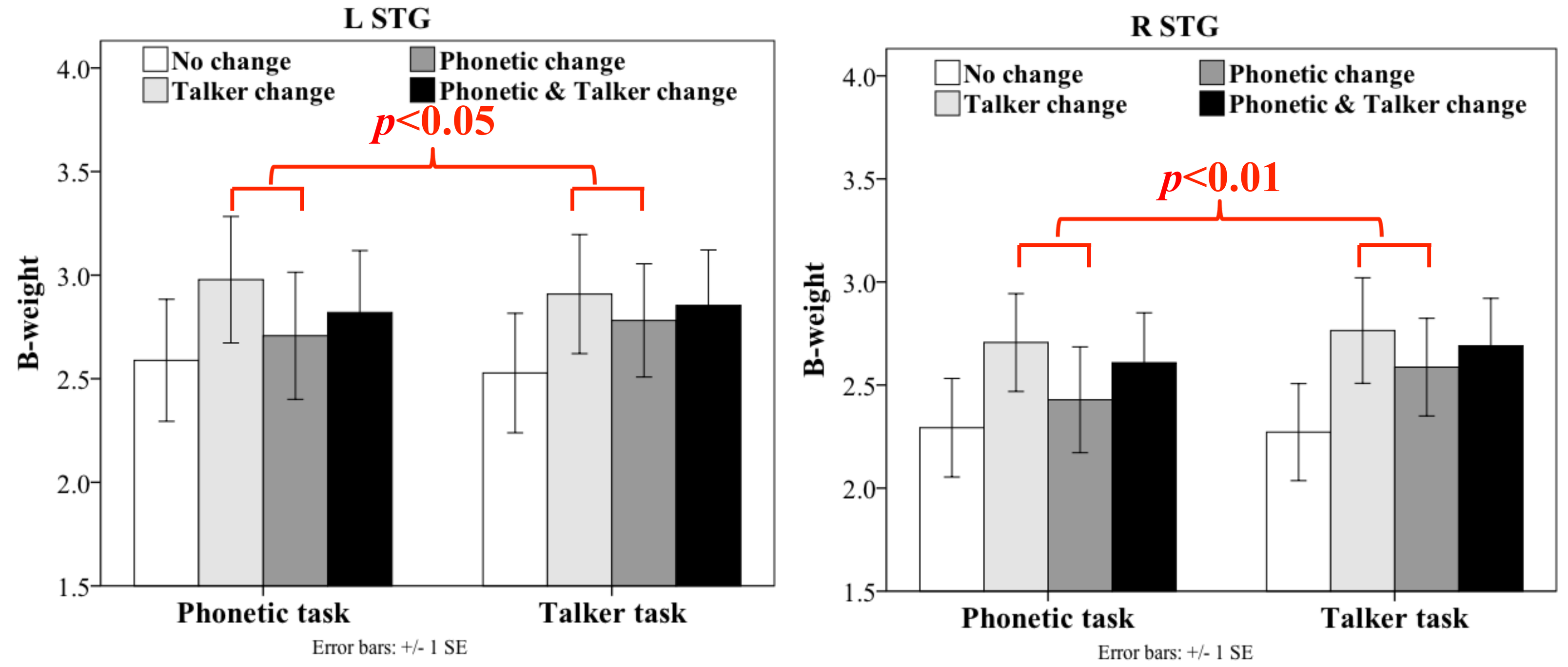
- **Talker** task: **Bilateral STG**, L declive, L postcentral gyrus, **R parahippocampal gyrus**.



(3) *Phonetic & Talker change* vs. No change

- **Phonetic** task: L STG;
- **Talker** task: R STG.

- **3. fMRI ROI analysis (STG)**



DISCUSSION

- 1. **Bilateral STG** activated in response to stimuli with *phonetic* changes in the *talker* discrimination task, and to stimuli with *talker* changes in the *phonetic* discrimination task.
  - It confirms the role of STG in ***integral processing*** of phonetic and talker information indexed by F0 in tone languages (Kreitewolf, Gaudrain & von Kriegstein, 2014);
  - No evidence for MTG involvement (presumably due to task influence).
- 2. **Left** and **right STG** weighted differentially in linguistic (phonetic discrimination) and non-linguistic (talker discrimination) tasks in response to stimuli with phonetic and talker changes.
- 3. **Right parahippocampal gyrus** activated to stimuli with talker changes in the talker discrimination task.
  - On-line learning of talker-related acoustic information.

REFERENCES

- Chandrasekaran, B., Chan, A. H. D., & Wong, P. C. M. (2011). Neural processing of what and who information in speech. *Journal of Cognitive Neuroscience*, *23*(10), 2690–2700.
- Green, K. P., Tomiak, G. R., & Kuhl, P. K. (1997). The encoding of rate and talker information during phonetic perception. *Perception & Psychophysics*, *59*(5), 675–692.
- Kaganovich, N., Francis, A. L., & Melara, R. D. (2006). Electrophysiological evidence for early interaction between talker and linguistic information during speech perception. Brain Research, 1114(1), 161–172.
- Kreitewolf, J., Gaudrain, E., & von Kriegstein, K. (2014). A neural mechanism for recognizing speech spoken by different speakers. *NeuroImage*, *91*(0), 375–385.
- Mullennix, J. W., & Pisoni, David B., (1990). Stimulus variability and processing dependencies in speech perception. *Perception and Psychophysics*, *47*, 379–390.
- Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *The Journal of the Acoustical Society of America*, *85*(1), 365–378.
- Von Kriegstein, K., Eger, E., Koeinschmidt, A., & Giraud, L. A. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research*, *17*, 48–55.
- Von Kriegstein, K., Smith, D. R. R., Patterson, R. D., Ives, D. T., & Griffiths, T. D. (2007). Neural representation of auditory size in the human voice and in sounds from other resonant sources. *Current Biology*, *17*(13), 1123–1128.
- Von Kriegstein, K., Smith, D. R. R., Patterson, R. D., Kiebel, S. J., & Griffiths, T. D. (2010). How the human brain recognizes speech in the context of changing speakers. *The Journal of Neuroscience*, *30*(2), 629–638.

Acknowledgements: This study is supported in part by grants from the National Basic Research Program of the Ministry of Science and Technology of China (973 Grant: 2012CB720700), and Research Grant Council of Hong Kong (GRF: 455911). Thanks to Mr. Ivan Zou for help with data collection, and to members of Centre for Language and Human Complexity for useful discussions.