



**Acoustics'08  
Paris**  
**June 29-July 4, 2008**  
[www.acoustics08-paris.org](http://www.acoustics08-paris.org)

## Linguistic versus non-linguistic processing of speech prosody in dichotic listening

Ritsu Kanamura<sup>a</sup> and Satoshi Imaizumi<sup>b</sup>

<sup>a</sup>Graduate School of Comprehensive Scientific Research, Prefectural Univ. of Hiroshima at Mihara, 1-1 Gakuen-machi, 723-0053 Mihara, Japan

<sup>b</sup>The Prefectural Univ. of Hiroshima, 1-1 Gakuen-machi, 723-0053 Mihara, Japan  
[imaizumi@pu-hiroshima.ac.jp](mailto:imaizumi@pu-hiroshima.ac.jp)

Linguistic and non-linguistic information processing of speech prosody were studied using two dichotic listening tasks, a Word task and a F0 task. During the Word task, subjects were required to identify either right- or left-ear stimulus from two-syllable homophonic words presented with different pitch accents. During the F0 task, subjects were required to identify either right- or left-ear stimulus from F0 partials extracted from the words used in the Word task. The correct percent was high and RT was short for high familiarity words presented to the right ear rather than the others under the Word task, while no such differences were found under the F0 task. RT of the Word task was shorter than the F0 task. These results suggest that the processing of linguistic speech prosody under dichotic listening conditions is based on the interactive auditory and linguistic neural resources with right-ear or left-hemispheric dominance, and are faster than auditory F0 pattern identification. The tasks developed can be used to detect central auditory processing disorders.

Keywords: Dichotic listening, F0 pattern, Pitch accent, Hemisphere dominance, Pitch perception

## 1 Introduction

Dichotic listening tasks are used for the detection of central auditory processing disorders [1, 2, 3]. Various linguistic stimuli such as digits, syllables [2, 3], words [1] and short sentences [1] have been used in these tests, although relationships between the central auditory processing system and the linguistic information processing system remain uncovered.

Particularly, dichotic listening tasks using speech stimuli with rich prosody should be reanalyzed since previous studies have provided conflicting results. Some indicate that the processing of pitch information associated with lexical items has left hemispheric dominance [4]; some others suggest that the processing of sound duration information has left hemispheric dominance regardless of the presence of linguistic information [5]; the processing of emotional prosody is related to both hemispheres [6]; the processing of temporal information is related to the right hemisphere and the processing of frequency information is related to the left hemisphere [7]; the processing of linguistic prosody and word-level prosody shows left hemispheric dominance [8]; the right hemisphere is closely related to slow information compared with the left hemisphere [9]; the processing of simple stimuli shows left hemisphere dominance and the more complex it is, the stronger is the dominance of the right hemisphere [10].

The aims of this study are to investigate the relationship between the central auditory processing system and the linguistic information processing system, and the difference of auditory information processing between the hemispheres. Two dichotic listening tasks, a Word task and a F0 task, were conducted. The stimuli of the Word task were pairs of two-syllable homophonic words whose pitch accents differ from each other, and those of the F0 task were the F0 partials extracted from the words used in the Word task.

## 2 Methods

### 2.1 Participants

Twenty-four normal-hearing, healthy, right-handed Japanese adults (12 males and 12 females aged 22.7 in average) participated. They spoke Tokyo dialect. All participants gave written consent before the experiment.

### 2.2 Tasks and Stimuli

There were two dichotic listening tasks, a Word task and a F0 task. These tasks were two-alternative forced-choice task. The listening subjects were required to select one visual stimulus from two alternatives corresponding to the auditory stimulus presented to one ear specified in advance. The auditory stimuli used in the Word task were 41 pairs of two-syllable homophonic spoken words, whose pitch accents differ from each other, for example, /a'me/ (high-low, rain) versus /ame'/ (low-high, a candy). These speech sounds were selected from the NTT database [11] which contains spoken words recorded by a woman in Tokyo dialect. The stimuli used in the F0 task were 50 pairs of sinusoidal sounds constructed by fundamental frequency of words used in the Word task. The visual stimuli used in the Word task were the orthographic expressions of the spoken words used as the auditory stimuli. The visual stimuli used in the F0 task were the line drawings of F0 contours used as the auditory stimuli.

The words with auditory word familiarity higher than 6 were defined as high familiarity words (HF) and the words with auditory word familiarity less than 6 were defined as low familiarity words (LF) referring to the NTT database [11].

We used Praat (version 4.6.31) [12] and Sony Sound Forge (version 7.0) for processing sounds. A head phone (TMR-IF630R, SONY) and a laptop personal computer (PC) (Dell INSPIRON I9200, Dell) were used for the tests.

### 2.3 Procedure

Participants sat in front of a PC with a head phone. The auditory and visual stimuli were presented at the same time. The auditory stimuli were dichotically presented through a headphone with most clearly perceiving level for each subject. The visual stimuli were presented as alternative choices on a PC screen in pairs.

Participants were required to identify an auditory stimulus presented to right or left ear (attended ear specified in advance), and select a visual stimulus corresponding to the auditory stimulus and to push a button corresponding to the correct choice as quickly and correctly as possible. For example, when they perceived an auditory stimulus corresponding to the left visual stimulus they had to push the left response button as quickly and correctly as possible.

The order of tasks and the attended ears were randomized so as to be counterbalanced between the participants.

## 2.4 Statistic analyses

The correct percent and the response time (RT) of the Word task and the F0 task were analyzed by a repeated measures ANOVA with four factors of Task (Word task vs. F0 task), Attended ear (Right ear vs. Left ear), Auditory word familiarity (High familiarity vs. Low familiarity) and Sex (Female vs. Male). Since Task showed significant interaction effects with Auditory word familiarity and Sex, the correct percent and the RT of each task were also analyzed by a repeated measures ANOVA with three factors of Attended ear (Right ear vs. Left ear), Auditory word familiarity (High familiarity vs. Low familiarity) and Sex (Female vs. Male).

## 3 Results

### 3.1 Differences between the Word task and the F0 task

For the correct percent, as shown in Fig.1, the main effects of Auditory word familiarity [ $F(1,22)=32.834$ ,  $p<.001$ ] and Sex [ $F(1,22)=14.438$ ,  $p=.001$ ] were significant. The words with high auditory word familiarity were judged more correctly than the words with low auditory word familiarity, and the female participants showed higher correct percent than the male participants.

The interaction effects between Task and Auditory word familiarity [ $F(1,22)=24.344$ ,  $p<.001$ ], and between Task and Sex [ $F(1,22)=5.423$ ,  $p=.030$ ] were also significant. As will be shown later, Auditory word familiarity effect is large in the Word task and Sex effect large in the F0 task.

For the RT, as shown in Fig.2, the main effects of Task [ $F(1,279)=39.729$ ,  $p<.001$ ], Attended ear [ $F(1,279)=14.123$ ,  $p<.001$ ], and Auditory word familiarity [ $F(1,279)=28.150$ ,  $p<.001$ ] were significant. The Word task resulted in shorter RT than the F0 task. The right-ear attention condition gave shorter RT than the left-ear attention condition. The higher the auditory word familiarity, the shorter was the RT.

The interaction effect of Task and Auditory word familiarity [ $F(1,279)=29.422$ ,  $p<.001$ ] was also significant. The higher the auditory word familiarity, the shorter was the RT, particularly in the Word task.

### 3.2 The Word task

For the percent correct, as shown in the right panel of Fig.1, the main effects of Attended ear [ $F(1,22)=5.022$ ,  $p=.036$ ] and Auditory word familiarity [ $F(1,22)=46.650$ ,  $p<.001$ ], were significant.

For the RT, as shown in the right panel of Fig.2, the main effects of the Attended ear [ $F(1,314)=20.487$ ,  $p<.001$ ], and Auditory word familiarity [ $F(1,314)=60.674$ ,  $p<.001$ ] were significant. The right-ear attention condition resulted in shorter RT than the left-ear attention condition. The higher the auditory word familiarity, the shorter was the RT.

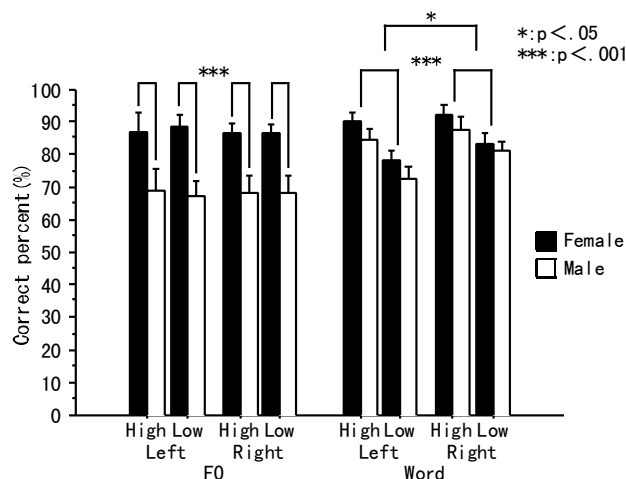


Fig.1 Correct percent of the Word task and the F0 task.

Correct percent with standard error bars. Filled bar: females; Empty bar: males. High: high familiarity words; Low: low familiarity words. Left: attend to left ear; Right: attend to right ear. F0: the F0 task; Word: the Word task.

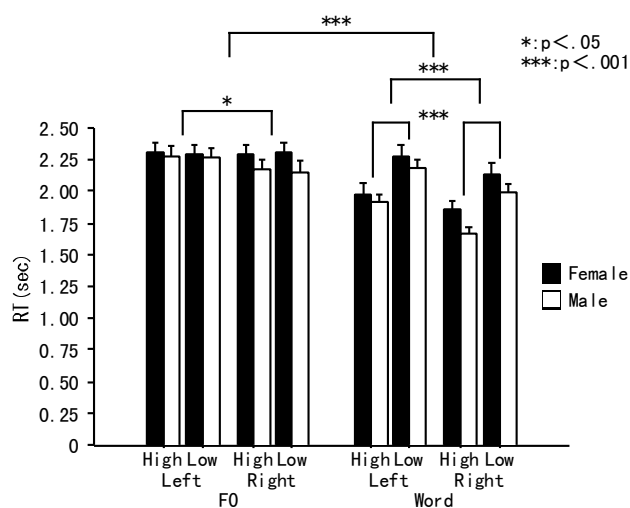


Fig.2 Response time of the Word task and the F0 task.

Response time with standard error bars. Filled bar: females; Empty bar: males. High: high familiarity words; Low: low familiarity words. Left: attend to left ear; Right: attend to right ear. F0: the F0 task; Word: the Word task.

### 3.3 The F0 task

For the correct percent, as shown in the left panel of Fig.1, the main effect of Sex [ $F(1,22)=14.852$ ,  $p<.001$ ] was significant. The female participants showed higher correct percent than the male participants.

For the RT, as shown in the right panel of Fig.2, the main effect of Ear [ $F(1,346)=4.134$ ,  $p=.043$ ] was significant. The right-ear attention condition resulted in shorter RT than the left-ear attention condition.

The significant interaction between Task and Auditory word familiarity corresponded to the difference showing that the correct percent was higher and the RT was shorter for the high familiarity words compared to the low familiarity words under the Word task, while no significant differences were found under the F0 task. The significant interaction between Task and Sex corresponded to the

contrast showing that the correct percent of the female participants was higher than the male participants under the F0 task, while there were no significant differences under the Word task.

## **4 Discussion**

In the Word task, the correct percent was higher and the RT was shorter for the high familiarity words compared to the low familiarity words. Moreover the correct percent was higher and the RT was shorter under the right ear attention condition compared to the left ear attention condition. It is known that the auditory word familiarity has an significant effect on the accuracy and speed of perception of words under various cognitive judgement tasks related to phonological and orthographic expressions of words and their semantics. The above-mentioned results suggest that the accuracy and speed of the perceptual judgement of spoken words for which only accent types were different from each other were affected by the auditory word familiarity and the attended ear. The correct percent was higher and the RT was shorter for the high familiarity words compared to the low familiarity words under the Word task, while no significant differences were found under the F0 task. The right-ear attention condition resulted in higher percent correct and shorter RT than the left-ear attention condition under the Word task.

One highly possible interpretation of these results is that the linguistic information processing system interferes with the auditory processing under the dichotic listening of spoken words which have only prosodic contrasts, and the whole process including linguistic and auditory processes are carried out with right-ear or left-hemisphere dominance. Furthermore, since the RT was significantly shorter under the Word task than the F0 task, it is also suggested that interaction between the linguistic and auditory processing systems accelerates the processing speed. Although the F0 contour has to be processed to judge the pitch accent type of the stimulus, the processing speed is higher when the F0 contour is processed in connection with linguistic information, that is, the orthographic expressions rather than the line drawings of the F0 contours in this study.

While the only difference between the Word task and the F0 task was whether the F0 contrasts are processed in connection with linguistic information or not, significant differences were found between the two tasks. The processing of pitch information in connection with linguistic information showed significant left hemisphere dominance with faster processing speed and with higher accuracy compared with when the same pitch information is processed without any connection with linguistic information. These results agree with previous studies that showed left hemisphere dominance in the processing of pitch pattern information associated with lexical items [4], and with dichotic listening studies using syllables and sentences [13] which showed that the performance of the right ear was higher than that of the left ear. Although the dominant hemisphere of auditory information processing has been suggested as dependent on the temporal resolution of target information [9], the present study suggests that the hemisphere dominance depends not only on the temporal characteristics of bottom up auditory signals but also on the

degree of mutual interaction between the linguistic information processing system and the auditory system.

The possible benefits of the dichotic test proposed here lie in the medical screening of three types of difficulties in auditory information processing. If performance deterioration is found only in the Word task, it suggests difficulties in linguistic processing of pitch accents through dichotic listening. If the performance deterioration occurred in only the F0 task, it suggests difficulties in detecting the pitch or F0 patterns as non-linguistic signals through dichotic listening. Finally, if the performance deterioration occurred in both the tasks, it suggests difficulties in the non-linguistic and linguistic processing stages. The additional benefit of the dichotic test is that the hemisphere dominance can be estimated simultaneously. As indicated above, by using two dichotic listening tasks with the least number of stimuli, it is possible to detect several deficits in auditory and linguistic information processing system. So, we can reduce the number of tests for detecting central auditory processing disorders.

There was no significant difference between male and female participants on the correct percent and the RT under the Word task and on the RT under the F0 task. However, the correct percent of the female participants was significantly higher than the male participants under the F0 task. This result suggests that the female participants perceive the F0 pattern more correctly than the male participants, although no significant difference is found under the linguistic judgment in the Word task. A previous study showed that females could judge the speaker's intention conveyed in spoken phrases with emotional prosody more correctly than males [14]. The present results may suggest that the difference between male and female participants in the perception of speaker's intention from emotional prosody of speech is partially due to the sex-dependent differences in the initial stage of the auditory processing of prosodic characteristics of speech.

## **5 Conclusion**

Linguistic and non-linguistic information processing of lexical prosody were studied using two dichotic listening tasks. The results suggest that the processing of lexical prosody under dichotic listening conditions is based on the interactive auditory and linguistic neural resources with right-ear or left-hemispheric dominance, and are faster than auditory F0 pattern identification. The tasks developed here can be used to detect central auditory processing disorders.

## **Acknowledgments**

We would like to thank all the subjects who kindly agreed to participate in our research.

## References

- [1] Keith, R.W. "Development and standardization of SCAN-A: test of auditory processing disorders in adolescents and adults" *J Am Acad Audiol* 6, 286-292 (1995)
- [2] Neijenhuis, K., Snik, A. & van den Broek, P. "Auditory processing disorders in adults and children: evaluation of a test battery" *Int J Audiol* 42, 391-400 (2003)
- [3] Fuente, A. & McPherson, B. "Auditory processing tests for Spanish-speaking adults: an initial study" *Int J Audiol* 45, 645-659 (2006)
- [4] Sato, Y., Sogabe, Y. & Mazuka, R. "Brain responses in the processing of lexical pitch-accent by Japanese speakers" *Neuroreport* 18, 2001-2004 (2007)
- [5] Brancucci, A., D'Anselmo, A., Martello, F. & Tommasi, L. "Left hemisphere specialization for duration discrimination of musical and speech sounds" *Neuropsychologia* (2008)
- [6] Kotz, S.A., et al. "On the lateralization of emotional prosody: an event-related functional MR investigation" *Brain Lang* 86, 366-376 (2003)
- [7] Van Lancker, D. & Sidtis, J.J. "The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: all errors are not created equal" *J Speech Hear Res* 35, 963-970 (1992)
- [8] Arciuli, J. & Slowiaczek, L.M. "The where and when of linguistic word-level prosody" *Neuropsychologia* 45, 2638-2642 (2007)
- [9] Boemio, A., Fromm, S., Braun, A. & Poeppel, D. "Hierarchical and asymmetric temporal sensitivity in human auditory cortices" *Nat Neurosci* 8, 389-395 (2005)
- [10] Harel, S. & Nachson, I. "Dichotic listening to temporal tonal stimuli by good and poor readers" *Percept Mot Skills* 84, 467-473 (1997)
- [11] Amano Shigeaki, Kondou Tadahisa. NTT Database Series [Nihongo-no Goitokusei: Lexical Properties of Japanese] Volume 1: Word Familiarity Sanseidou, Tokyo (1999)
- [12] <http://www.fon.hum.uva.nl/praat/>
- [13] Hallgren, M., Johansson, M., Larsby, B. & Arlinger, S. "Dichotic speech tests" *Scand Audiol Suppl* 49, 35-39 (1998)
- [14] Imaizumi, S., Homma, M., Ozawa, Y., Maruishi, M., Muranaka, H. "Gender differences in emotional prosody processing -A fMRI study-." *Psychologia*, 47, 113-124 (2004)