

Proceedings of Meetings on Acoustics

Volume 19, 2013

<http://acousticalsociety.org/>

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication

Session 4pSCb: Production and Perception I: Beyond the Speech Segment (Poster Session)

4pSCb27. Perceptual integration of indexical information in bilingual speech

Charlotte Vaughn* and Susanne Brouwer

*Corresponding author's address: Northwestern University, Evanston, IL 60208, crvaughn@u.northwestern.edu

The present research examines how different types of indexical information, namely talker information and the language being spoken, are perceptually integrated in bilingual speech. Using a speeded classification paradigm (Garner, 1974), variability in characteristics of the talker (gender in Experiment 1 and specific talker in Experiment 2) and in the language being spoken (Mandarin vs. English) was manipulated. Listeners from two different language backgrounds, English monolinguals and Mandarin-English bilinguals, were asked to classify short, meaningful sentences obtained from different Mandarin-English bilingual talkers on these indexical dimensions. Results for the gender-language classification (Exp. 1) showed a significant, symmetrical interference effect for both listener groups, indicating that gender information and language are processed in an integral manner. For talker-language classification (Exp. 2), language interfered more with talker than vice versa for the English monolinguals, but symmetrical interference was found for the Mandarin-English bilinguals. These results suggest both that talker-specificity is not fully segregated from language-specificity, and that bilinguals exhibit more balanced classification along various indexical dimensions of speech. Currently, follow-up studies investigate this talker-language dependency for bilingual listeners who do not speak Mandarin in order to disentangle the role of bilingualism versus language familiarity.

Published by the Acoustical Society of America through the American Institute of Physics

INTRODUCTION

In the course of everyday life, bilinguals switch between languages not only when speaking but also when listening. For bilinguals, then, which language is being spoken can be thought of as an indexical category, as it is a pointer to an aspect of the communicative context. How this language-being-spoken category is perceptually related to other indexical categories, particularly talker-related indexical categories, e.g. the gender of a talker and the specific identity of a talker, is currently unknown. Discovering more about the internal structure of the indexical system, particularly in bilinguals, is important because this internal structure is likely to play a role in linguistic processing given that variation in indexical information (e.g. talker variability) affects linguistic processing in monolingual settings (cf. Mullennix & Pisoni, 1990).

The present research seeks to determine how different types of indexical information may be perceptually integrated or segregated across different listener groups. This is done by using the speeded classification paradigm (often called the Garner task after Garner, 1974) to manipulate the amount of variability in characteristics of the talker (gender in Experiment 1 and specific talker in Experiment 2) and in the language being spoken (Mandarin vs. English). In the task, participants are directed to pay attention to one particular aspect of a speech stimulus while ignoring any variation in the stimulus that is irrelevant to that classification. The stimuli are arranged into different stimulus sets (this paper will be concerned with the sets called orthogonal and control), each with differing amounts of variation in the dimension not being attended to in the task at hand. The measure of exactly how much one dimension interferes with the other is taken by subtracting a subject's performance in the control condition from their performance in the orthogonal condition, a measure often called "Garner interference." If one dimension of the stimulus cannot be ignored, the subject will have longer reaction times in the condition where the other dimension is variable (orthogonal) than when it is constant (control), and hence there will be Garner interference. The relationship between two dimensions is said to be symmetrical if the amount of Garner interference is the same in the two tasks, and asymmetrical if there is more interference for one task than the other.

This paper presents results of two experiments using the Garner task (speeded classification) with Mandarin and English stimuli, collected from two different groups of participants, English monolingual listeners and Mandarin-English bilingual listeners. These experiments will help determine (1) whether language-being-spoken is integral with or separable from two talker-related indexical properties, (2) if integral, whether that integration is symmetrical (both categories interfere with each other equally) or asymmetrical (one category interferes with the other more than vice versa), and (3) whether these patterns of dependency are similar for monolingual and bilingual listeners.

EXPERIMENT 1: LANGUAGE-BEING-SPOKEN VS. GENDER

Materials and Methods

Stimulus Materials and Talkers

Stimuli were taken from the Archive of L1 and L2 Scripted and Spontaneous Transcripts and Recordings (ALLSTAR, Bradlow, Ackerman, Burchfield, Hesterberg, Luque, & Mok, 2010). Stimuli were short, meaningful sentences originally developed for the Hearing in Noise Test (HINT, English version, Soli & Wong, 2008; Mandarin version, Wong, Liu, Han, Huang, & Soli, 2007). Sixty-four HINT sentences read in English and 64 read in Mandarin by two bilingual Mandarin-English talkers (with English as the second language), one male and one female, were used in this experiment.

Participants and Experiment Design

Thirty-six English-speaking monolingual undergraduates completed Experiment 1. Eighteen Mandarin-English bilingual subjects participated in Experiment 1.

Each participant completed three conditions, corresponding to different stimulus sets, or blocks (orthogonal, correlated, and control) for each of two dimensions of the stimulus classification (gender, male or female; and language, Chinese or English). In the orthogonal blocks, the two dimensions vary independently (e.g. Chinese sentences are presented in both the male and female voice, and the English sentences are also presented in both the male and female voice). Thus, in orthogonal blocks classification along one dimension requires ignoring variation in the other, irrelevant dimension. In the control blocks, only one dimension is presented (e.g. only Chinese

sentences are presented for gender classification) such that there is no irrelevant variation. In the correlated blocks, the two dimensions vary together (e.g. Chinese sentences are always presented in the male voice, and English sentences are always presented in the female voice). Thus, in correlated blocks variation in the unattended dimension correlates with variation in the classification dimension. Only the results from the orthogonal and control conditions will be discussed in this paper. The conditions were blocked by stimulus dimension, so that a participant received all three blocks for one classification dimension, and then all three blocks for the other classification dimension.

Procedure

Participants were seated in a sound attenuated booth equipped with a Mac Mini running Superlab 4.5, headphones, and a Cedrus button box. Stimulus sentences were presented one at a time over headphones, and participants were instructed to classify each sentence into one of two categories (depending on which task they were completing, this was male or female for the gender task or Chinese or English for language-being-spoken). They were asked to respond using the buttons on the button box as soon as they knew the answer; they did not need to wait until the sentence was over before responding. Participants saw 20 practice trials with feedback at the beginning of each task before moving on to the main experiment. The order of the buttons was counterbalanced across participants. Participants were given a short break between each condition.

Results

Only correct responses were included in the analysis of reaction times. Performance was above 98% correct on all conditions across tasks for both listener groups. In Figure 1, reaction times are given for each listener group's performance on (a) the gender task and (b) the language-being-spoken task, for each block.

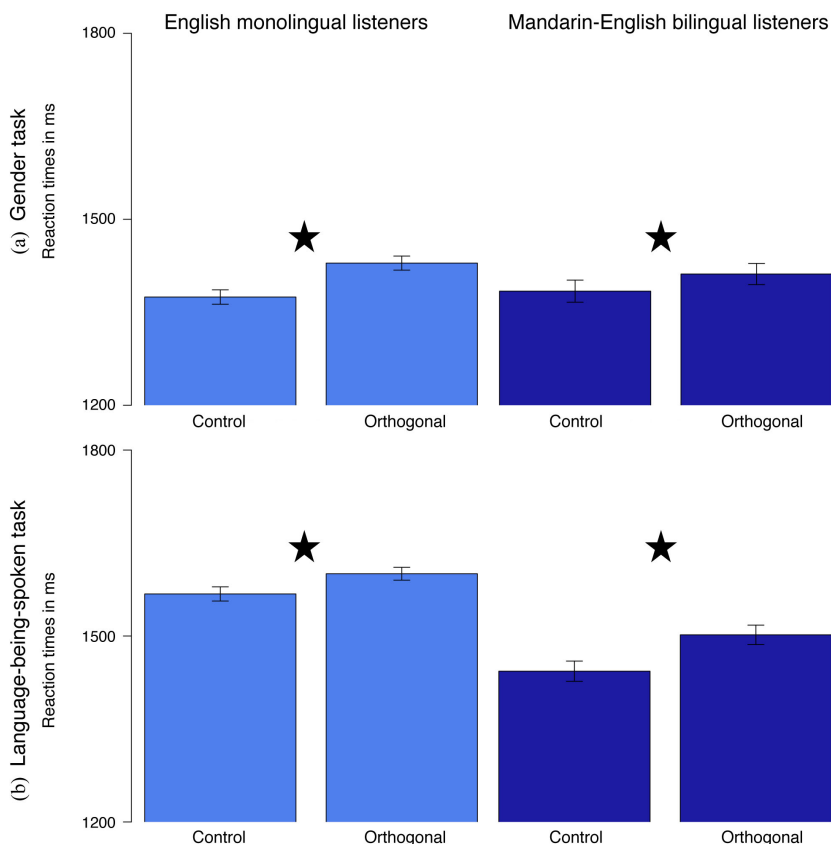


FIGURE 1. Results for Experiment 1. Row (a) shows reaction times for the gender task, and (b) shows reaction times for the language-being-spoken task. In each row, results from English monolingual participants are given on the left and Mandarin-English bilingual participants are given on the right. Stars reflect significant effects from several linear mixed effects models.

As can be seen in Figure 1, participants from both listener groups performed significantly faster on control blocks than orthogonal blocks in each task, demonstrating Garner interference. Within each listener group, the amount of Garner interference was symmetrical across tasks. Both listener groups performed gender classification significantly faster than language-being-spoken classification in the control condition. Several linear mixed effects models were fit to the log-transformed reaction times using the lme4 package in R (Bates, Maechler, & Bolker, 2011) and are the basis for the statistical significance results described above.

The results of Experiment 1 demonstrate that language-being-spoken is perceptually integrated with gender, as there is significant Garner interference in both tasks for both listener groups (i.e. listeners cannot ignore irrelevant variation in the unattended dimension). Symmetrical patterns of interference were found for both monolinguals and bilinguals. Experiment 2 pairs language-being-spoken with talker identity to test whether language-being-spoken shows similar patterns of integration with an indexical category having different properties than gender. For example, a large amount of information about gender can be conveyed in a relatively small number of acoustic features, with F0 as a major cue (Coleman, 1971), but differences between talker voices are conveyed in a wide variety of acoustic cues. Talker identity may therefore be processed at a deeper level than gender identity, potentially leading to different patterns of integration with language-being-spoken, as perceptual dimensions may show different interference patterns depending on their relative levels of processing (Wood, 1975).

EXPERIMENT 2: LANGUAGE-BEING-SPOKEN VS. TALKER

This experiment is designed to test the integration of language-being-spoken and another talker dimension, talker identity information, in the processing of bilingual speech.

Methods and Materials

Stimulus Materials and Talkers

The same 64 HINT sentences in Experiment 1 were used in Experiment 2. However, the female talker from Experiment 1 was replaced with a new male talker, making the two talkers in Experiment 2 both male Mandarin-English bilinguals.

Participants and Experiment Design

Twenty-eight English-speaking monolingual undergraduates completed Experiment 2 (none of whom had completed Experiment 1). Eighteen Mandarin-English bilingual subjects participated in Experiment 2. The experiment design was identical to Experiment 1, with the use of the talker category in Experiment 2 instead of the gender category in Experiment 1.

Procedure

The procedure for Experiment 2 is identical to that of Experiment 1, except that one of the stimulus dimensions was different; instead of classifying gender, participants classified talker, with both talkers sharing the same gender (in this case, male). The two male talkers were given the names “Wei” and “Li.” Thus, participants were instructed to classify each sentence into one of two categories (depending on which task they were completing, this was “Wei” or “Li” for the talker task or Chinese or English for language-being-spoken). Due to the nature of the talker classification task, the experimenter gave each participant an additional verbal instruction prior to the start of the talker task explaining that the participant should use the practice trials as a way to guess who is Wei and who is Li by trial-and-error.

Results

Only correct responses were included in the analysis of reaction times. Performance was above 96% correct on all conditions across tasks for both listener groups. In Figure 2, reaction times are given for each listener group’s performance on (a) the talker task and (b) the language-being-spoken task, for each block.

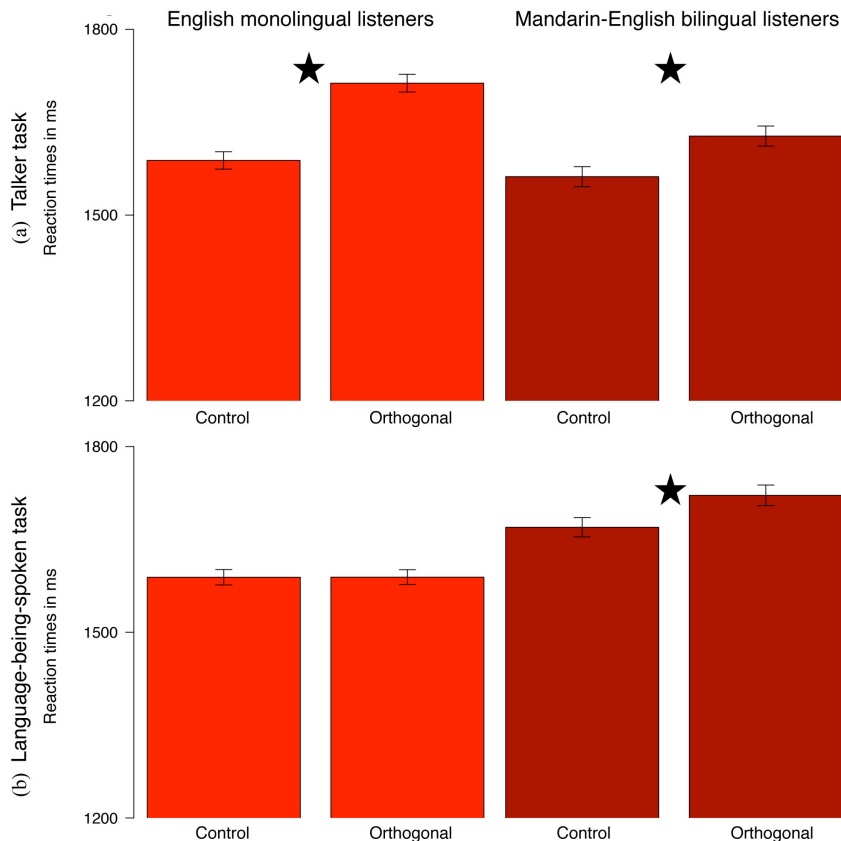


FIGURE 2. Results for Experiment 2. Row (a) shows reaction times for the talker task, and (b) shows reaction times for the language-being-spoken task. In each row, results from English monolingual participants are given on the left and Mandarin-English bilingual participants are given on the right. Stars reflect significant effects from several linear mixed effects models.

As can be seen in Figure 2, English monolingual participants showed Garner interference for the talker task, but not for the language-being-spoken task, resulting in an asymmetrical pattern of interference. That is, the English monolinguals were able to ignore irrelevant variation along the talker dimension when classifying the stimuli by language-being-spoken (Chinese or English), but they were unable to ignore irrelevant variation along the language-being-spoken dimension when classifying the stimuli by talker (Wei or Li). The Mandarin-English bilingual listeners showed Garner interference in both tasks, and the amount of Garner interference was symmetrical across tasks. English monolingual listeners performed both tasks equally quickly in the control condition, while Mandarin-English bilinguals were significantly faster in the control condition when classifying talker than when classifying language-being-spoken. As in Experiment 1, several linear mixed effects models were fit to the log-transformed reaction times and are the basis for the statistical significance results described above.

The results of Experiment 2 demonstrate that language-being-spoken is not completely segregated from talker identity, as there is significant Garner interference in the talker task for both listener groups and in the language-being-spoken task for bilinguals. Patterns of interference differed between monolinguals and bilinguals.

CONCLUSIONS

Results for the gender vs. language-being-spoken classification (Experiment 1) showed a significant, symmetrical interference effect for both listener groups, indicating that gender information and language-being-spoken are processed in an integral manner. For talker vs. language-being-spoken classification (Experiment 2), language-being-spoken interfered more with talker than vice versa for the English monolinguals, but symmetrical interference was found for the Mandarin-English bilinguals. These results suggest that the language-being-spoken is not fully segregated from talker-related properties in processing.

Framed in terms of listener background, monolingual English listeners showed symmetrical interference for language-being-spoken and gender, but asymmetrical interference for talker and language-being-spoken, while

Mandarin-English bilinguals showed symmetrical interference for both pairs of indexical properties. Familiarity with both languages seems to promote an even balance of attention to these indexical properties. That these bilinguals exhibited more balanced classification along various indexical dimensions of speech could be attributable to two possible factors. First, it could be the case that their familiarity with both languages being tested enabled these listeners to be more “tuned in” to the indexical properties under investigation in these two languages. Alternatively, it could be the case that the experience of being bilingual enhances attention to these indexical properties in general, which would even apply when listening to an unfamiliar language. Currently, follow-up studies investigate the dependency between language-being-spoken and talker identity for bilingual listeners who do not speak Mandarin in order to disentangle the role of bilingualism versus language familiarity.

ACKNOWLEDGMENTS

The authors wish to acknowledge Ann Bradlow, Vanessa Dopker, Chun Liang Chan, Tyler Kendall, Emily Kahn, and Lindsay Valentino. Work supported by Grant R01-DC005794 from NIH-NIDCD.

REFERENCES

- Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using Eigen and Eigenfaces. R package version 0.999375-42, <http://CRAN.R-project.org/package=lme4>.
- Bradlow, A. R., Ackerman, L., Burchfield, L. A., Hesterberg, L., Luque, J. S., & Mok, K. (2010). ALLSTAR: Archive of L1 and L2 Scripted and Spontaneous Transcripts And Recordings, http://groups.linguistics.northwestern.edu/speech_comm_group/allstar/index.html.
- Coleman, R. O. 1971 . “Male and female voice quality and its relationship to vowel formant frequencies,” *J. Speech Hear. Res.* 14, 565–577.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Lawrence Erlbaum.
- Mullennix, J. W., & Pisoni, D. B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception & Psychophysics*, 47, 379- 390.
- Soli, S. D., Wong, L. L. N. (2008). Assessment of speech intelligibility in noise with the hearing in noise test. *Intl. J. Audiology* 47, 356-361.
- Wong, L. L. N., Liu, S., Han, N., Huang, V. M., & Soli, S. D. (2007). Development of two versions of the Mandarin Hearing In Noise Test (MHINT). *Ear & Hearing*, 28(2 Suppl.), 70-74.
- Wood, C. C. (1975). Auditory and phonetic levels of processing in speech perception: Neurophysiological and information-processing analyses. *Journal of Experimental Psychology: Human Perception and Performance*, 1(1), 3.