ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication
Session 5aSCb: Production and Perception II: The Speech Segment (Poster Session)

5aSCb3. Effects of phonemic variability and language dominance on Canadian French-English bilinguals? perception of French vowels in various phonological contexts
Allison A. Johnson* and Franzo Law II

*Corresponding author's address: Communication Sciences and Disorders, University of Wisconsin-Madison, Madison, WI 53706, aajohnson4@wisc.edu

Phonemic variability can cause productions within a category of one language to be mapped onto multiple categories in a different language (Escudero, 2009). For example, French productions of /ɛ/ are labeled both as /e/ and /ɛ/ by monolingual English listeners, likely due to spectral variability in production of French /ɛ/ (Strange, Levy, & Law, 2009). The present study examines how phonemic variability in Canadian French (CF) affects the perceptual tendencies of bilingual Canadian French-English listeners with varying levels of language dominance. Vowel productions by monolingual CF speakers were used in a modified identification task (Law, 2011). The vowels were word-final in several phonological contexts (preceded by labial, coronal, and back consonants; followed by labial and coronal consonants) in real and nonwords embedded in carrier phrases. A subset of these vowels /ɛ-ɛ-i-y/ was analyzed for duration, vocalic midpoint, and formant trajectories to examine the relationship between phonemic variation (as a function of phonological environment) and the perception results by bilinguals dominant in either CF or Canadian English (CE). We predict that the performance of CE-dominant listeners will vary in terms of speed and accuracy based on how similar each token is to CE vowel category expectations.

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

The articulatory gestures of a speaker produce an acoustic signal that is highly variable—subtle differences occur naturally even when the same speaker repeats the same word. This variability may not impact phonemic perception by a native speaker, but may significantly affect the perception of nonnative listeners. For example, Canadian French (CF) speakers tend to produce vowels that are spectrally higher (i.e., have lower first formant values) than Canadian English (CE) productions. In addition, CF vowels have increased within-category duration variability compared to CE productions because vowel length is not phonologically contrastive in CF. These phonetic differences lead CE listeners to perceive multiple productions of a single CF vowel as equally good instances of two different CE vowel categories, depending on the proximity to native productions in multiple acoustic dimensions (Escudero & Polka, 2003). These patterns may be the result of CE listeners’ application of “cue weighting strategies” (Escudero & Polka, 2003), or possibly due to overlearned processes that detect critical, higher-order speech parameters through language-specific selective perceptual routines (Strange, 2011).

Several theoretical models have been proposed to describe cross-linguistic perceptual patterns. Some models show how pairs of nonnative phonemes with varying levels of discriminability are mapped onto native phonemic categories (Best, 1994, 1995). This Perceptual Assimilation Model (PAM) also explores perceptual learning, the process by which individuals become increasingly attuned to detecting higher-order invariants in their native language and articulatory/acoustic patterns that critically distinguish language-specific segmental features. Through linguistic experience, a person learns to selectively attend to language-specific acoustic/articulatory cues and learns to ignore irrelevant ones (Strange, 2011; Werker & Curtin, 2005).

Most models focus on perception of nonnative sounds by a naïve listener; however, the perceptual patterns of bilingual listeners are more complex, as they must suppress knowledge of one language while actively attending to relevant information and communicating in the other. By focusing on experienced bilinguals, it is possible to extend the hypotheses of cross-language speech perception models. In addition, these models are based on experiments using discrimination paradigms comparing perception in a pairwise fashion; the present study employed an identification task using all possible vowels phonemic in the language.

OBJECTIVE

The goal of the present study is to test some of the hypotheses posited by the current theoretical models of cross-language speech perception. As language dominance may affect a listener’s sensitivity to natural variability, the present study focuses on the effects of phonetic variability in vowel productions on perception by CF-CE bilinguals living in Quebec. It is hypothesized that the CE-dominant group will be faster and more accurate at identifying productions with variability that result in more CE-like vowel category expectations.

A subset of the vowels used in the perception study was selected to be analyzed for acoustic/articulatory variation. The vowels /e, ε/ were chosen because they occur in both languages, but are realized differently both phonetically and phonotactically. As previously mentioned, CF vowels, such as /e, ε/, are often produced higher spectrally, compared to their CE counterparts. Consequently, CF vowels are sometimes labeled as spectrally higher CE vowels (Escudero & Polka, 2003). In addition, both /e/ and /ε/ can occur word-finally in French, but not in English.

Vowels /i/ and /y/ were selected to explore further some of the perceptual patterns observed in the perception task. Although it was hypothesized that the CF-dominant group would be able to identify all CF vowels equally well, both groups were particularly slow in identifying /y/ compared to other vowels, and there was no significant difference in speed or accuracy between groups. Although both groups were faster at identifying /i/ than most other vowels, the CF-dominant group made more errors than expected on /i/.

The acoustical analysis of these tokens was used in conjunction with the results described in Law (2011a, b) to examine the relationship between phonemic variation (as a function of phonological environment) and the perceptual patterns exhibited by bilinguals with varying levels of language dominance.
METHODOLOGY

Methods: Production

Participants

The tokens were produced by two monolingual CF speakers: a 23-year-old male born in Jonquière, Canada, and an 18-year old female born in Montréal, Canada. Both participants were residing in Montréal at the time of testing and were recruited from the Quebecois community via a francophone newspaper, a community website, or a college volunteer email post, and reported Quebecois French as their spoken language on a background questionnaire.

Procedure

Recordings were obtained by a second-language speaker of French in a sound-treated room at McGill University. Sessions were conducted in French and lasted approximately two hours (including breaks), after which participants were compensated for their time. The token words with the target vowel were produced in word-final position with three preceding contexts (labial, coronal, and back stop consonants) in both real and nonsense mono- or disyllabic words in an effort to elicit more ecologically-valid productions. The words were embedded in two different carrier phrases where the following consonant was either labial (“Je dis le mot po uring lui”) or coronal (“Je dis le mot tout le temps”). Two repetitions of each sentence were recorded by each participant, read in random order blocked by word type (real or nonsense). The last sentence of each list was identical to the first and not used in analysis to minimize effects of list-final prosodic patterns. For the present study’s analysis, 48 tokens for each vowel were analyzed: 2 speakers × 2 preceding contexts × 3 following contexts × 2 word types × 2 repetitions.

The experimenter recorded the participants using a Marantz PMD671 solid-state recorder and Shure SM81 cardioid condenser microphone with a pop filter. The built-in 18 dB/octave roll-off high pass filter of the microphone was also used to minimize background noise below 100 Hz. The recording was digitized at 44.1 kHz sampling rate, 16-bit resolution (for more information on the protocol, see Law II & Strange 2008, Submitted).

Acoustical Analysis

The tokens were analyzed in a quiet room on a Dell Optiplex 755 Desktop using Praat for Windows 7 and Sennheiser HD 280 Pro circumaural headphones. The vowel tokens were analyzed in Praat to measure vocalic midpoint, duration, and formant trajectories. To determine vowel boundaries, both the waveform and spectrogram were consulted in addition to perceptual judgment. Vowel onsets were defined as the beginning of the first complete periodic cycle following the stop release or aspiration. Vowel offset was determined by the last complete cycle before significant visual decrease of energy in the upper formants.

Methods: Perception

Participants

Thirty-two bilingual adults were recruited from the Quebecois community to participate. The participants were divided into two equal groups of 6 men and 10 women as determined by self-reported language dominance. All participants underwent hearing screening and several measures to establish relative levels of French proficiency.

Procedure

The experiment was presented to the participants, using a Dell Studio Hybrid PC with Paradigm software and LG W2361VG 23”, 1920 x 1080 monitor, set up in a sound-attenuated chamber. The sentences were presented at approximately 80 dB SPL through Sennheiser HD280 Pro headphones, while the corresponding sentence was simultaneously presented in orthographic form on the computer screen with a blank where the target word would be. The participants used a Microsoft SideWinder laser mouse, calibrated to move at a speed of 1600 dpi. The task was described to the participants as a rhyming task; they were to choose a response button on the screen that rhymed with the target words. The response buttons were real French words that would be transcribed as one of the eight vowels that can occur word-finally (e.g., the response alternative “ai” would be transcribed as /e/). The response
buttons were arranged in a semicircle and the order of the buttons was reversed for half of the participants in each group to counterbalance for effects due to motor control. Two familiarization blocks aided the participants in understanding the nature of the experiment and becoming acquainted with the layout of the response alternatives (for more information, see Law II, 2011a).

**Analysis**

Participants’ responses in the identification task were analyzed in terms of accuracy and a more sensitive measure, reaction time. A preliminary analysis was conducted in order to test the hypothesis that variances in the stimuli used might be correlated with the speed and accuracy in identifying the vowels by both CF-dominant and CE-dominant bilinguals. The formant measures at vocalic midpoint, as well as vocalic duration reported in Law & Strange (Submitted) were used. All formant frequencies were converted to a Bark psychoacoustic scale prior to analysis (Zwicker & Terhardt, 1980). Regarding the perception data, the analysis previously reported reaction time and percent correct collapsing across language group and vowel; in this analysis, the mean reaction time and percent correct were analyzed collapsing across stimulus presentation. Only the last 8 blocks were used, giving the participant 4 blocks to become acclimated to the task. The reaction time of each trial with a correct response for each participant was log-transformed to determine outliers, and then converted back to linear values. The average reaction time for control vowels was calculated and subtracted from the average reaction time for each stimulus type for each subject. These different score reaction times were used as an effort to control for overall subject variability in speed. Percent correct for each stimulus was calculated across participants within a language group by speaker for all twelve blocks.

**RESULTS**

In order to explore associations between production variation and perceptual patterns, Spearman rank-order correlations were calculated, quantifying the association between vocalic duration and F1-F2-F3 measures (in Bark) to speed and accuracy in vowel identification for vowels /e, ε, y/. A non-parametric test was used, given the ceiling effects in overall accuracy and the non-normal distribution of the reaction-time difference scores. Correlations were calculated separately for each language group, for each speaker, and each vowel. This was deemed a plausible strategy for an exploratory analysis, given that the patterns were expected to be different for each vowel, the speakers may have different productions strategies for each vowel, and the stimuli were presented blocked by speaker.

Results indeed confirm that certain variations in production were correlated with differing patterns of perception for both the CE-dominant and CF-dominant group. Previous analysis of the perceptual patterns led to the conclusion that the CE-dominant group was slower than the CF-dominant group in identifying /ε/ (Law II, 2011a, b). The present analysis elaborates on this finding. For Speaker A’s productions of /ε/, the CE-dominant group was faster at correctly identifying the vowel when F2 and F3 were higher [F2: $\rho = -0.438, p = 0.032$; F3: $\rho = -0.413, p = 0.045$]. They also tended to be more accurate in identifying this vowel when F3 was higher [$\rho = 0.402, p = 0.052$]. For Speaker B, again, faster responses were correlated with higher F3 values [$\rho = -0.442, p = 0.031$]. The CF-dominant group were as fast as baseline in identifying /ε/ (Law II, 2011a, b), but when they did misidentify a token produced by Speaker A, it was when F2 and F3 were relatively higher [F2: $\rho = -0.402, p = 0.051$; F3: $\rho = -0.488, p = 0.015$]. Similarly, the CF-dominant group was faster at identifying /ε/ tokens produced by Speaker B when produced with a lower F3 [$\rho = -0.48, p = 0.018$] and shorter duration [$\rho = -0.549, p = 0.005$]. The results of these correlations suggest that CE-dominant listeners prefer /ε/ to be produced higher F2 and F2 values, whereas CF-dominant listeners prefer a lower F3.

In comparing the acoustic measures to the speed and accuracy in identifying /ε/, were more accurate in identifying productions by Speaker B with a lower F3 [$\rho = -0.475, p = 0.019$] and with a relatively shorter duration [$\rho = -0.446, p = 0.029$]. None of the acoustic measures correlated with performance for the CF-dominant group in identifying /ε/. In comparing the correlations of /e-ε/ for the CE-dominant group, it appears that F2 and F3 may need to be maximally distinct in order to identify the contrast quickly and reliably. In addition, they may be attending to duration when identifying /ε/, which is a reliable cue in English, but not in French.

Finally, in identifying /y/, both groups were more accurate in identifying tokens produced by Speaker A with a lower F1 [CE-dom: $\rho = -0.468, p = 0.021$; CF-dom: $\rho = -0.551, p = 0.005$].
Discussion

Preliminary analysis suggests that there are modest correlations between production variability and perceptual accuracy and speed. Furthermore, the CF-dominant and CE-dominant groups have divergent patterns of perception. For the /ɛ-ε/ contrast, the CE-dominant group had a more reliable perceptual pattern in identifying stimuli that had more peripheral F2 and F3 distributions, and /ɛ/ with a shorter duration.

Future Directions

Further analyses will include additional acoustical measures, such as formant slope, formant measures at vowel onset, ratio between formants, as well as exploring the variation in /i/ productions. In addition, the collected measures of language proficiency, as well as analysis of a small corpus of the perceivers’ productions may help to understand perceptual patterns within a language group.

ACKNOWLEDGMENTS

This work was supported by NICHD grants F31DC008075 and T32 HD049899-06.

REFERENCES


Law II, F. (2011a). Perception of Canadian-French word-final vowels in lexical and morphosyntactic minimal pairs by Canadian English learners of French. The Graduate School and University Center - CUNY.


