MAPPING PERCEPTUAL DISTANCES BETWEEN SIBILANT AND PALATAL FRICATIVES

Corey Telfer
Dept. of Linguistics, University of Alberta, 4-32 Assiniboia Hall, Alberta, Canada, T6G 3E7, corey.telfer@ualberta.ca

1. INTRODUCTION

The perception of fricatives is not well understood, in part because they are generated with air turbulence, which complicates their articulatory and acoustic properties (e.g. Shadle 2012). Sibilant fricatives are especially difficult to characterize because, unlike other obstruents, "…place as well as manner cues are signaled primarily by the spectral structure of the segment itself" (Toda et al. 2010:343), rather than by the formant transitions. While acoustic models of sibilants such as those of Howe & McGowan (2005) and Toda et al. (2010) continue to improve our understanding of these speech sounds, the precise organizational principles behind their perception remain unknown.

Sibilant fricatives can be ordered along a one-dimensional continuum defined by the spectral mean, as seen in Figure 1 (after Boersma & Hamann 2008).
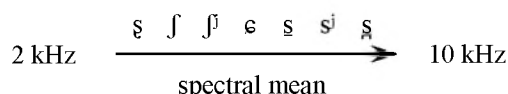
ʂ ʃ ʃʲ ɕ s̠ sʲ s̺

2 kHz ——————————→ 10 kHz
spectral mean

Figure 1. Continuum of seven sibilants based on increasing spectral mean.

The spectral mean is inversely proportional to the volume of the sublingual cavity, so as the place of articulation approaches the anterior of the mouth, the spectral mean increases. While this simple representation captures the basic facts, it fails to account for more subtle distinctions, such as Fujisaki & Kunisaki's (1978) finding that the Japanese alveolar fricative [s] is best modeled using a spectral distribution with two spectral peaks and one valley, rather than just a single peak.

This study investigated whether or not certain pairs of sibilant and palatal fricatives were more difficult to differentiate than others. This was done by presenting listeners with synthesized stimuli in an AX discrimination task and recording their reaction times (RTs). The RTs from correct responses were then transformed and analyzed using multidimensional scaling (MDS) in order to reveal the relative perceptual distances between these fricatives.

2. STIMULI

A continuum of thirteen synthesized fricative stimuli was generated by filtering Gaussian white noise. The filter shapes were based on observations and measurements taken from recordings of four 'anchor' fricatives: postalveolar [ʃ] and alveolar [s] from Canadian English, alveolo-palatal [ɕ] from Mandarin Chinese, and palatal [ç] from Russian and German. The palatal fricative [ç] is not a sibilant, but was included due to its shared similarities with this class of

speech sounds. Like sibilants, it has very little energy in the lower frequencies, a relatively sharp and well-defined peak, and a gradual decrease in amplitude in the high frequencies. It differs from sibilants primarily in peak amplitude.

In an effort to focus on the primary perceptual cues used to identify sibilants and palatal fricatives, the filter shapes were modeled as simplistically as possible, using only seven sets of frequency and amplitude values for each filter.
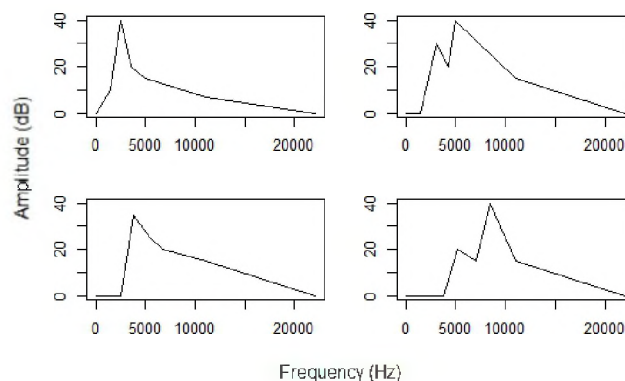

Figure 2. Filter shapes for the four anchor stimuli: [ʃ, ɕ, ç, s]. Note that while the sibilants had amplitude ranges of 40 dB, the palatal fricative [ç] had a range of only 35 dB. The frequency and amplitude values of the four anchor filters went through a two-stage process of interpolation in order to create a continuum of thirteen filters. The anchor stimuli [ʃ, ɕ, ç] and [s] correspond to positions 1, 5, 9 and 13 in the stimulus continuum.

A 600 ms sample of white noise was generated using Praat (Boersma & Weenink 2010), with the first and final 100 ms tapered down to 0 dB in order to approximate the spectral envelope of naturally-produced fricatives. This tapered white noise was then passed through each of the thirteen filter shapes using MATLAB's `filter` function (MathWorks 2008) in order to create the final stimulus continuum.

3. METHOD

The participants were 30 native speakers of Canadian English from the University of Alberta. They received partial course credit in exchange for their participation.

The participants were presented with an AX (same-different) discrimination task where they were asked to determine if pairs of synthesized fricative stimuli were instances of the same stimulus, or two different stimuli. They were heard 312 'same-different' stimulus pairs and an equal number of 'same-same' pairs, presented in random order, for a total of 624 trials per participant.

The stimuli were presented using E-Prime 2.0 (Psychology Software Tools 2002). Participants were given feedback after each trial indicating whether their previous response was correct or incorrect, and, if correct, their reaction time was also displayed. This was done to help participants respond as quickly and accurately as possible.

4. RESULTS & ANALYSIS

The overall error rate was 13.7%; the pair [ʃ-ɕ] had the highest error rate, at 7.5%.
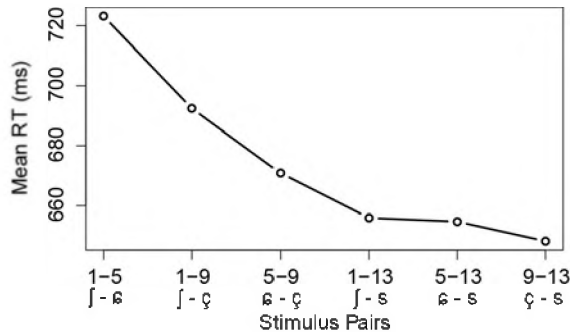

Figure 3. Mean RTs for 'anchor' stimulus pairs.

Pairs that involved a contrast with alveolar [s] had the lowest reaction times, suggesting this fricative was easily distinguished from the others.

The mean RTs for each stimulus pair were transformed into dissimilarity proximities by taking their reciprocal (1/RT). These proximities were then analyzed using the `cmdscale()` function included in the Base Package of the statistical analysis program R (R Core Team 2012). The data was modeled in three dimensions (stress value: 7.47). MDS analyses are most easily viewed in a two-dimensional format, and in this case the most noteworthy comparison is found between dimensions one and three, seen in Figure 4.
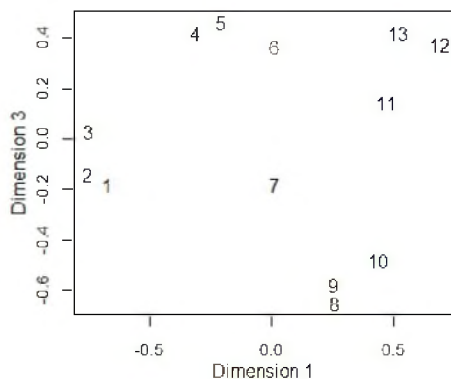

Figure 4. MDS model of dimensions 1 and 3 for all stimuli. Key for anchor stimuli: [ʃ]=1, [ɕ]=5, [ç]=9 and [s]=13.

We see here that the majority of the stimuli are clustered in four groups, and each group is formed around one of the four anchor stimuli. The relatively long distances between the clusters indicates that English listeners did not find it difficult to distinguish the four anchor stimuli, including the non-English fricatives [ɕ] and [ç].

5. DISCUSSION

The dimensions assigned by the MDS model are inherently meaningless, but they tend to reflect underlying structure in the data. In this case, the first dimension in Figure 4 seems to reflect a general increase in the spectral mean, much like the diagram seen in Figure 1. Dimension three serves primarily to differentiate the amplitude of the spectral means, with the non-sibilant palatal fricative [ç] separated from the sibilants [ʃ, ɕ] and [s].

This study demonstrated that it is possible to learn about the perceptual organization of sibilant and palatal fricatives using only listeners' RTs. The palatal fricative [ç] was seen to fall within the sibilant fricative continuum, and non-phonemic fricatives were reliably distinguished by English listeners. Further investigation will consider more fine-grained differences in spectral shape, as well as including vocalic environments in order to better emulate fricatives as they appear in natural speech.

REFERENCES

Boersma, Paul, and Silke Hamann. (2008). The Evolution of Auditory Dispersion in Bidirectional Constraint Grammars. *Phonology*, 25, 217–270.

Boersma, Paul, and David Weenink. (2010). Praat: doing phonetics by computer. http://www.praat.org/.

E-Prime 2.0. (2002). Pittsburgh, PA: Psychology Software Tools.

Fujisaki, H., and O. Kunisaki. (1978). Analysis, recognition, and perception of voiceless fricative consonants in Japanese. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 26, 21–27.

Howe, M. S., and R. S. McGowan. (2005). Aeroacoustics of [s]. *Proceedings: Mathematical, Physical and Engineering Sciences*, 461, 1005–1028.

MATLAB. (2008). Version 7.6.0.324 (R2008a). Natick, MA: MathWorks, Inc.

R Core Team. (2012). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Shadle, Christine H. (2012). The Acoustics and Aerodynamics of Fricatives. *The Oxford Handbook of Laboratory Phonology*, ed. by A. C. Cohn, C. Fourgeron, and M. Huffman. Oxford: Oxford University Press.

Toda, Martine, Shinji Maeda, and Kiyoshi Honda. 2010. Formant-cavity affiliation in sibilant fricatives. *Turbulent sounds - an interdisciplinary guide*, ed. by Susanne Fuchs, Marzena Zygis, and Martine Toda, 343–374. Walter de Gruyter.