THE EFFECT OF VOWEL CONTEXT ON ACOUSTIC CHARACTERISTICS OF [ç,x]

Christine H. Shadle, Sheila J.Mair and John N. Carter
Department of Electronics and Computer Science,
University of Southampton, Southampton SO17 1BJ, U.K.;
Neil Millner, DRA Malvern,
St. Andrew's Road, Malvern WR14 3PS, U.K.

ABSTRACT

In this paper we study the acoustic effects of vowel context on [ç,x] by spectral analysis of sustained and unsustained productions by two native speakers of German. Comparisons to non-German speakers of the same corpus allow inference of the acoustic mechanisms involved. [x] is more influenced by vowel context than [ç]. Evidence exists of changes in place, in the area of the constriction, and in the source localization or effectiveness, due to vowel context.

INTRODUCTION

In formulating models of fricative production, vowel context has important and sometimes unexpected effects, as shown by recent studies [1,2,3]. Acoustic spectra of [s] have been reported to display the greatest effect of vowel context in studies based on extensive analysis of two speakers (one a native of French, the other, of American English), and using spectral analysis primarily in the centre of the fricative [1,3]. Contrasting with these results is an analysis of aerodynamic data of fricatives, showing a greater effect of vowel context on area of the constriction as place of constriction moves posteriorly; it was suggested that fricative configurations independent of the tongue body are relatively immune to vowel context [2]. The work presented here therefore represents a more detailed look at the acoustic effects of vowel context on [ç,x].

METHOD

The speech corpuses consisted of the fricatives [s,ʃ,ç,x] in two environments: (1) preceded by the vowels, respectively, [a,a,i,a], sustained for 3 s, and repeated six times each, and (2) inserted into the nonsense words [pV_1FV_2] and repeated 10 times on a single breath, for V_1, V_2 chosen from [a,i,u]. Four subjects were recorded: the two subjects, CS and PB, reported on previously, for whom extensive articulatory and airflow data were available, a native German man, CD, for whom direct palatography and some airflow data were available, and a native German woman, EG. The inclusion of the German native speakers was prompted by several observations that CS and PB produced the phonemes not native to them more variably and with less place consistency.

The acoustic recordings were made under the 'High-Fidelity conditions' reported previously [3,4]. Time-averaged acoustic spectra were computed for the sustained fricatives using 8 non-overlapping 20-ms Hanning windows, positioned in the centre 160 ms of each token, and averaging the resultant Discrete Fourier Transforms (DFT's). Ensemble-averaged acoustic spectra were computed for the fricatives in vowel context by positioning one 20-ms Hanning window at the same position (i.e. beginning, middle or end of the steady-state portion of the fricative) in each of 8 tokens, and averaging the resultant DFT's. The technique is described in more detail in ref. [1]. The first and last tokens of an item were omitted.

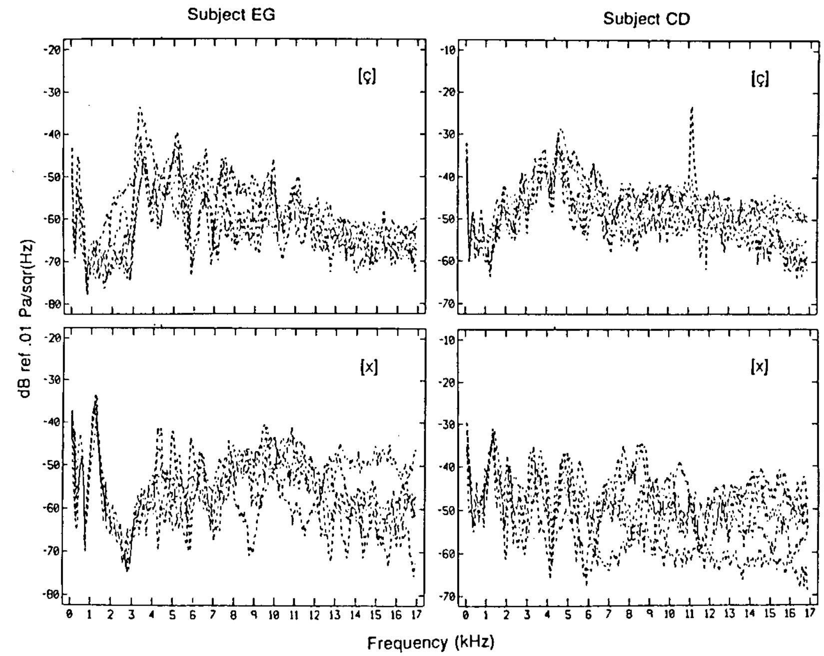

Figure 1: *Spectra of sustained fricatives produced by subjects EG and CD. Each graph shows six curves, one for each token.*

RESULTS

Figure 1 shows spectra of all six tokens of the sustained [ç, x] as produced by the native German speakers. The single peak at 11 kHz in CD's [ç] occurred in the first token produced and is indicative of a slightly whistly fricative – perhaps a response by the subject to the unnatural task of sustaining the fricative for 3 s. After comparison with the corresponding spectra of CS and PB (not shown), spectra of the sustained [ç,x] appeared to be slightly more variable token-to-token for the native Germans than for the other subjects, contrary to expectations. The overall spectral shape is fairly similar. In particular, [ç] is distinguished by low amplitude at low frequencies, extending up to 1 kHz (males) or 2-3 kHz (females), followed by a broad peak of high amplitude, made up of many smaller peaks, extending up to 6 or 7 kHz. [x] has fairly evenly spaced formant-like peaks, beginning at about 1 to 1.5 kHz, sometimes separated by deep troughs, as visible in EG at 3 kHz. Clearly there is more variation in the shape at frequencies above about 5 kHz. It is interesting to note, and not unexpected, that the same-sex subjects show greater similarity in spectral shape than the same-language subjects.

The spectral structure (but not amplitude) of sustained and unsustained productions of a given fricative was similar for the same vowel context, for each subject. Within the unsustained productions, vowel context affected the frequencies of spectral peaks for all subjects; for example, for the most part an [u] context lowered frequencies. Figure 2 demonstrates a more extreme effect of vowel context, where 'natural' and 'unnatural' contexts are contrasted. Even though subjects commented on the 'impossibility' of items such as [paça, pixi], the fricatives did retain their distinctive spectral shapes even in such cases. They did alter, however, as shown in the figure: the high-amplitude region begins at a higher frequency for [ç] in [piçi] than [paça]; the amplitude relationship of the first two peaks in [x] is significantly reversed in the two contexts, and the trough at 2.8 kHz is much deeper in [paxa].

Subject CD shows less of a difference between [a] and [i] contexts than does EG, but more of a difference for an [u] context. Figure 3 contrasts [paxa] and [paxu], showing beginning, middle and end of the fricative steady-state for each. While the two cases begin similarly, in [paxu] the peak at 2.2 kHz has dropped 17 dB from beginning to end spectrum, contrasted with a 4 dB drop in [paxa]. Peak and trough frequencies remain the same, except for the peak at 1 kHz. This effect appears consistently in all of CD's [x] spectra with [u] context, and appears, but to a lesser degree, in EG's spectra.

DISCUSSION

The lack of articulatory and much aerodynamic data for EG and CD, and the difficulty of gathering such data for [ç,x], make it difficult to explain

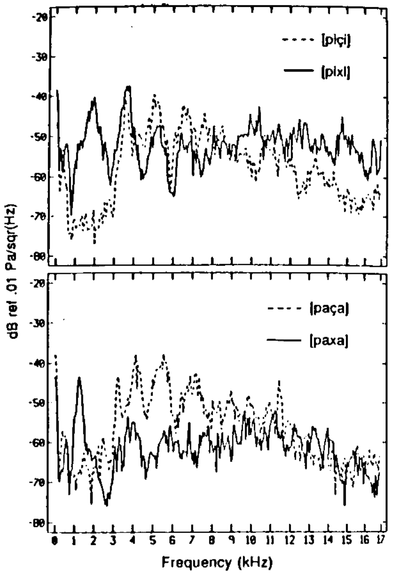

Figure 2: *Each graph contrasts ensemble-averaged spectra from the middle of the steady-state portion of the fricatives [ç,x] in the vowel contexts shown. Subject is EG.*

the various effects of vowel context observed above. However, from studies of CS and PB, for whom more articulatory data are available, some useful facts emerge. The cavity affiliations of the formants in [ç] were identified using a white noise source [4]; the lowest-frequency high-amplitude peak corresponds to the first front-cavity formant. In [s,ʃ] an [i] context shifts the place of constriction slightly forward relative to that for an [a] context. An [u] context for PB's [s] alters the source in a way that affects the fricative spectrum significantly [3].

The increased frequency of the peaks in EG's [piçi] compared to [paça] indicates that the place of constriction is more anterior in [piçi]. The same appears to be true for [pixi] compared to [paxa], where the high-amplitude peaks

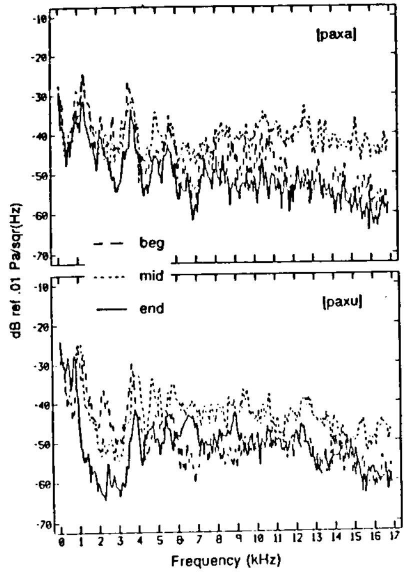

Figure 3: *Each graph contrasts ensemble-averaged spectra from the beginning, middle and end of the steady-state portion of the fricative [x] in the vowel contexts shown. Subject is CD.*

at 2 and 3.8 kHz appear to be front-cavity affiliated. However, the trough between these peaks deepens significantly for [paxa] and even more so for [puxu], which cannot be so easily explained by small shifts in place. The small peak visible in Fig. 3 at 2-2.5 kHz is a back-cavity formant, and is more prominent in [a] contexts, possibly indicating that the degree of coupling of back and front cavities is increased. The deepness of the trough in [u] contexts may be partly due to a reordering of poles and zeros due to the effect of rounding on pole frequencies, but it may also be due in part to changes in the degree of localization or effectiveness of the noise source. Since the constriction shapes and 'aims' the turbulent jet, such changes in spectral zeros may point to changes in the constriction shape. It is clear that for both subjects vowel context has a greater effect on [x] than on [ç].

CONCLUSIONS

The effect of vowel context on [ç,x] was investigated for two native speakers of German. Place of constriction moves anteriorly slightly in [i] contexts, increasing front-cavity formants, and rounding in [u] contexts decreases formant frequencies and bandwidths. However, some changes in constriction shape appear to occur as well, affecting the relative amplitude of back cavity formants and the significance of spectral zeros. These changes, and the fact that [x] exhibits more changes with vowel context than [ç], are consistent with Scully's explanation [2] of aerodynamic data on constriction area.

ACKNOWLEDGEMENTS

Work supported in part by a SERC studentship award to N. Millner, and by a collaborative EC SCIENCE award, CEC-SCI*0147C(EDB), and a European CEC-ESPRIT project Speech MAPS.

REFERENCES

[1] Shadle, C.H., Moulinier, A., Dobelke, C. & Scully, C. (1992), "Ensemble averaging applied to the analysis of fricative consonants", *Proc. of ICSLP-92*, vol.1, Banff, pp.53-56.
[2] Scully, C. (1992), "L'Importance des processus aerodynamiques dans la production de la parole", *Actes 19èmes Journées d'Etude sur la Parole*, Brussels 19-22 May, pp. 7-12.
[3] Shadle, C.H. & Scully, C. (1995), "An articulatory-acoustic-aerodynamic analysis of [s] in VCV sequences", *J. Phonetics*, vol. 23.
[4] Shadle, C.H., Badin, P. & Moulinier, A. (1991), "Towards the spectral characteristics of fricative consonants", *Proc. of the XIIth Int. Cong. of Phon. Sci.*, Aix-en-Provence, vol.3, pp.42-45.