Poznań Studies in Contemporary Linguistics 40, 2005, pp. 209-216

REPORT

TEMPORAL PHENOMENA IN POLISH HUSHING AFFRICATES

GEOFF SCHWARTZ
Adam Mickiewicz University, Poznań

ABSTRACT

This paper presents the results of an instrumental and perceptual study of two contrasting affricates in Polish, the palato-alveolar /ʧ/ and the alveolo-palatal /ʨ/. The production study revealed that in addition to the expected differences in the spectrum of the frication noise, the contrast is also maintained in the temporal parameters of frication duration (FD) and rise time (RT). On the basis of the production study, a series of perception tests were designed to test the hypothesis that the temporal parameters are important for the contrast; it was found that of the temporal parameters tested, overall FD, seems to be the primary parameter listeners use to hear the contrast.

1. Background information

Speakers of Polish have three voiceless coronal affricates, a dental /ts/, an alveolo-palatal /ʨ/, and a palato-alveolar /ʧ/ (the latter two are often referred to as 'hushers' while the palato-alveolar is sometimes analyzed as a retroflex /tʂ/; cf. Rochoń and Pompino-Marschall 1999). Some remarks on distribution are necessary with respect to the affricates in question. In Polish, the vowel [i] cannot occur after palato-alveolars and dentals and is replaced by [ɨ]. This vowel, as in the word *czy* (an interrogative conjunction) in Polish, is traditionally transcribed as [ɨ]. (One could argue for the use of the signal [ɪ] due to the large separation between the first and second formants of this vowel, observed in Wierzchowska (1971) and by this author in the laboratory.) Conversely, [ɨ] cannot occur after the alveolo-palatals. In this distribution it is likely that the high front vowels are allophones of a single phoneme. The affricates themselves are separate phonemes, as there are numerous minimal pairs (*gracz* /graʧ/ – *grać* /graʨ/ 'player' – 'to play').

2. The Polish affricates

Polish has three places of affricate articulation, dental (not involved in the contrast under study, but included here for descriptive purposes), alveolo-palatal, and palato-alveolar. There is some variation in the features used to distinguish among the three. Hall (1998) cites the dental as [+anterior], while the alveolo-palatal, which he considers to be inherently palatalized, is specified as [+high] and [–back], and the palatal-alveolar is [–high]. Wierzchowska (1971) classifies all Polish sounds based on three basic tongue positions: forward, flat (or neutral), and back. These categories are based not on articulation, but on the position of the mass of the tongue as seen in X-ray tracings (cf. Wierzchowska 1971). For the alveolo-palatal, the tongue is in the forward position, while for the other two affricates it is in neutral position. The alveolo-palatal is articulated laminally, while the other two are apical.

The phonetic distinction among places of articulation of affricates has been traditionally classified as similar to that of fricatives, where the relevant acoustic parameter in distinguishing them is the frequency spectrum of the frication noise (as well as formant transitions on neighboring vowels). In Wierzchowska's (1971) description of Polish phonetics, palato-alveolars have the highest noise energy between 2000-5000 Hz, alveolo-palatals between 3000-5000 Hz, and dentals above 4500 Hz. This pattern is to be expected when we look at the articulations. The palato-alveolars have the largest cavity in front of the constriction, including a sublingual cavity. The air in a larger cavity should resonate at a lower frequency. The dentals naturally have a minute cavity ahead of the constriction, thus will have a concentration of noise energy at much higher frequencies. The alveolo-palatals, articulated with the blade of the tongue and thus lacking a significant sublingual cavity, fall in between the two extremes.

Foreign students of Polish are often confounded by its fricatives and affricates, and in particular this distinction between the alveolo-palatals and palato-alveolars. Indeed these sounds both have a similar 'hushing' quality, and there is a large overlap in the noise spectra of the two affricates. Nevertheless this distinction persists in Polish. Thus it seems likely that there is additional acoustic information that Polish speakers attend to in maintaining this distinction. These additional cues can be found in the affricates' temporal characteristics.

Given the dynamic nature of affricates – segments made up of portions of two different classes of sounds – one might expect temporal information to play an important role in distinguishing them from stops and fricatives. Indeed, as mentioned earlier, homorganic fricatives and affricates differ in the duration of the rise time of the frication noise – fricatives have longer rise time than affricates.

In addition to the expected spectral differences, acoustic examination of these affricates showed temporal differences as well. The duration of frication noise (FD) and the rise time (RT) tended to be shorter in the palato-alveolars than in the dentals or the alveolo-palatals. The latter two places of articulation seemed to be relatively close in these two parameters. Thus the palato-alveolar appears more 'affricate-like' (often with a stronger stop burst than the other affricates), while the other two, with a weak burst and long RT, are more 'fricative-like'. Indeed, across languages the most common place of articulation for affricates is palato-alveolar.

Palatograms of these sounds (again found in Wierzchowska 1971) can shed some light as to the causes of these phenomena, indicating that a crucial difference between [ʧ] and [ʨ] lies in the surface area of the constriction, which for the palato-alveolar is significantly smaller. When there is a complete closure of the airstream, the pressure builds until the constriction is released. In the case of the alveolo-palatal, more of the tongue's mass forms the closure and there is more resistance as the air passes through the opening. Thus, the frication noise is slower to develop and lasts longer. By the same token we might expect the dental affricate to have similar RT and FD to the palato-alveolar. However, in the case of the dental, the aperture through which the air passes during the fricative portion is much smaller. In order to maintain the small aperture, the tension in the tongue must be increased, providing greater resistance to the airflow and increasing the rise time and frication noise.

In order to get a detailed picture of these affricates we must look at both spectral and temporal parameters. In the spectral domain we are interested in the frequency range of the frication noise, while FD and RT are the relevant temporal parameters. In maintaining this contrast, both spectral and temporal differences should be expected.

3. Production study

3.1. Subjects and data

Seven male, university-educated, native speakers of Polish (residents of Poznań) in their twenties, each produced eight repetitions of each affricate, four in word-final position after /e/ in monosyllabic words, four in word-initial position before /a/ in bisyllabic words. The words uttered in the carrier sentence *Powiedz słowo* ________ ('Say the word ________'). Additional dummy words were also added to the production stimuli, thus each speaker produced 40 tokens, sixteen of which (8 ʧ, 8 ʨ), presented in random order were subject to analysis.

3.2. Phonetic paremeters

Based on phonetic observations of the Polish affricates outlined in the previous section, the following phonetic parameters were measured to determine the status of the contrast between [ʨ] and [ʧ]:

(1) Noise Range Low (Hz)
(2) Noise Range High (Hz)
(3) Frication Duration (FD) in milliseconds
(4) Rise Time (RT) in milliseconds

The first two parameters determine the frequency range of the frication noise on the affricates. The range is calculated by measuring the frequency range at which the frication noise is within 6 dB of the noise level at its highest. (This number was arrived at based on numerous observations of Polish fricative spectra, as well as by a heuristic approach in creating synthetic stimuli for perception tests.) Parameters 3 and 4 are temporal measurements (in milliseconds) discussed in the previous section. To provide additional insight into the temporal measurements, the ratio of rise time to frication duration, and as the center of the frication noise range were also calculated.

3.3. Analysis

A series of Analysis of Variance (ANOVA) tests were carried out on each parameter to determine whether in fact the distinction was maintained. In the case of the spectral parameters the tests were performed all measurements, while for the temporal parameters, the data were sorted into initial and final environments to normalize for phonetic context effects. For the ratio of Rise Time to Frication Duration, the ANOVA was carried out on all measurements.

4. Results

The results are summarized in Table 1. In the spectral parameters, a distinction was maintained for the noise range low ($F[1, 222] = 143.6$; $p < 0.001$) and noise center ($F[1, 222] = 8.79$; $p = 0.003$), but not for the noise range high ($F[1, 222] = 0.45$; $p = 0.5$) of the frication spectrum. This can be seen in Table 1, as well as results for the temporal parameters, where a clear distinction was maintained in both the word initial ($F[1, 110] = 97.4$; $p < 0.001$ for frication duration; $F[1, 110] = 218.3$; $p < 0.001$ for rise time) and word-final ($F[1, 110] = 159.9$; $p < 0.001$ for FD; $F[1, 110] = 370.4$; $p < 0.001$ for RT) contexts, as well as overall ($F[1, 110] = 510.2$; $p < 0.001$ for RT/FD).

5. Perception Tests

Based on data obtained from the production study, my own additional analysis of Polish affricates, as well as Wierzchowska (1971), I synthesized (using SynthWorks by SciCon RD) an affricate with a frication noise range (3000-4500 Hz) that falls between characteristic values (for males) for /ʧ/ and /tɕ/. Several Polish speakers verified the perceptual ambiguity of the affricate. The synthesized affricate, with a constant noise spectrum and RT (0.30), was placed in both CV and VC contexts (the vowel was /e/ in both cases). For each context, affricates with three different frication durations were produced (50, 75, and 100 ms for CV; and 70, 95, and 120 ms for VC – word final affricates were consistently longer). Twelve Polish listeners were presented with 10 repetitions of each duration for each context in a randomized order, and given a forced choice task, asked to mark whether they heard /ʧ/ or /tɕ/ (which are systematically represented in Polish orthography). Listeners were given a brief test using natural speech to make sure they perceived the distinction under study. Those who passed (12 out of 16 interviewed) were then given the synthesized stimuli. The results of this test, with RT/FD constant are presented in Table 2.

Table 1. Overall mean values for spectral (above) and temporal parameters. FD refers to duration of frication noise, RT to rise time of the affricate

	ʧ	tɕ	
Noise Low (Hz)	2711.8	3217.9	$F(1, 222)=143.6$; $p<0.01$
Noise High (Hz)	4726.6	4760.2	$F(1, 222)=0.45$; $p=0.5$
Noise Center (Hz)	3719.2	3989.0	$F(1, 222)=8.79$; $p=0.003$
Initial FD (ms)	58.2	78.5	$F(1,110)=97.4$; $p<0.001$
Initial RT (ms)	20.9	42.0	$F(1,110)=218.3$; $p<0.001$
Final FD (ms)	85.5	112.1	$F(1,110)=159.9$; $p<0.001$
Final RT (ms)	32.0	59.6	$F(1,110)=370.4$; $p<0.001$
RT/FD Overall	0.36	0.53	$F(1,110)=510.2$; $p<0.001$

Table 2. Results of Perception Test 1, Noise Spectrum and RT (30 ms) held constant

Noise duration and context	Perceived as /ʧ/; total number and %	Perceived as /tɕ/ total number and %
50 ms (CV)	94 (78%)	26 (22%)
75 ms (CV)	56 (47%)	64 (53%)
100 ms (CV)	19 (16%)	101 (84%)
(VC) 70 ms	65 (54%)	55 (46%)
(VC) 95 ms	35 (29%)	85 (71%)
(VC) 120 ms	9 (7%)	111 (93%)

An additional test was then prepared in which the RT/FD ratio was constant (0.45), and both Rise Time and Frication duration were varied, increasing in 3 steps (34 ms RT, 75 ms FD; 45 ms RT, 100 ms FD; 56 ms RT, 125 ms FD). The results are shown in Table 3.

A third perception test was prepared using synthesized affricates with constant Frication Duration (100ms) and noise spectrum (3000-4500 Hz), with three different Ratios of Rise Time to Duration (0.2, 0.4, 0.6). All of the listeners reported that they

Table 3. Results of Perception Test 2. Noise spectrum and RT/FD (0.44) held constant

Rise Time/ Frication Duration (Context)	Perceived as / tʃ /; total number and %	Perceived as /tɕ/; total number and %
22 ms / 50 ms (CV)	102 (85%)	18 (15%)
33ms / 75 ms (CV)	51 (43%)	69 (57%)
44 ms / 100 ms (CV)	16 (13%)	104 (87%)
(VC) 32 ms / 70 ms	71 (59%)	49 (41%)
(VC) 43 ms / 95 ms	47 (39%)	73 (61%)
(VC) 54 ms /120 ms	4 (3%)	116 (97%)

could not reliably hear any difference in these stimuli; they either guessed blindly or simply picked the same choice for each token.

Biases for Polish listeners might tend to lean toward to the alveolo-palatal, which is extremely common as a historical reflex of a /t/ before a front vowel, and appears in nearly all infinitive verb endings. While /tʃ/, the historical reflex of a velar, occurs somewhat less frequently, Polish does not suffer from a shortage of these segments.

6. Discussion

Based on the production results, the distinction between the palato-alveolar and the alveolo-palatal affricates appears to be maintained in both the spectral and temporal parameters. However, when interpreting spectral measurements of frication noise, one should keep in mind that pitch sensitivity in the human auditory system becomes less keen as frequencies rise above 4000 Hz (see Durrant and Lovrinic 1995: 265-269). As a result, the low end of the frequency range is likely to be far more salient for listeners in perceiving the place of articulation. In this parameter the subjects show an average difference of 511 Hz (see Table 1). While such a difference is certainly perceptible, there is an average overlap of over 1500 Hz in the noise ranges of the two affricates, serving to attenuate the salience of the spectral distinction.

While the production data indicate that the contrast is apparently maintained in both the spectral and temporal parameters, facts about the auditory system, combined with the clear distinctions in the temporal parameters, might tend to promote the status of the temporal cues to this Polish contrast. Of the two temporal parameters, overall frication duration (FD) seems to have the most reliable effects in the perception tests. Adding RT to FD as a perception test variable did not seem to have much impact on the results (cf. Tables 2 vs. 3). When forced to rely on RT alone, listeners could not consistently hear any distinction. While RT is an important cue for the stop-affricate distinction (Johnson 1997), the most perceptually salient acoustic feature in the place distinction between the Polish palato-alveolar and alveolo-palatal affricates seems to be FD.

This result runs parallel to findings showing differences in duration based on fricative place of articulation, and might be explained by the same physiological factors outlined in Section 2 of this paper. Because of the smaller surface area of its constriction, /tʃ/ has a shorter period of frication than /tɕ/. It seems this fact has perceptual significance for the Polish phonemic inventory, and needs to be considered in phonological analyses.

When Polish /ts/ is added to the mix, the three affricates present some interesting possibilities as far as interaction of perceptual cues is concerned. When the dental [ts] is contrasted with the alveolo-palatal [tɕ], the temporal distinctions are minimal and listeners must rely primarily on spectral cues. However, when [tɕ] is contrasted with the palato-alveolar [tʃ], the temporal cues are promoted. At same time, rapid speech can serve to attenuate the salience of temporal phenomena. The main point is that the weight of a single acoustic cue can depend on all sorts of factors, including which contrast listeners are trying to make ([tɕ] v. [tʃ] or [tɕ] v. [ts]), the phonetic context, and speaking rate.

Nearey (1997), in outlining his "double weak" model of speech perception, allows for this type of variation in the relative weight of perceptual cues. In his account, the perception of a phonetic contrast is based on the integration of numerous types of information from purely phonetic phenomena to biases based on phonotactics and word frequency. Each type of cue can have something of an independent status, and its importance in perception can vary. Incorporating this type of perceptual approach can have significant implications for the study of sociolinguistics and diachronic development, providing a link between perceptual (Ohala 1981) and sociolinguistic (Labov, 1997) approaches to language change, and a deeper understanding of many diachronic problems.

In fact, Polish /tʃ/ and /tɕ/ might provide another case study of language change in progress. It is likely that many Polish speakers show a near-merger (Labov, 1997) of the two affricates. Two facts from this study raise this possibility: (1) four potential perception test subjects did only marginally better than chance on the preliminary test for the distinction, and (2) listeners could not reliably hear the differences in the stimuli in the third perception test. In addition, while teaching Polish to heritage speakers in the United States, I have consistently observed orthographic mistakes indicating difficulties perceiving the contrast under study. This is a study I plan to undertake with Polish speakers both in the U.S. and in Poland.

REFERENCES

Durrant, J.D. and J.H. Lovrinic. 1995. *Bases of hearing science.* Baltimore: Williams and Wilkens.

Hall, T.A. 1998. *The phonology of coronals.* Amsterdam: Benjamins.

Johnson, K. 1997. *Acoustic and auditory phonetics.* Cambridge, MA: Blackwell.

Labov, W. 1997. *Principles of linguistic change: Internal factors.* Oxford: Blackwell.

Nearey, T. 1997. "Speech perception as pattern recognition." *Journal of the Acoustical Society of America* 6. 3241-3254.

Ohala, J.J. 1981. "The listener as a source of sound change." In Masek, C.S., R.A. Hendrick, and M.F. Miller (eds.), *Papers from the parasession on language and behavior.* Chicago: Chicago Linguistic Society. 178-203.

Rochoń, M. and B. Pompino-Marschall. 1999. "The articulation of secondarily palatalized coronals in Polish." *Proceedings of the 14th International Congress of Phonetic Sciences.* San Francisco. 1897-1900.

Wierzchowska, B. 1971. *Wymowa polska.* Warszawa: Państwowe Zakłady Wydawnictw Szkolnych.