Małgorzata Ćavar, GK Universität Potsdam-Humboldt-Universität zu Berlin
and
Silke Hamann, OTS Utrecht & ZAS Berlin

Polish velar and coronal palatalization – its perceptual background

Introduction

Several languages, e.g. Korean (Kim 2001) and Quebec French (Charbonneau & Jacque 1972), show an alternation between coronal stops with palatoalveolar affricates in the context of a front vowel, cf. (1a). In other languages, velar stops palatalize to palatoalveolar affricates in the same context, for example in dialects of Italian (Calabrese 1993) and Kinyarwanda (Ladefoged & Maddieson 1996), cf. (1b).

(1) a. t → ʧ/ _ i, e b. k → ʧ / _ i, e

What happens when a single language has phonological rules that palatalize both coronals and velars with a change in the major place of articulation? In Polish there is no neutralization of the place distinction in palatalized segments: we have two different outputs of palatalization of coronals and velars within the post-alveolar region, cf. the two rules in (2) (the data exemplifying (2a-b) follows in §1).[1]

(2) a) t → tɕ / _ i, e b) k → tɕ / _ i, e

The alternation in (2a) is traditionally referred to as Coronal Palatalization, in (2b) – as First Velar Palatalization.

The present study is concerned with the question how the correlation between the inputs and outputs of palatalization in Polish, i.e. between /t/ and [tɕ] versus /k/ and [ʧ], as in (2), is motivated.

Much recent research in phonology, e.g. Boersma (1998), Flemming (1995), Padgett (2001), and Steriade (1997, 2001) focuses on functional explanations for phonological processes. In addition to the drive to save articulatory effort, these functional phonological approaches include the influence of perception on language: speaker's communication is guided by the need to produce the most distinctive speech signal for the sake of best perception by the listener.

Previous research on Polish recognized the articulatory mechanism behind palatalization (e.g. Rubach, 1984; Szpyra, 1995), however, they could not explain the basis of the relation between the sets of inputs and outputs in the process. In the present article we posed the hypothesis that, under the assumption that both palatalization processes in Polish in (2) are advantageous from the articulatory perspective, the outputs are chosen for their perceptual similarity to the respective inputs. In order to show this, we compared the acoustics of the relevant sounds in the context of front vowel [i][2], looking for similarities between input and output segments. Furthermore, we tested Polish speakers' intuitions about the relative perceptual similarity of the plosives [t] and [k] with the affricates [tɕ] and [ʧ]. The test was repeated with native speakers of German in order to assess the influence of native-speakers' lexical bias. The results showed that there is a functional, perceptual explanation for the correlation of the input /t/ with the output [tɕ]. For the pair /k/ - [ʧ], however, we could not prove systematic perceptual similarity.

The article is organized as follows. §1 gives the data on palatalization alternations in Polish. Furthermore, it shows that articulatory assimilation is only a secondary criterion for

[1] The output of the velar palatalization is described in this article as palatoalveolar and referred to with the symbol [ʧ], though one should bear in mind that Polish [ʧ] is clearly different from the prototypical [ʧ] like e.g. in English. On its possible retroflex status, see Hamann (2002). In Ladefoged and Maddieson the Polish sounds is referred to as [tʂ̩].

[2] The question of perceptual similarity of the relevant sounds in the context of back vowels or before a pause requires further tests.

the alternating pairs. §2 illustrates the perceptual approach. The experiments are described in §4 and the expected results are summarized in §5. The evaluation of results is given in §§ 6 and 7. The last section concludes and proposes an alternative, diachronic explanation for the synchronic data.

1. Palatalization of velar and coronal stops in Polish

The process of First Velar Palatalization (Rubach 1984) in Polish yields an alternation between the velar [k, g][3] and postalveolar affricates [ʧ, ʤ] in the adjacency of an underlying front vowel, as in (2b), cf. examples in (3):

(3) kro[k] 'step' kro[ʧ]+ek 'step,' diminutive
mózg[g] 'brain ' móż[ʤ]+ek 'brain,' diminutive

Palatalization of the coronal plosives [t, d] in the same context, usually referred to as Coronal Palatalization (Rubach 1984)[4], yields the alveolopalatal affricates [tɕ, dʑ] as in (2a), cf. the examples in (4).

(4) bra[t] 'brother' bra[tɕ]+e 'brother,' loc. & voc.
skła[d] 'composition' skła[dʑ]+e 'composition,' loc. & voc.

Both processes occur across morpheme boundaries only.[5]

In addition to these morphologically conditioned palatalizations, Polish has a Surface Palatalization Rule (Rubach 1984), where the addition of a high front vowel gesture to a consonant results in its secondary palatalization, see (5).

(5) [t, d] + i → [tʲ, dʲ] e.g. tʲik 'tic', dʲiwa 'diva'
[k, g] + i → [kʲ, gʲ] e.g. kʲino 'cinema', gʲitara 'guitar'

The primary place of articulation is slightly changed in both velar and coronal secondary palatalization in Polish, as the tongue modulates its shape in the direction of the vowel gesture. Whereas the purely assimilatory nature of such processes as in (5) cannot be questioned, many authors, e.g. Sagey (1985), Hume (1991), Lahiri & Evers (1992), Rubach (1993), Szpyra (1995) or Ćavar (1997) adopted purely articulatory explanations of palatalization processes in general. These accounts were proposed also for the change of the major articulation place, not only for secondary palatalization. Admittedly, given the "right" set of features, one might "account" for palatalization processes without reference to perception. No such account can however explain why [t] alternates with [tɕ] and not with [ʧ], and, why [k] selects as its alternant [ʧ], and not [tɕ].

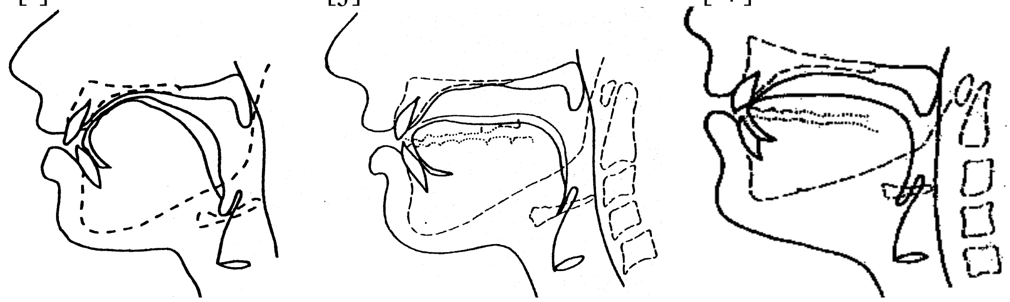
It is argued here that, in contrast to secondary palatalization, Polish Coronal and First Velar palatalization as in (4) and (5) cannot be explained by articulatory assimilation alone, as the alternation of /t/ and [tɕ] versus /k/ and [ʧ] (and the voiced counterparts) is articulatorily unexpected. Comparing the place of articulation of [t] (see 6a) with that of the friction part in both [ʧ] (6b) and [tɕ] (6c) in x-ray studies (Keating 1991: 36 for [tɕ] and Wierzchowska 1980: 58, 64 for [t] and [ʧ], respectively) shows that [ʧ] and [t] are articulated at the alveolar ridge, whereas [tɕ] is articulated at the alveolo-palatal region.

[3] All velar sounds of Polish, i.e. [k, g, x], undergo First Velar Palatalization. The present study, however, discusses voiceless plosives only.

[4] As in the case of velars, the alternation is not restricted to plosives.

[5] These processes apply exceptionlessly before [i]; before [e] there are a number of exceptions where palatalization does not take place, which led Rubach (1984) to postulate underlying back vowels in the non-palatalizing suffixes. In the present article, we follow Rubach's assumption. For a different view see e.g. Szpyra (1995).

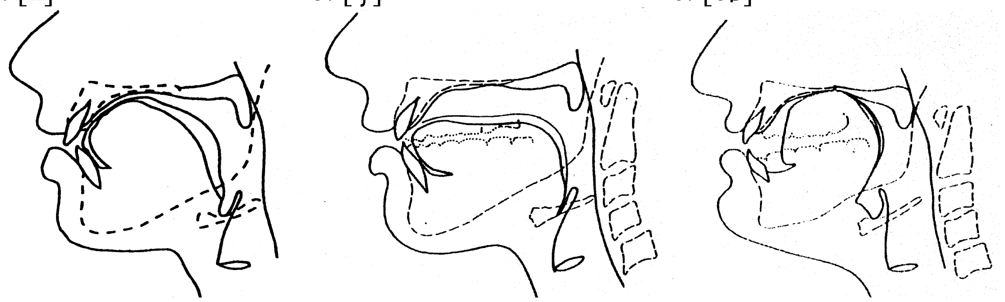
(6) a. [t] b. [ʧ] c. [tɕ]



Consequently, one would expect the output of palatalized /t/ to be [ʧ], not the actually occurring [tɕ] from an articulatory point of view, if one assumed only minimal changes to occur. Cross-linguistically, evidence for the unmarkedness of correlating /t/ rather with the prototypical [ʧ] is given by the number of languages which employ it, e.g. English or Korean and Quebec French mentioned above. On the other hand, [tɕ] is relatively rare cross-linguistically (cf. Ladefoged and Maddieson, 1996).

An articulatory comparison of [k] (in front vowel context, Wierzchowska 1980: 82) with [ʧ] and [tɕ], see (7a-c) respectively, shows that [k] is closer in place of articulation and tongue shape to [tɕ] than to the actual output [ʧ].

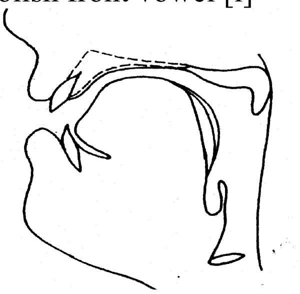
(7) a. [k] b. [ʧ] c. [tɕ]



The languages which have [tɕ] in their inventory in cases discussed in literature correlate it with an underlying /k/ (and not /t/) in front vowel context. In Mandarin (Cheng 1973), for example, a palatal fricative phoneme diachronically developed from a velar plosive before front vowels. Synchronically, Mandarin still shows variation between velar-schwa and palatal-i sequences, e.g. [kəŋ] ~ [ʧiŋ] 'neck' (Cheng 1973: 15). Summing up, the two outputs of Polish palatalization are not articulatory closest to their inputs, and unexpected from the point of view of markedness theory.

Up to now we did not take the articulatory information of the trigger, the front vowel [i], into consideration. Front vowels are articulated with the tongue body substantially raised and positioned at the front part of the hard palate, cf. (8) (Wierzchowska 1980: 87). This gesture is very similar to the tongue position of prepalatals such as [tɕ]. The articulatory configuration of [ʧ] is, on the other hand, less similar to the front high vowel, compare (7b) and (8). In this light, assimilation to [i] should result for both velar and dental plosives in the prepalatal affricate. Even if one does not assume a near total assimilation to the trigger, the change from a raised tongue body of a velar [k] towards a flat tongue body of [ʧ] in a raised tongue body context cannot be articulatory accounted for from the phonetic point of view.

(8) Polish front vowel [i]



The account of palatalization in terms of articulatory assimilation encounters yet another problem, namely that of affrication. The change of manner from plosives to affricates, often referred to as assibilation in the literature (e.g. Kim 2001), cannot be explained as an articulatory assimilation to the vowel, as front vowels do not show a constriction close enough to cause affrication. The problem is not solved by assuming that affrication is caused by spreading of phonological feature [+continuant] from the vowel, because then it still remains unclear why only front vowels contain a [+continuant] specification which is able to spread.

Clements (1999) and Kim (2001) show that there are perceptual reasons for affrication in palatalization. The release of a stop into a high front vowel causes a period of stronger air turbulence than in any other vowel contexts. The listener often misperceives this turbulence as affrication noise. This explanation is attested by the fact that palatalization with the change of the major place of articulation almost always co-occurs with affrication. In his typological study on palatalization, Bhat (1978) names only one language where the change of the major articulation place is not accompanied by affrication, namely Acoma, where dental stops are realized as palatal stops before front vowels.

The argumentation above shows that a functional account for the alternations in Polish First Velar and Coronal Palatalization based on articulation alone is insufficient.

2. Perceptual account

An additional factor in the palatalization correlations and a possible explanation is the perceptual similarity between input and output forms. The listener and learner can misparse the cues and hear a coronal with a following vowel instead of a velar (or labial) with a following high front vowel (see Ohala 1992, 1993 for similar reasoning as a source for sound change). Evidence for this misperception is given by Winitz et al.'s (1972) perception test, where listeners had to categorize bursts with different vowel contexts. In this test, the listeners misperceived [pi] as a [t] (with any following vowel) in a large number of times, but a reverse misperception of [ti] as [p] did not take place.

A similar effect has been obtained by Guion (1998), who tested velar stops and English palatoalveolar affricates [ʧ] in a perception experiment. Her results show that in the context of front vowels velar stops were more often misperceived as palatoalveolar affricates than expected. Chang, Plauché and Ohala (2001), on the other hand, claim that a perceptual similarity holds between velars in front vowel context and alveolar plosives.[6] But the present study is interested in palatalization with an output in the postalveolar region only; hence

[6] Chang et al. studied explicitly the relation between /tɕ/ and non-aspirated /k/ and /t/ in American English. Their results show that there is no /ki/ > /ʧ/ confusion asymmetry in laboratory conditions. Chang et al. subscribe the results of Guion (1998) to the fact that she did offer only a forced choice between /k/ and /ʧ/, without offering /t/. The fact that /k/ > /t/ is not a common sound change, whereas /k/ > /ʧ/ is, is explained by them by /k/'s often being aspirated, where this aspiration can be interpreted as friction portion of an affricate. In their experiment, /k/ was unaspirated and hence did not yield these results.

perceptual closeness to alveolar affricate is irrelevant. For the same reason, the alternation /k/- [ts], which also occurs in Polish but is purely lexical, was neglected.

In order to find a functional explanation for the aforementioned correlation in Polish palatalization of plosives, we posed the hypothesis that each affricate output is synchronically perceptually closer to its input than the other possible affricate, i.e. that [tɕ] is perceptually closer to [t] than [ʧ] is, and [ʧ] is closer to [k] than [tɕ] is. In other words, alternating pairs of segments need to be acoustically and perceptually similar. This assumption yields the following hierarchy of perceptual closeness in the context of front vowel [i].[7]

(9) t – tɕ – ʧ- k

This hypothesis was tested in a perceptual experiment comparing the four sounds in the context of front vowel [i]. The results of the experiments did not confirm the original hypothesis, and the emerging perceptual scale is represented in (10):

(10) t - tɕ - k
      |
      ʧ

3. Experiment

Three young, female Polish native speakers were recorded reading sentences with the segments /t, tɕ, ʧ, k/ with following /i/ and /a/ in word-onset and stressed position in normal reading style. The material was recorded onto a DAT tape in a soundproof booth and sampled at 32 kHz. Words used were:

(11) Tina / tata [t]
kino / kasza [k]
cicho/ ciasto [tɕ]
Czingiz/ czaszka [ʧ]

The Carrier sentence was *Czy mogłabyś powtórzyć X*, ‘could you repeat X’.

3.1. Acoustic analysis

In the first part of the experiment, the recordings were analyzed acoustically with the help of Praat version 3.9.36. Measurements of formant frequencies of F2 and F3 at the time of the burst (and at the end of the transition phase/at the stable phase of the vowel) were carried out. The length of the frication noise was measured (from burst till beginning of a vowel) along with the length of the burst. Furthermore, mean spectra of the burst noise in the case of plosives and of the friction in the case of the affricates (all filtered with a bandwith of 800 Hz) were analyzed and compared. The comparison of mean spectra was made only with spectra from single speakers, as there is much interspeaker variety, and a cross-speaker comparison could distort the results (Ladefoged & Maddieson, 1996: 174).

3.2. *Perception test*

In the second part of the experiment, the samples were cut to yield the relevant segments a) without any portion of the following vowel, b) with a whole syllable, i.e. with the following vowel [i]. [8] Out of these samples, an ABX test was constructed.[9]

[7] The assumed similarity between [tɕ] and [ʧ] is not topic of this study and will therefore be neglected in the following.

[8] For reasons why to restrict the test to [i]- vowel context, recall discussion on likelihood of confusion in this context in § 2.

[9] In an ABX test, a subject listens to triads of tokens (A-B-X), and is asked to say whether the third token in each triad (X) is more like the first token (A), or more like second token (B).

The listener had to compare a plosive with two affricates.[10] The combination of comparing one plosive to another plosive and an affricate, i.e. a mixture of modes for A and B, was excluded, as preliminary tests showed that listeners tend to associate the item in question with the one offered that is articulated in the same modus, i.e. plosive with plosive because of the burst noise and affricate with affricate because of the friction noise. The same conclusion has been reached by Łobacz (1985). These assumptions yield two possibilities, cf. (12).

(12)	A	B	X
a.	ʧ	tɕ	k
b.	ʧ	tɕ	t

To exclude any bias to the segment that was perceived last (B), the order of A and B was also reversed. Furthermore, the segments were presented once without vowel and once with vowel, including the following possibilities: all three as short signals; A and B as short, X as long signal; A and B as long and X as short; and all three long. Altogether there were 2 (2 different X) x 2 (reverse order) x 4 (length conditions) = 16 tokens. These tokens were presented in randomized order and in five repetitions.

The test was presented to 9 Polish native speakers and to 7 German native speakers.

4. Expected results

The shape of the formant transitions in velars in the context of front vowels resembles that of coronals (Stevens, 1998; Pickett, 1999), the shape of the most important for the perception first two formants is the same: F1 is located very low in the spectrum, and F2 is relatively high pointing downwards to the consonant. Therefore we did not expect a huge difference in the formant transitions of /ti/, /ki/ and /tɕ/ /ʧ/ in our acoustic measurements of these sounds. We expected however to find some similarity, e.g. burst length or spectral shape, in the acoustic signals of /ti/ and /tɕ/ exclusively, and also a similarity between /ki/ and /ʧ/ - to the exclusion of the remaining two sounds.

In the perception test, listeners were expected to judge [k] as perceptually closer to [ʧ] than to [tɕ]. [t], on the other hand, should be judged closer to [tɕ] than to [ʧ] in order to support the hierarchy of perceptual closeness posed in (9). The length of the signal is assumed to play an important role in this association. We suspected that shorter A and B would yield more confusion and disturb the categorization of A and B into well-known A and B categories (which would make it difficult for the listeners to relate them to the third category, X). On the other hand, longer X was expected to help the categorization of signal A or B into the category X. Furthermore we expected a bias towards judging X more often as B, because of short-term memory effects.

5. Evaluation of the acoustical analysis

In the following, the results of the formant values and length, the burst length and the spectra of all four sounds are described. These are evaluated in respect to the perceptual similarity hierarchy posed in §2, to see whether there is any supporting evidence for the hierarchy in the acoustics of the signals.

5.1. Formant values and length

In spite of the great variance within a speaker as well as across speakers, the formants show an overall homogeneous behaviour. The measured frequencies of second and third formants

[10] The reverse setting, where the listener has to compare an affricate to two plosives, was not included in this test.

for consonants followed by a front vowel are on average higher than those of consonants before back vowels, as expected from phonetic literature (e.g. Johnson 1997).

The table in (13) shows the average values of the second formant for the four consonants investigated followed by [i] and by [a]. The first column for every vowel gives the onset value of the transition, the second the value at the beginning of the steady state. As predictable, the vowel steady state lies around 2670 Hz for the [i] and around 1560 Hz for the [a].

(13) Average values of F2

	__i		__a	
t__	2378	2616	2000	1566
k__	2693	2683	1763	1515
tɕ__	2700	2650	2316	1600
ʧ __	2561	2711	2060	1600

Looking at the second formants in general, one can observe that the values for all four consonants lie very close to each other before the front vowel. For [t], the values are lowest across the contexts, with a rising transition in the [i] context and a falling transition for the [a] context. This is closely followed by the values for [ʧ], which shows the same transition directions. The F2 values for [tɕ] are somewhat higher than for [tʃ], 140 Hz for the high vowel and 150 Hz for the low. The transitions of [tɕ] are falling in both vowel contexts. [k] has a starting value for the F2 before [i] which is nearly identical with that of [tɕ]. In [a] context, however, the value is far lower than for any of the other consonants. Like [tɕ], [k] has slightly falling transitions in both contexts. With respect to the second formant, there is thus a closer similarity between [k] and [tɕ] in [i] context and between [t] and [ʧ] in both contexts.

In (14), the average values of the third formant for all four consonants are given.

(14) Average values of F3

	__i		__a	
t__	3166	3450	3061	2766
k__	3433	3583	2850	2416
tɕ__	3566	3600	2900	2600
ʧ __	3433	3466	2883	2581

The third formant values are very similar for all four consonants in the context of [a]. They start around 2900 Hz and are all falling, the transition for the [k] being a bit steeper than the others. Before [i], there are some differences observable. [ʧ] and [t] have nearly identical transition offsets, but whereas the former has a nearly steady transition, the latter has a rising one. The values for [k] and [tɕ] are very close together, too, and both sounds have rising transitions. Thus, the similarity between [k] and [tɕ] and between [t] and [ʧ] found for the second formant can be attested for the third formant. This is not in account with our proposed hierarchy in (9), here repeated in (15a) for convenience.

(15) a. t – tɕ – ʧ – k b. t – ʧ – tɕ – k

Instead, we get an acoustic hierarchy as in (15b). But it has to be pointed out again that for all four consonants the F2 values are very close in [i] context and the F3 in [a] context.

Interestingly, the front high vowel environment allows a greater variability of the values of the F2 transitions between speakers than the low vowel context, see table (16).

(16) Range of F2 values at the beginning of the transition

	__i	__a
t	2200-2670	1800-2100
k	2400-2960	1700-1800
tɕ	2600-2800	2100-2400
ʧ	2300-2800	2000-2200

Whereas in the environment of the high front vowel the extreme values of F2 vary as much as 500-600 Hz, in the context of the low vowel the range is between 200-300 Hz. F2 seems thus much more stable before low vowel than before high vowel.

Measurements of the transition length show that they are similar across the four consonants, with [t] posing the only exception, as its transitions are regularly shorter than that of the other consonants.

5.2. Release noise

In order to compare the release noise of the two plosives and affricates, we measured the beginning and ending of burst noise for the plosives and the beginning and ending of friction noise for the fricatives. The results show that the burst is consistently longer in high front vowel environment than in low vowel environment, cf. (17).

(17) Average noise duration

	__i		__a	
	average	range	average	range
t	0,053	0,041-0,071	0,014	0,011-0,020
k	0,088	0,064-0,098	0,054	0,040-0,065
tɕ	0,113	0,088-0,146	0,100	0,076-0,145
ʧ	0,098	0,067-0,138	0,040	0,031-0,051

Comparing the burst length across the four different places of articulations shows that for [t] it is far shorter than for any of the other consonants. The noise length is on average longest in alveolopalatals. The values for [ʧ] and [k] are very close, in high vowel context [ʧ] is shorter and in low vowel context longer than [k]. These findings are summarized in (18).

(18) Noise duration scales
  a. high vowels: t – ʧ – k – tɕ
  b. low vowels: t – k – ʧ – tɕ

If compared to the hierarchy proposed in § 2, cf. (15a), the noise duration scales in (18) support it in as far as they show a closeness of [ʧ] and [k]. But the scales give no evidence for the proposed similarity between [t] and [tɕ].

5.3. Spectra

The mean spectra of the fricatives and burst-aspirations were filtered with a bandwidth of 800 Hz. The spectra yielded the following characteristics for the sounds under observation. For [t] with a following [i], the overall shape of the spectrum is rather flat, with two main frequency areas, the first with two peaks between 2500 and 3200 Hz and the second with one prominent peak around 6000 Hz. For the [a] context, the spectrum is more sloping down from a first peak around 500 Hz, with a second prominent peak again around 6000 Hz, see (19). The findings of Patryn (1987: 60) that the alveolar plosive shows generally higher energy between 3200 and 3800 Hz could not be confirmed.

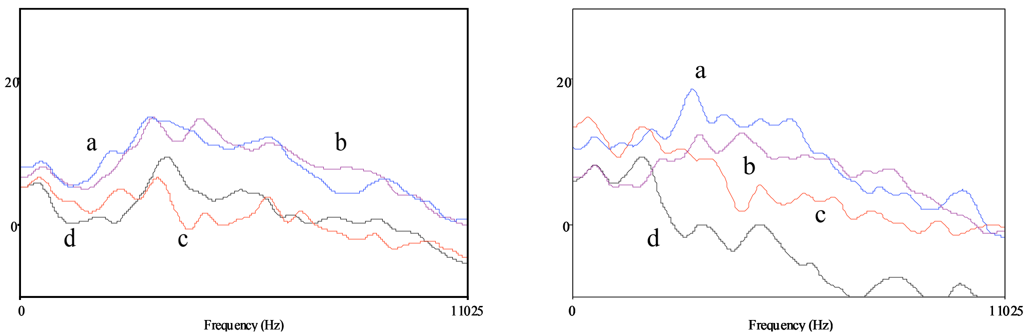
The spectrum of the velar [k] shows a peak around 500 Hz followed by a trough. Patryn (1987) also mentions a peak in the low frequencies as a characteristic of velars, but estimates that it is around 1000 Hz. In the [i] context, we find a second peak only at 3500 Hz, from which on the envelope is descending. In the [a] context, however, there is an early peak around 2000 Hz and a later around 4000 Hz, again with a characteristic sloping down of the envelope.

The spectrum of [tɕ], which according to our thesis should be more similar to [t] than [k], shows in both contexts peaks at 3200 Hz (in accordance with Patryn 1987: 60), 4200 Hz and 6500 Hz, with a downward-sloping envelope. The peak at 3200 Hz is similar to the one of [t] in the context of [i]. Furthermore, [tɕ] and [t] share a high frequency peak (6500 Hz and 6000 Hz, respectively). These two points confirm our hypothesis that [tɕ] and the [k] are acoustically similar. On the other hand, [tɕ] shares the general downward-sloping envelope with [k], which might speak for a similarity between [tɕ] and [k].

In both vowel contexts, [ʧ] has a deep trough at around 1200 Hz and the first high energy peak at 2800 Hz, the second at 6000 Hz. Between the two peaks there seems to be an energy plateau, more prominent in the [a] context than in the [i]. The spectrum of [ʧ] should show some similarities to that of [k] to confirm our hypothesis. Both spectra show a trough at the lower frequencies, but their shape and extension is rather different. With [t], the [ʧ] shares a peak at 6000 Hz, but contrary to the spectrum of [ʧ], [t] does not have an energy plateau before this peak. As an illustration of these descriptions serve the spectra of speaker CM (1) in (19).

(19) Fricative spectra of speaker CM (1) in [i] context (left) and [a] context (right)

k: black/d; ʧ: blue/a; tɕ: purple/b; t : red/c



In sum, the spectra could confirm only part of our original hypothesis. There is spectral similarity between [tɕ] and [t] (and [tɕ] and [k]), but none between [ʧ] and [k].

Summing up the information on the formants, the bursts and the spectra, no clear picture emerges. The formant values and transitions show a similarity between [t] and [ʧ] on the one hand and [tɕ] and [k] on the other. The noise duration scales do not support this similarity, instead they show that [ʧ] and [k] are close, which is in accordance with our expectations. However, no evidence could be yielded from the noise duration scales for the proposed similarity between [t] and [tɕ]. The results from the spectral comparison are exactly opposite, there is spectral similarity between [tɕ] and [t] (as proposed) but none between [ʧ] and [k]. Transition length delivers no criteria according to which the four consonants can be classified, as the values where too similar for all four.

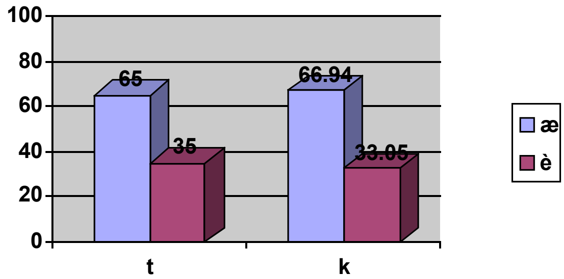
Interestingly, the front high vowel environment allows a greater variability of the values of the F2 transitions between speakers than the low vowel context. As to the burst length, there is a more reliable difference in burst before low vowel than before high front vowel context. Based on this information, we expect that all four consonants are likely to be confused before the front vowel.

6. Evaluation of the perception test

The results of the perception test are discussed in the following order. First we look at the general results in Polish and German. Then the short-long distinction in the signals will be discussed, and finally the importance of the order of the cue is evaluated.

The overall results of the test with Polish listeners are graphically presented in (20). In this and all following graphs in this paragraph, the left-hand, blue (lighter) column indicates the answer [tɕ] and the right, red (darker) column the answer [ʧ].

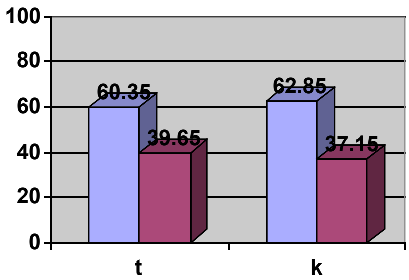
(20) *Polish listeners*



As can be seen from this graph, the Polish native-speakers judged both [t] and [k] to be similar to [tɕ] than to [ʧ]. There is no substantial difference in judgments on the similarity of more[t/tɕ] and [k/tɕ]; the former pair is judged by 65% of subjects as more similar, the latter by 66,94% of the subjects. This result is not in accordance with our expectations, as the [k] was expected to be similar to [ʧ].

The German listeners had similar results as the Polish native speakers. In 60,35 % of the cases, subjects answered that [t] is more similar to [tɕ] (rather than [ʧ]) and in 62,85% of the cases [k] was more similar to [tɕ], cf. (21).

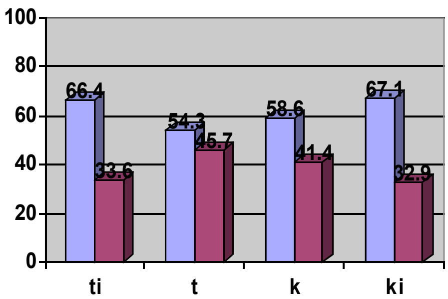
(21) German listeners



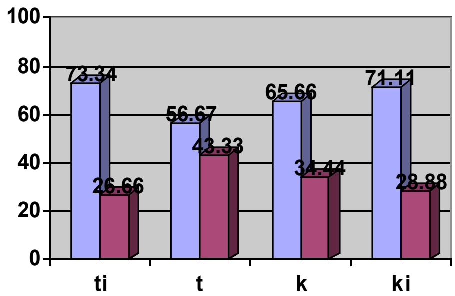
To judge the influence of the length of the cues, the results are split according to the length of the cues. In (22), the difference between long and short signal X, i.e. the one that has to be categorized, is given.

(22) *Length of X*

a. Polish listeners



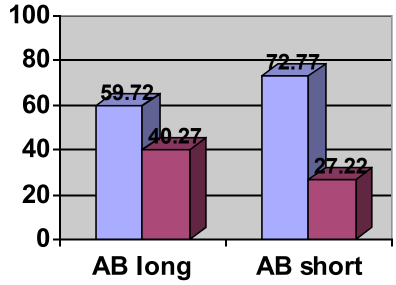
b. German listeners



In these graphs we can see that the addition of the vowel to the signal X changed the listener's classification. The percentage of classifying [ti] or [ki] as [ʧ] rose compared to [t] or [k]. This corresponds to what we expected, because the vowel [i] is actually the trigger for the misperception of a non-palatalized stop as it palatalized counterpart affricate. Thus, it helps to categorize into the provided categories A and B. Interestingly, it did not add so much in the German results. This is due to the fact that the German listeners had in general classification results that were closer to chance, i.e. to 50%.

Apart from the X cues, the perception test varied the length of the AB cues, too. This is exemplified in the graph in (23).

(23) Length of AB cues – results in Polish subjects

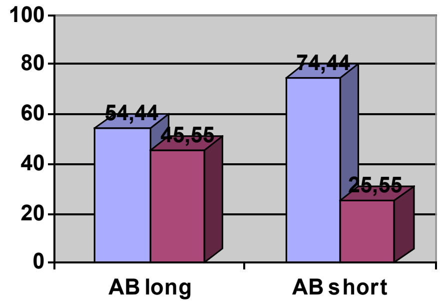


The length of the cue influenced the results in this way, that the shorter A and B signals triggered much clearer results than the signal including a portion of the vowel. E.g. the listeners judged both [t] and [k] as [tɕ] in 59,72% of the cases, and [ti] and [ki] as [tɕ] in 72,77% of the cases.

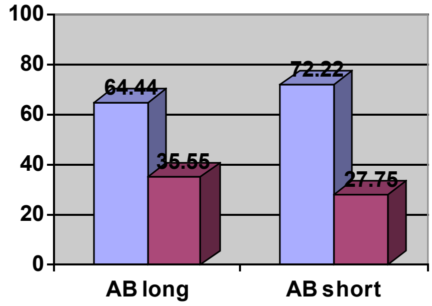
This difference still holds when the test triads with [t] and [k] as X are considered separately, as in (24) and (25). (24) gives the results for Polish listeners, with the categorization of [k] in the left graph and the categorization of [t] in the right one.

(24) *Length of AB cues for Polish subjects*

a. k(i)



b. t(i)

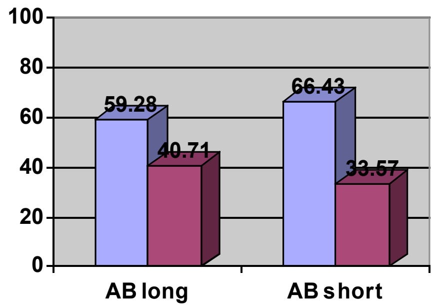


The left graph in (24) is of interest, as it shows that the Polish listeners had problems categorizing [ki] as either [tɕ] or [ʧ], see left columns, as both values are very close to 50% which is mere guessing. Apparently, vocalic transition cues in the target category confused subjects so that they could not make a decision between A and B. For the short [k], however, the Polish listeners had no such problem, which means that there is more similarity between a [k] and a [tɕ] than between a [ki] with the same affricates.

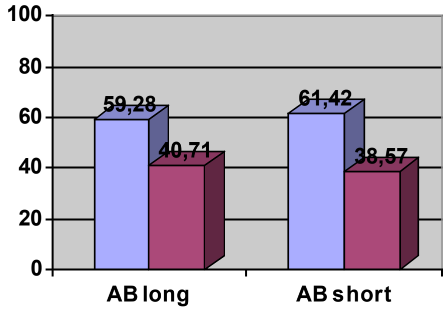
The results of the German test were alike, as represented in the graphs in (25).

(25) *Length of AB cues for German subjects*

a. k(i)



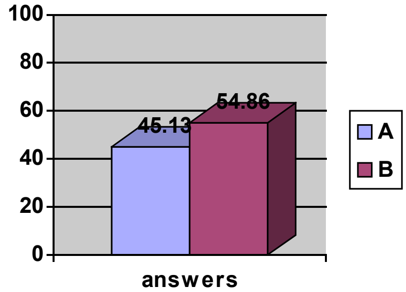
b. t(i)



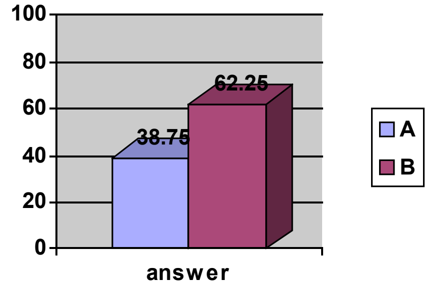
However, the German listeners did not show such a big differences for the [k]-[ki] cues as the Polish ones. This means that they were still able to find some perceptual similarity between [ki] and [tɕ]. From this we can conclude that the Polish speakers behave differently because they are influenced by their lexical knowledge.

To check the possible influence of the bias to the most recent information, let us look at the percentage of A and B answers given by both Polish and German listeners, cf. (26).

(26) a. *Polish listeners*



b. *German listeners*



For German native speakers/listeners, the difference between answers A and B is higher than for Polish ones, which might be a result of lower sensitivity to the relevant cues: listeners do not always perceive substantial difference between A and B, and show a preference for B, as that is the cue they heard last.

In sum, the results of the perceptual test do not confirm the original hypothesis . [k] both with and without vowel [i] is perceptually not more similar to [ʧ] then to [tɕ]. Instead, both [t] and [k] are judged similar to [tɕ]. Thus, the selection of alternating pairs in palatalization in Polish is not based on the criterion of relative perceptual similarity according to the perceptual scale proposed in § 2.

7. Conclusions

In the acoustic part of the experiment we managed to single out a property that alternants in velar palatalization have in common, namely the length of the noise. The alternants in coronal palatalization are similar in respect to the shape of the envelope, but there is also a similarity between the velar and the alveolopalatal sounds (the input of Velar Palatalization and the output of Coronal Palatalization). These findings were confirmed in the perception test, where across all conditions and irrespective of the native language of subjects both [t] and [k] were judged more similar to [tɕ]. The results from the perception test show that the palatalization pairs in Polish are not motivated by perceptual similarity, at least not the [ʧ] - [k] alternation.

Thus the perceptual hierarchy posed before has to be modified to the perceptual scale given in (27).

(27) t - tɕ - k
|
ʧ

According to this hierarchy, both [t] and [k] are perceptually similar to [tɕ], whereas [ʧ] only shows some similarity to the alveolopalatal affricate (because of similar modus). This hierarchy, or better similarity association, mirrors the articulatory similarity described in §1. There we elaborated that if the trigger had a major influence on the output, then both palatalization processes should result in [tɕ], because this affricate has a similar tongue shape and place of articulation as the front vowel [i].

Contrary to our expectations, there is no functional perceptual/articulatory account for the present day outcome of velar and coronal palatalization other than that the outputs of coronal and velar palatalization need to be perceptually distinct from each other[11]. The choice of outputs for each set of inputs is in synchronic terms arbitrary. One has to assume that velars have to alternate with post-alveolars due to information stored in the lexicon.

Nevertheless it is possible to explain the diachronic development of the present day palatalization system as driven by functional factors. The input-output correlations once were functional, either by articulatory assimilation, perceptual similarity, or by maximal contrast to other existing input-output pairs (see §2). These pairs then were lexicalized, and the phonetic surrounding responsible for the correlation was lost and is therefore not directly accessible anymore.

Historically, coronal palatalization was an alternation between dentals and secondarily palatalized dentals in the context of a front vowel. From the secondarily palatalized dentals the present-day prepalatals developed most probably in the 13th century (Długosz-Kurczabowa 1993, Klemensiewicz 1985). The historical development is summarized in (28).

(28) ti > t^{j}i > tɕ i

Old Polish soft (palatalized) post-alveolars (originating from, among others, Velar Palatalization) got depalatalized in Old Polish by the 16th century, see (29).[12]

(29) CS ki > ʧ^{j}i > ʧ

From Stieber's (1966) claim that /tɕ/ and /ʧ^{j}/ expressed a semantic distinction in the 15th century, one can conclude that the two coexisted at a certain point, though perceptual distance between the two must have been very little and the need for more perceptual distinction was likely to trigger further mutation of the alternants of velars. However, why is it velar alternants that transformed, and not the alternants of dentals? The change of [ʧ^{j}] to [ʧ], which increased the perceptual distance between alternants in velar palatalization might be explained by the fact that velar palatalization and its output were already well established in the language system (i.e., lexicalized) at the time coronal palatalization entered the language in its modern shape. [tɕ] could thus change to [ʧ], without threat of perceptual confusion, unlike in the hypothetical case when newly established, the alveolopalatal alternant of [t] would have transformed to [ʧ]. In this sense, functional factors referring to the stability of the language system and not phonetic factors like perception and articulation are responsible for the present alternations in Polish velar and coronal palatalization.

---

[11] The coalescence of the outputs of the two palatalization processes would mean loss of contrast between coronal and velar consonants in the palatalizing context, which is against the tendency to preserve maximal number of contrasts within the inventory of a language (Flemming, 2001).

[12] ʧ^{j} marks here a secondarily palatalized soft sound, probably similar to the English palatoalveolar, as opposed to ʧ of present day Polish which is not secondarily palatalized.

Acknowledgements

We would like to thank for the comments on this paper to: Tracy A. Hall, Bernd Pompino-Marschall, and Sophie Repp.

References

Bhat, D.N.S. 1978. A General Study of Palatalization. In: J.H. Greenberg (ed.), *Universals of Human Language, Vol.2., Phonology.* Stanford, Ca.: Stanford University Press.

Boersma, P. 1998. *Functional Phonology*. The Hague: Holland Academic Graphics.

Calabrese, A. 1993. Palatalization processes in the history of romance languages: A theoretical study. In: W. J. Ashby, M. Mithun, G. Perissinotto & E. Raposo (eds.), *Linguistic Perspective on the Romance Languages*. Amsterdam: John Benjamin.

Ćavar, M. 1997. *Coronal sounds and palatalization: an analysis of Polish and English*. Unpublished Master thesis, Warsaw University.

Chang, S.S., M.C. Plauché & J.J. Ohala. 2001. Consonant Confusion Asymmetries. In: Hume & Johnson (eds.); 79-101.

Charbonneau, R. & B. Jacque. 1972. [t_s] et [d_z] en français canadien. In: A. Valdman (Hrsg.). *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*. Den Haag: Mouton; 77-90.

Clements, G. N. 1999. Affricates as Noncontoured Stops. In: O. Fujimura et al. (Hrsg.) *Item, Order in Language and Speach*. Prague: Charles University Press.

Długosz-Kurczabowa, K. 1993. *Gramatyka historyczna języka polskiego*. Wydawnictwa Uniwersytetu Warszawskiego.

Flemming. E. 2001. *Auditory Representations in Phonology*. Routledge: London.

Guion, S.G. 1998. The Role of Perception in the Sound Change of Velar Palatalization. *Phonetica*: 55, 18-52.

Hamann, S. 2002. Are postalveolar fricatives in Slavic Languages retroflex? Ms. to be published in: S. Baauw, M. Huiskes & M. Schoorlemmer (eds.) *OTS Yearbook 2002*, Utrecht.

Hume, E.V. 1991. *Front vowels, coronal consonants and their interaction in non-linear phonology.* Ithaca; N.Y.: Cornell University Ph.D. dissertation.

Hume, E. & K. Johnson. 2001. *The Role of Speech Perception in Phonology.* San Diego, Ca. : Academic Press.

Johnson, K. 1997. *Acoustic and Auditory Phonetics*. Oxford: Blackwell.

Keating, P. 1991. Coronal Places of Articulation. In: Paradis & Prunet (eds.); 29-48.

Kim, H. 2001. A phonetically based account of phonological stop assibilation. *Phonology* 18: 81-108.

Klemensiewicz, Z. 1985. *Historia języka polskiego.* Warszawa: Państwowe Wadawnictwo Naukowe.

Ladefoged, P. and I. Maddieson. 1996. *The Sounds of the World's Languages*. Oxford, UK: Blackwell.

Lahiri, A. and V. Evers. 1992. Palatalization and Coronality. In: Paradis & Prunet (eds.) ; 79-101.

Łobacz, P. 1985. *Fonetyczno-leksykalne interakcje w percepcji mowy.* Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.

Ohala, J. 1992. What's cognitive, what's not, in sound change. In: M. Kellerman (ed.) Diachrony within synchrony: language history and cognition. *Duisburger Arbeiten zur Sprach- und Kulturwissenschaft* 14: 309-355.

Ohala, J. 1993. The perceptual basis of some sound patterns. In: B.A. Connell & A. Arvaniti (eds.) *Phonology and phonetic evidence: Papers in laboratory phonology IV*. Cambridge: CUP; 87-92.

Padgett, J. 2001. Contrast Dispersion and Russian Palatalization. In: Hume & Johnson (eds.); 187-218.
Paradis, C. & J.-F. Prunet. 1993. *The special status of coronals: Internal and external evidence (Phonology and Phonetics 2)*. San Diego, Ca.: Academic Press.
Patryn, R. 1987. *Phonetic-acoustic analysis of Polish speech sounds*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
Pickett, J.M. 1999. *The Acoustics of Speech Communication. Fundamentals, Speech Perception Theory, and Technology*. Boston etc.: Allyn and Bacon.
Rubach, J. 1984. *Cyclic and Lexical Phonology. The Structure of Polish.* Dordrecht: Foris.
Rubach, J. 1993. *The Lexical Phonology of Slovak*. Oxford: Clarendon Press.
Steriade, D. 1997. Phonetics in Phonology: The Case of Laryngeal Neutralization. Ms UCLA.
Steriade, D. 2001. Directional Asymmetries in Place Assimilation: A Perceptual Account. In: Hume and Johnson (eds.); 219-250.
Stevens, K. 1998. *Acoustic Phonetics*. Cambridge, Massachusetts: MIT Press.
Szpyra, J. 1995. *Three Tiers in Polish and English Phonology.* Lublin: UMCS.
Wierzchowska, B. 1971. *Wymowa Polska.* Warszawa: Państwowe Zakłady Wydawnictw Szkolnych.
Wierzchowska, B. 1980. *Fonetyka I Fonologia Języka Polskiego*. Wrocław: Zakład Narodowy imienia Ossolińskich.
Winitz, H., M.E. Scheib & J.A. Reeds. 1972. Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *JASA* 51: 1309-1317.