Do palatal consonants correspond to the fourth category in the perceptual F2-F3 space?

Geng, C.[1], *Mády, K.*[2], *Bogliotti, C.*[3], *Messaoud-Galusi, S.*[3], *Medina, V.*[3] *& Serniclaes, W.*[3]

[1]Centre for General Linguistics, Typology and Universals Research (ZAS)
[2] Pázmány Péter Catholic University, Institute of German Studies (Budapest)
[3]Laboratoire de Psychologie Expérimentale, CNRS & Université René Descartes (Paris 5)

Corresponding author: geng@zas.gwz-berlin.de

Abstract

The research on the perception of place of articulation in oral stop consonants has almost exclusively focused on the bilabial, alveolar and velar places of articulation, most common in the sound inventories of the world's languages, whereas the palatal place of articulation has been left aside. One language that phonemically has a palatal stop is Hungarian. Acoustically, the palatals exhibit formant transitions similar to the alveolars. Place perception in Hungarian was investigated in a combined identification-discrimination-experiment using synthetic /Cə/-stimuli. The onset frequencies of F2 and F3 transitions were systematically modified either alone or with concomitant changes of burst frequencies. The analyses performed so far show that territorial mapping of the responses in stimulus space for palatal and alveolar place together resembles that of the alveolar place alone in a three-category language.

1. Introduction

In recent years, the cross-linguistic plasticity of perceptual boundaries has drawn a lot of attention. It is often assumed that natural phonetic boundaries in the pre-linguistic child are replaced by language-specific ones. For consonantal perception, the most prominent empirical domains encompass research on voicing and stop place of articulation perception. Specifically, for place perception, a vast amount of studies on the transitional cues has evolved in the tradition of Delattre et al. [1]. In our view, it is noteworthy that this research has almost exclusively focused on the three oral stop place categories bilabial, alveolar and velar, leaving the palatal place of articulation aside. Aiming at an integration of the palatal into the general topology of place of articulation, we will first (a) achieve the description of contemporary approaches to speech category formation, (b) review selected evidence in favor or against these approaches and (c) motivate the selected empirical/experimental approach adopted here with respect to the choice of stimuli and language.

1.1. Models of speech category development

As systematized by Phillips [2], different hypotheses are conceivable for the mapping of natural phonetic settings onto phonological ones: One, as advocated by [3], postulates that the natural, phonetic boundaries stay present in the adult population to various degrees, i.e. a subset of the innate perceptual boundaries are taken to form a phonological decoding level. In Phillips' terms, an approach like this would be called a "structure-adding" approach. Contrastively, approaches, in which the change from infant to adult perceptual space involves rather a reshaping of the infant representations than the selection among existing ones, are termed "structure-changing" approaches: Theorizing in line with Kuhl's Native Language Magnet [e.g. 4], prelinguistic representations are replaced by new ones through perceptual magnet effects. In Kuhl's approach, categorical perception phenomena are related to the prototypes: In the vicinity of the phoneme prototype, vowel discrimination is more difficult, i.e. exemplars close to the prototype are hard to distinguish, while the opposite holds for exemplars farther away from the prototype. Note that the prototype model, although primarily invented for vowel perception has been extended to consonant-place-identification [5]. A third possibility synthesizes claims from structure-adding and structure-changing approaches [6]: In category-development, new category boundaries are created through perceptual couplings between predispositions. However, primitive boundaries still remain discriminable in some conditions, and may be conducive to "allophonic perception", as in people affected by dyslexia [7]. We will return to this point later. Before that, we have to motivate the presence of natural boundaries in the perception of place of articulation.

1.2. The choice of the neutral vowel

Among others, [8] and [9] demonstrated that the perception of place of articulation in stop consonants depends on both burst and transition cues. After Smits et. al [7], the relative importance of burst- and transition cues depends on voicing, place of articulation and vowel context: So first, with respect to cue-trading, by taking stimuli from same voicing category (i.e. voiced stop) and in the same vocalic context (stop + schwa syllables) one can fix trading relations between burst and transition cues. Second, the schwa has been assigned a prominent developmental role: The first vowel productions of infants are central vowels resulting from an open and neutral shaping of the vocal tract, more peripheral vowel qualities are acquired later, i.e. developmentally, the vowel space expands toward the periphery [10]. Related to this line of reasoning, the central vowel plays a structuring role in the development of consonant-place-perception. Further, the central vowel might also have been a catalyzer in phylogenetic development [11]. In the neutral vocoïd context, both rising F2 and F3 transitions represent bilabials, both falling F2 and F3 represent coronals and falling F2 and rising F3 represent velars. Place boundaries are then related to natural psychoacoustic boundaries in the schwa context, i.e. boundaries are related to flat transitions as the latter constitute the limit between rising and falling frequency movements. It is then possible that natural boundaries were used as primitives for developing place articulation distinctions in speech communication and that the neutral vocoïd context acted as a reference in this process. Related to this line of reasoning, the central vocoïd plays a structuring role in consonant-place-perception for both phylogenetic and ontogenetic reasons.

1.3. Previous Studies

The focus on this particular issue emerged as the result of previous research [12,13]: When separately modified, the perceptual F2-F3 space is divided into four distinct "natural" regions, delimited by flat F2 or F3 formant transitions (see Figure 1). Note that the upper left quadrant of the acoustic space, corresponding to raising F2 and falling F3 transitions is "empty". That is, none of the places of articulation for a language with three place categories does correspond to rising second and falling third formant transitions. The partitioning of this acoustic space into the three phonologically relevant stop place categories present in the French language was previously examined with perceptual data.

Results showed that the velar region occupied roughly the lower right quadrant, corresponding to falling F2- rising F3 transitions. The two other categories, labials and coronals, share the remainder of the space with a perceptual boundary located inside the empty raising F2- falling F3 region, though not in the middle of that region (see Figure 1). Results also indicated that non-phonological boundaries remained discriminable for the French adult subjects, thereby showing that even when natural boundaries are located inside phonological categories, they can still affect consonant discrimination. This result has been interpreted as evidence for a residual perceptive sensitivity for natural boundaries.

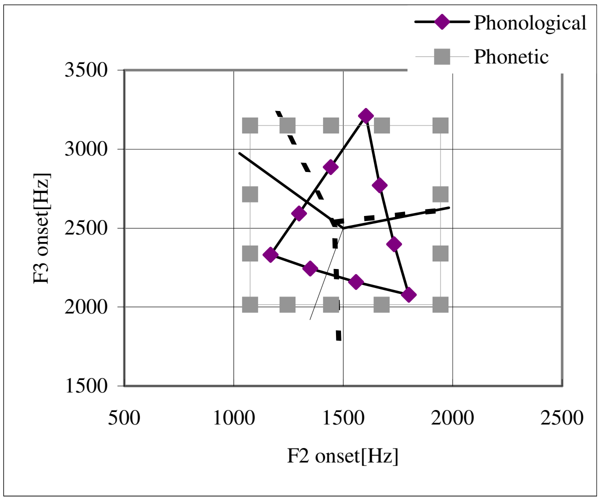

Figure 1. Stimulus continua. The rectangle corresponds to the "phonetic" continuum. The 4 component lines follow directions normal to the natural boundaries, corresponding to flat formant transitions with onsets at F2=1500 Hz and F3=2500 Hz. The triangle corresponds to the "phonological" continuum. The 3 component lines follow directions normal to the expected phonological boundaries, which divide the acoustic space in 3 equal-sized regions. Broken lines indicate the location of perceptual boundaries between places of articulation in French.

1.4. The inventory of Hungarian stop sounds

The choice of Hungarian is motivated by the presence of palatal consonants in the sound inventory. Their phonemic status as plosive or affricate has been a matter of debate though: For example, the Hungarian palatals are treated as the affricates /cç, ɟʝ/ in the Handbook of the IPA [14], whereas other authors treat them as the plosives /c,ɟ/. Siptár & Törkenzy [15] provide a summary about this issue.

Combined acoustic and magnetometer recordings of Hungarian palatals were collected in the ZAS laboratory [16]. The basic results concerning burst information are heterogeneous between speakers: A subset of our participants exhibited clear bursts, whereas the patterns for a second subset of speakers still requires more detailed acoustic analysis. This implies that if the burst information is an unstable cue for perception of palatal place, then the usage of transition information might be of crucial importance for cueing the palatal place. As we have shown [16], regardless of the vowel environment, the articulatory configurations closely resemble an [iː]-like configuration. Therefore it seems reasonable to assume the transitions of F2 and F3 to be falling towards the neutral vowel. This assumptions gains plausibility by studies investigating palatal glides, for example, according to Chafcoulof [17], the palatal glide /j/ is characterized by the following formant structure: F1=250-275 Hz, F2=2500-2700 Hz, F3=3000-3100 Hz.

1.5. Aims of the study

Summarizing the above, the basic situation is as follows: There are four psychoacoustic regions shared by four place-categories. The coronal quadrant with falling F2 and F3 is populated by both palatal and alveolar place categories resulting in an ambiguous mapping between psychoacoustic region and place inventory of Hungarian.

So, the general objective is to explore how to relate the additional palatal stop place of articulation of a four category-language like Hungarian to the partitioning of stimulus space described above.

2. Method

We used the same stimuli as in the French study [12]. 23 stimuli CV were generated with a parallel formant synthesizer. F1-F2-F3 transitions ended at 500, 1500 and 1500 Hz respectively after a 27 ms transition. The VOT was set to -95 ms, the stable vocalic portion of 154 ms. The stimuli differed as to the onset of F2 and F3 transition. 14 stimuli were generated by separate modification of the F2 and F3 onsets along a "phonetic" continuum, normal to the locations of the natural boundaries – corresponding to either flat F2 or F3 transitions- as shown in Figure 1. Nine other stimuli were generated by joint modification of the F2 and F3 onsets along a "phonological" continuum normal to the expected category boundaries separating the F2/F3-space into three distinct regions (see Figure 1). The same amount of stimuli was generated with the same basic data but an additional constant burst.

Successive stimuli were 1 Bark apart on each continuum. Stimuli were presented binaurally with Sennheiser HFD 590 headphones. The subjects had the possibility to adjust the volume to a comfortable level. Participants were (a) participants of an undergraduate linguistics course or (b) volunteers contacted by an automatic email, their native language was Hungarian. Most of them were participants of undergraduate exchange programs. They were between 18 and 53 years old with no reading or hearing impairment reported. The continuum without burst was presented to a subset of the participants (7 subjects). Both continua were presented to a second subset of the participants. For these subjects (12), the continua with and without burst were presented in alternating order.

3. Results

A first question that arises is whether identification and discrimination data collected here for Hungarian conform to the patterns which have been described in the framework of Categorical Perception, i.e. whether the data exhibit (a) discrimination maxima at the category boundaries and (b) discrimination peaks which are predictable by the identification function. For this purpose, we carried out an analysis of variance with percent correct responses as the dependent variable, burst (two levels, factor BURST), predicted versus observed (two levels, factor PRED) and stimulus number (14 resp. 9 levels, factor STM) as within-subject factors. The order of presentation (2 levels, ORDER) was treated as a between-subject factor. Correct observed discrimination response scores were calculated for adjacent stimuli on each continuum as the mean correct response to both same pairs and different pairs. The predicted discrimination scores were calculated by a formula adapted from [18]. Discrimination results are presented in Figure 2.

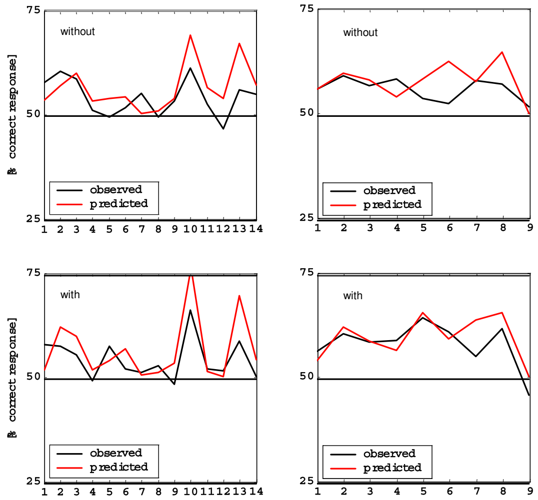

Figure 2: Mean observed and predicted discrimination scores for the stimuli without (top panels) and with burst (bottom panels). Results for the "phonetic" continua are shown on the left, for the "phonological" continua on the right.

Table 1: Repeated measurement ANOVA of predicted and observed scores

	Phonetic			Phonological		
Factor	**df**	**F**	**p**	**df**	**F**	**p**
Order	1	.003	.960	1	.009	.924
Pred	1.0	22.0	.000	1	3.17	.093
Stm	6.3	8.50	.000	5.61	4.57	.001
PredxStm	11.3	1.86	.047	6.43	1.63	.141

For the phonetic continuum, Mauchly tests of sphericity yielded significant values for STM and the STM x PLOS interactions. For the phonological continuum, STM,(1%), and the triple interaction PLOS x PRED x STM(5%), reached the level of significance. Therefore, degrees of freedom were Huynh-Feldt-corrected, and corrected values of significance and degrees of freedom will be reported. The significant effects and the results for the between-subject factor ORDER are shown in Table 1.

The interaction between PRED and STM is significant (p<.05) for the phonetic continuum. This indicates a failure of the prediction formula for some stimuli. Visual inspection of observed versus predicted discrimination scores (Figure 2) showed that this failure is local to the stimuli with high F3, i.e. the boundaries between palatal and alveolar and alveolar and bilabial are selectively affected. Observed discrimination peaks fall closer to the natural boundaries, i.e. to flat transitions lines in the F2-F3 onset space, than did the peaks expected from the labeling data.

The same kind of analysis was carried out with the observed discrimination scores only. The factors are identical to those in the previous analysis except that the within-subject factor for PRED is missing. The significant results and the results for the ORDER between-subject factor are reported in Table 2. Again, some tests of sphericity were significant (phonetic continuum: STM, STM x PLOS, (1%-level); phonological continuum: STM (1%) and the triple interaction PLOS x PRED x STM(5%)).

Table 2: Repeated measurement ANOVA of predicted scores

	Phonetic			Phonological		
Factor	**df**	**F**	**p**	**df**	**F**	**p**
Order	1	.002	.964	1	.019	.893
Stm	11	3.94	.000	7.70	1.77	.092

For the analysis only including observed discrimination scores, i.e. dropping the PRED within-subject factor, generally fewer effects reach the significance level: Only the STM main effect on the phonetic continuum reached the level of significance.

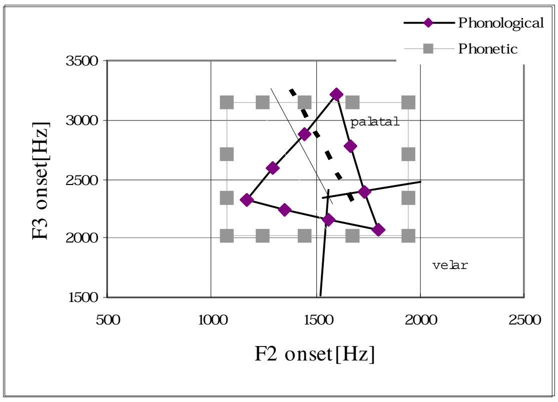

Figure 3: Labeling results obtained by multinomial logistic regression.

For the purpose of a graphical display of the partitioning of the stimulus space, we carried out a multinomial logistic regression of schwa-centered formant values on the identification data (Figure 3). The most salient aspect here is that the four categories do not partition the perceptual space into regions of equal size, in particular, the alveolar place only occupies a rather small portion of the stimulus-space.

Note that the boundaries do not cross at 1500Hz-2500Hz either, which is evidenced by significant intercept values of the logistic models. However, the cross point is not far away from 1500Hz-2500Hz (Figure 3).

4. Summary and Discussion

We have conducted combined identification/discrimination experiments on native Hungarian speakers with synthetic stimuli on two different continua in two burst conditions. The first basic results reported in this paper include the way, the perceptual F2/F3 space investigated by two different stimulus continua is partitioned by a language with four place-categories in contrast to a language which has only three categories like French. At the moment, we have to restrict ourselves to a qualitative comparison of the territorial map shown in Figure 3 with data presented in [13]: One possible outcome of the local increase in category density would be a substantially increased region covered by palatal and alveolar places together in comparison to the region covered by the alveolar alone in a three-category language. However, this is not the case: The category boundary between alveolar/palatal and bilabial is at a location similar to the French sample used in [13]. If the coronal space should be considered as enlarged at all, this occurs at the expense of a smaller velar region. We have described theories of speech category development in the introductory section. The approach which postulates the greatest independence of adult percepts from natural boundaries and thus the greatest modifiability of category boundaries is the structure-changing Perceptual Magnet. The data we have shown in this study are challenging with respect to this claim: Phonemic boundaries appear to be remarkably constant over languages, with the exception of an intrusive alveolar-palatal boundary for Hungarian located inside the French coronal region. One possibility of an explanation arises from constraints on the production, i.e. sounds with rising F2 and falling F3 are “difficult to produce” [19], and so the listener has learned to restrict the coronal category only to a subpart of the upper left quadrant. Yet, examination of the discrimination results reveals four natural categories indicated by four discrimination peaks. This converges with the second result we have reported: There is considerable variation in the labeling performance and the location of the category boundaries we have obtained. As a consequence, we have observed a failure of the CP prediction formula concerning the boundaries involving palatal and alveolar identification. This is evidenced by the significant interaction term between stimulus number and the factor coding predicted versus observed discrimination scores. It will be one of our next steps in the analysis to disentangle the exact sources and mechanisms of these inter-speaker differences in categorization by setting up post-hoc contrasts between subpopulations of the present sample, which could be revealing about the strategies employed to balance psycho-acoustic and phonological categories.

Acknowledgements
We wish to thank Márton Méhes for his help hiring subjects. This research was funded by GWZ 4/10-1, P.1 to the German Research Council.

5. References

[1] Delattre, P.C., Liberman, A.M., & Cooper, F.S., “Acoustic loci and transitional cues for consonants”, *Journal of the Acoustical Society of America*, 27:769-773, 1955.

[2] Phillips, C., “Levels of representation in the electro-physiology of speech perception”, *Cognitive Science*, 25:711-731, 2001.

[3] Werker, J.F. & Tees, R.C., “Cross-language speech perception: evidence for perceptual reorganization during first year of life*”, Infant Behavior and Development*, 7:49-63, 1984.

[4] Kuhl, P. K., “Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not”, *Perception & Psychophysics*, 50, 93-107, 1991.

[5] Davis, K. & Kuhl, P.K., “Tests of the perceptual magnet effect for American English /k/ and /g/”, *127th ASA Meeting,* 4pSP17, 1994.

[6] Serniclaes,W., ”Etude expérimentale de la perception du trait de voisement des occlusives du français”, Université Libre de Bruxelles, Bruxelles, 1987.

[7] Serniclaes, W., Van Heghe, S., Mousty, Ph., Carré, R. & Sprenger-Charolles, L. Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology, 87*: 336-361, 2004.

[8] Smits, R., ten Bosch, L. & Collier, R., “Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. I. Perception Experiment“, *JASA*, 100(6):3852-3864, 1996

[9] Fischer-Jørgensen, E., “Tape-cutting experiments with Danish stop consonants in initial position”, *Annual Report of the Institute of Phonetics, University of Copenhagen, 6:104-168, 1972.*

[10] Kent, R.D. & Murray, A.D., “Acoustic features of infant vocalic utterances at 3, 6, and 9 months”, J. Acoust.Soc. Amer.72, 353–365, 1982..

[11] Carré, R., Liénard, S., Marsico, E., & Serniclaes, W., “*On the role of the "schwa" in the perception of plosive consonants*”, in J.L.H. Hansen & B. Pellom (Eds.). 7th International Conference on Spoken Language Processing, 1681-1684, 2002.

[12] Serniclaes, W., Bogliotti, C. & Carré, R., “Perception of consonant place of articulation: phonological categories meet natural boundaries”. In M.J.Solé, D.Recasens & J.Romero (Eds.). *Proc. 15th International Congress on Phonetic Sciences*, 391-394, 2003.

[13] Bogliotti, C., “Perception categorielle et perception allophonique: Incidences de l'âge, du niveau de lecture, et des couplages entre predispositions phonétiques”, *PhD. Thesis, Université Paris 7 - Denis Diderot*, 2005.

[14] Szende, Tamás, “Hungarian”, *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*, Cambridge, Cambridge University Press, 104-107, 1999.

[15] Siptár, Péter & Törkenczy, Miklós, *“The phonology of Hungarian”*, Oxford: Oxford Univ. Press, 2000.

[16] Geng, C. & Mooshammer, C. “The Hungarian palatal stop: Phonological considerations and phonetic data”, *ZAS Papers in Linguistics,* 37, 221-246, 2004.

[17] Chafcoulof, M., “Les caractéristiques acoustiques de (j,w,l,r)“, *Travaux de l'Institut de phonétique d'Aix*, 7:7-56, 1980.

[18] Pollack, I., & Pisoni, D., “On the comparison between identification and discrimination tests in speech perception”, *Psychonomic Science*, 24:299-300, 1971.

[19] Carré, R. (Submitted). “On the phonetic characteristics of an acoustic tube: Vowel and consonant systems”.