Chapter 14

Cross-linguistic trends in the perception of place of articulation in stop consonants: A comparison between Hungarian and French

Authors: W. Serniclaes[*] *& C. Geng*[**]

[*] *Laboratoire de Psychologie de la Perception, CNRS & Université René Descartes (Paris 5)*

[**] *Center for General Linguistics, Typology and Universals Research (ZAS), Berlin*

.1 Introduction

A basic question in the study of speech development during infancy is to understand how the perceptual features evidenced in the predispositions of the pre-linguistic child contribute to phonological categories in a given language. Possible answers in current theories of speech development are that phonological categories (1) are acquired through selection of predispositions relevant for perceiving phonetic features (Pegg and Werker, 1997); (2) emerge by building up prototypes in an acoustic space, without straighforward relationship with the predispositions (Kuhl, 2004). According to selectionist apporaches in its strongest form, phonological contrasts should conform to the perceptual predispositions and adaptation to the language environment would then proceed by mere selection of predispositions (Pegg and Werker, 1997). The orthogonal position votes for a language-specific morphing of the acoustic space and almost free configurability of vowels in this space. Here, it becomes hard to avoid associating an analogy with the build-up of vowel systems: Liljencrants &

Lindblom (1972; Lindblom, 1986), in their seminal paper, were the first to relate phonetic principles of perceptual contrast to the structure of vowel inventories and their sizes. In short, languages prefer vowels, which are maximally distinct for the perceiver and are produced with the least effort for the producer. However, in the field of speech development, the hypothesis of free configurability in the acoustic space cannot account for the constraints arising from perceptual predispositions. While the prototype approach of speech development matches nicely with dispersion approaches to phonological systems, it does not account for the role of perceptual predispositions. A compromise between the strong inneist position and the prototypical approach is that adaptation of the perceptual predispositions to the linguistic paradigm of the ambient language proceeds not only by selection but also by combinations between predispositions (Serniclaes, 2000; Hoonhorst, Colin, Radeau, Deltenre, & Serniclaes, 2006). As there is fairly strong evidence for the existence of predispositions for the perception of virtually all possible phonetic contrasts in the world's languages (e.g. Vihman, 1996), we cannot avoid contemplating the constraints imposed by such predispositions on the build up of phonological categories prevailing in a specific language. Here we present some further evidence in support of this view in the empirical domain of consonant-place perception, without forgetting to mention that consonantism in general must be seen as the stepchild of dispersion-theoretic approaches in both aqusitional and systemic (but see Abry, 2003) domains.

A basic question in the study of speech development during infancy is to understand how the perceptual features evidenced in the predispositions of the pre-linguistic child contribute to phonological categories in a given language. A similar question is to understand how phonetic features contribute to the genesis of phonological systems. These questions are fairly similar as they pertain to the build-up of categories from a universal set of features. Possible answers in current theories of speech development are that phonological categories (1) are acquired through selection of predispositions relevant for perceiving phonetic features (Pegg and Werker, 1997); (2) emerge by building up prototypes in an acoustic space, without straighforward relationship with the predispositions (Kuhl, 2004). Somehow in analogy, phonetic models of cross-linguistic diversity describe phonological systems as (1) combinations between phonetic features (Jakobson, Fant, & Halle, 1952); (2) optimisation of distances between categories in some acoustic space (Lindblom, 1986).

A central problem for developmental and systemic models is the build-up of the relationship between perception and production. This issue was most directly addressed in the distance based (does it need a hyphen: distance-based?) models which call on the maximization of the distances between categories in an acoustic space, a process which is either in charge of the language system or of the child when acquiring the language. The language maximizes the perceptual distances between categories (Lindblom, 1986) and the perceptual system finds back the categories by creating maximally different prototypes (Guenther & Bohland, 2002).

While the prototypical approach of speech development matches nicely with distance models of phonological systems, it does not account for the role of perceptual predispositions. As there is fairly strong evidence for the existence of predispositions for the perception of virtually all possible phonetic contrasts in the world's languages (e.g. Vihman, 1996), we cannot avoid contemplating the constraints imposed by such predispositions on the build up of phonological categories prevailing in a specific language. In its strongest form, phonological contrasts should conform to the perceptual predispositions and adaptation to the language environment would then proceed by mere selection of predispositions (Pegg and Werker, 1997). However, this seems hardly tenable in view of the diversity of phonological contrasts and their plasticity across phonetic contexts. A compromise between the strong inneist position and the prototypical approach is that adaptation of the perceptual predispositions to the linguistic paradigm of the ambient language proceeds not only by selection but also by combinations between predispositions (Serniclaes, 2000; Hoonhorst et al., 2006). Here we present some further evidence in support to this view.

Models of speech perception development

Adults listerners display "categorical perception": they only perceive differences between speech sounds wich differ according to some feature, not the acoustic variations between sounds belonging to the same feature category (see for a review: Harnad, 1987). Predispositions for categorical perception have been evidenced in the pre-linguistic child (below six months of age, see for a review: Vihman, 1996). For instance, whatever their linguistic background (Spanish, Kikuyu), infants younger than six months are sensitive to both negative and positive natural VOT boundaries (Lasky, Syrdal-Lasky and Klein, 1975). Similar predispositions were evidenced for place of articulation (Eimas, 1974). The initial ability to discriminate the universal set of phonetic contrasts however appears to decline in the absence of specific language experience and the decline occurs within the first year of life (Werker & Tees, 1984a) and that it involves a change in processing strategies rather than a sensorineural loss (Werker & Tees, 1984b). Finally, repeated experience to the sounds of a given language also gives rise to facilitation effects (Kuhl, Stevens, Hayashi, Deguchi, Kiritani and Iverson, 2006).

How do phoneme categories arise from predispositions? One possible answer to this question is that phonological features are acquired through selection of predispositions relevant for perceiving categories in a given language (Pegg & Werker, 1997). Another possibility is that phonemes, i.e. phonological categories rather than phonological features, emerge by exposure to the sounds present in a given language without relationship with the predispositions (Kuhl, 2004).

While it seems evident that adaptation to a specific language does not proceed entirely through selective processes, there is some evidence that the emergence of phonological percepts is somehow constrained by predispositions. Specifically, voicing perception in French, and other language with similar voicing categories, seemingly emerges through couplings between predispositions (Serniclaes et al., 2004). For instance, in French speaking environment the discrimination of VOT in young infants below the of age is organized around boundaries located around some -30 ms and +30 ms whereas infants above six months discriminate the adult VOT boundary which is located at 0 ms in this French (Hoonhorst et al., 2006).

The phonological coupling hypothesis (Serniclaes, 1987; 2000) explains the emergence of non-native language specific boundaries by the interplay between predispositions for feature perception and exposure to linguistic categories. For voicing, predispositions for perceiving either negative or positive VOT would be used jointly in two-category languages with negative vs. positive VOT contrasts, thereby explaining the 0 ms VOT boundary. Specifically, there is evidence to suggest that there are interdependencies in the perception of negative and positive VOT (Serniclaes, 1987). Such interdependencies are a special instance of the general concept of "coupling" between perceptual entities (Koffka, 1935; Hochberg, 1981; for a review see: Masin, 1993).

Place of articulation perception

The perception of place of articulation in stop consonants offers another interesting instance for evidencing couplings between features. There are basically two different kinds of acoustic cues involved in place perception: those carried by the burst and those carried by the formant transitions. Although these cues might correspond to different features, as a startpoint we will consider here that burst and transitions convey different acoustic information about the same feature. Further, we will start from the transition-based description of the place features afforded by the Distinctive Region Model (DRM) of place production (Carré & Mrayati, 1991). The DRM is organized around the neutral vowel (schwa) as a central reference. In the neutral vowel context, place boundaries tend to correspond to flat F2-F3 transitions, the categories being characterized by rising vs. falling transitions (Figure 1). This suggests that place perception is grounded on a "natural" boundary in the neutral context (Serniclaes & Carré, 2002), but which has to undergo specific adjustments in other contexts. The place boundary is shifted towards falling transitions before back rounded vowels, rising transitions before front unrounded vowels, and intermediate positions before front rounded vowels. The radial model of place perception states that the contextual adjustments of the transition boundary follow a rotational movement in the F(onset) – F(endpoint) plane around a central point corresponding to the flat transition in the neutral context, direction of the boundary line depending on the perceived identity of the following vowel (Serniclaes & Carré, 2002).

Though place perception is strongly depended on the phonetic context, the fact that place boundaries correspond to flat F2 and F3 transitions in the neutral vowel context points a relationship with natural psychoacoustic settings. Both infants below 9 months of age and adults are much less sensitive to a difference between two different falling or rising frequency transition than to a difference between a falling and a rising transition (Aslin, 1989). This suggests that flat transition boundaries correspond to basic psychoacoustic limitations. Although psychocaoustic in nature, the sensitivity to changes in the direction of frequency transitions might be adapted

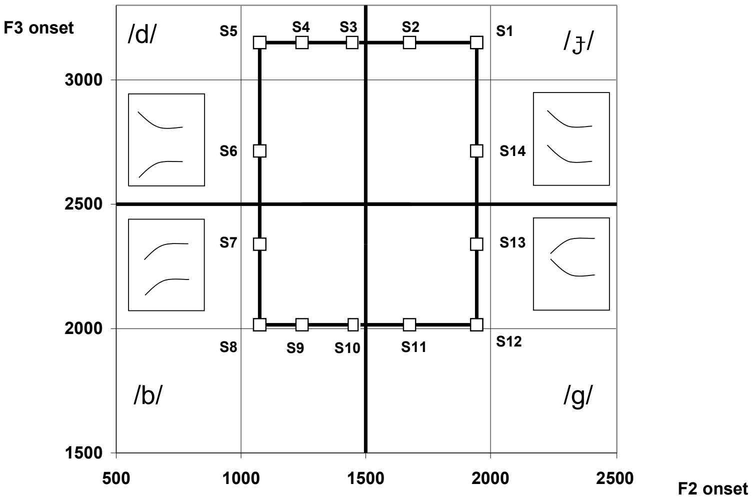


Figure 1a.

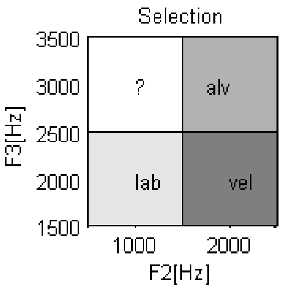


Figure 1b

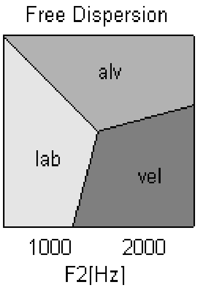


Figure 1c

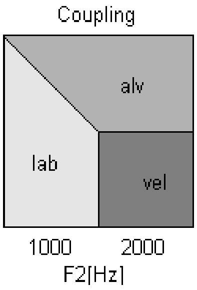


Figure 1d

for perceiving place of articulation during language development. Alternatively, it is also possible that sensitivity to differences in the frequency transition direction were integrated into a speech specific module during phylogenetic evolution (Liberman, 1998). What is clear is that flat transition boundaries are not directly usable for perceiving consonant place of aticulation in all the languages.

The flat F2-F3 transitions might be straightforwardly used for perceiving place of articulation contrasts in a four-category language such as Hungarian. Each category would then occupy a single quadrant of the formant transition onset F2-F3 acoustic space (Figure 1a). In a three category languages such as French, the perceptual boundaries afforded by flat F2 and F3 transitions might also be used as such for perceiving place contrasts, but one of the two categories adjacent to the absent one would then occupy two different regions of the acoustic space (see Figure 1b). This is what an entirely selective model of predispositions would predict. However, as explained above, a more realistic model of speech development with predispostions is that there are also couplings between predispositions. In the present case, coupling between the predispositions for perceiving the F2 and F3 transitions might give rise to a new boundary that would be settled in the middle of the “void” region as this has the advantage of equally sharing this region between the two adjacent categories (i.e. /b/ and /d/). As to the boundaries between the contrasts that remain present in three-category languages, i.e. the b/g and d/g contrasts, two different possibilities remain. Either the perception of these contrasts also move as a side effect of change in the b/d boundary (see Figure 1c) or they stick to the natural boundaries, i.e. the flat F2 transition for b/g and the flat F3 transition for d/g (see Figure 1d). While these two possibilities are compatible with a predisposition model of perceptual development, a general boundary shift of the boundaries is also compatible with a distance optimisation view. Under this view, categories would tend to divide the space into three equal regions and the boundaries would be shifted in consequence (Figure 1c).

The present study

Cross-linguistic trends in the perception of place of articulation

We recently collected perceptual evidence in support of combinations between phonetic features for place of articulation (in French: Serniclaes, Bogliotti, & Carré, 2003; in Hungarian: Geng, Mády, Bogliotti, Messaoud-Galusi, Medina, & Serniclaes, 2005). While phonological boundaries did not always correspond to those included in the phonetic predispositions, thereby confirming that simple selectionnist approaches cannot account of perceptual development on its own, nevertheless all the phonological boundaries were somehow related to the phonetic ones. This lends support to the hypothesis that phonological boundaries which are seemingly unrelated to the phonetic predispositions arise from ‘couplings’ (i.e. interactive combinations) between predispositions (Serniclaes, 2000). A further argument in support of couplings is that similar place of articulation boundaries were found for the distinctions that are shared in common by French and Hungarian, despite the fact that Hungarian uses four place categories whereas French only uses three place categories. However, no direct comparisons between the French and Hungarian boundaries were performed in our previous reports.

Here we present some new evidence in support of coupling between predispositions in the perception of phonological features based on cross-linguistic comparisons. A direct comparison between the labeling responses of either French or Hungarian listeners to the same stimuli was used to test the similarities and differences between these two languages. We expected that both languages would display the same perceptual boundaries for distinctions they share in common in the F2-F3 transition onset space. Further, we wanted to confirm that these boundaries correspond to natural boundaries or to some coupling between natural boundaries.

.2 Method

.3

Participants

Participants for the Hungarian subset were (a) participants of an undergraduate linguistics course or (b) volunteers contacted via a mailing list. Apart from their first language, all of them were familiar with at least one of the languages German, French or English. 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 French dataset was similar in age structure: Subjects' age ranging between 17 and 59 years. Likewise, there were no known auditory problems.

Stimuli

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 had a duration 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. Successive stimuli were 1 Bark apart on both continua. The present paper only deals with the data for the first, the "phonetic" continuum. The same amount of stimuli was generated with the same basic data but an additional, constant burst-like signal portion.

Procedure

Both continua were presented to each of the participants. The continua with and without burst were presented in alternating order resulting in a between-subject factor (order of presentation) which was used for control purposes. Hungarian participants were told that they would hear one of the four sounds "b","d","gy" or "g" and were instructed to report which of the four sounds they had heard. They were told that the sounds not necessarily were presented with equal frequency, and to judge each sound separately. For the French participants, the procedure was alike except that for them only three response alternatives were available.

Statistical models

The data were fitted by Nonlinear Regression with a hierachical model, already tested on data pertaining to voicing perception in French (Serniclaes, in prep.), in which the effect of the effect of F2 was nested in F3 and the latter was nested in the effect of the residual (fixed) cues (Equation 1). The effect of each cue was modelled with Logistic regression (LR: Equation 2a, 2b, 2c). Bark transforms of F2 and F3 values were used. The effect of the burst was included as interaction for both F2 and F3 effects (SB parameter). The effect of language was included as a bias for the residual cues effect (SL parameter).

Equ.1. Labeling response = LR (residual cues, LR(F3, LR(F2)))

Equ.2a.

$$\mathrm{LR}\,(F2) = 1/(1+\mathrm{EXP}-(S2*(-C2 + F2 + SB*F2*\mathrm{Burst})))$$

Equ.2b.

$$\mathrm{LR}\,(F3) = 1/(1+\mathrm{EXP}-(S3*(-C3*\mathrm{LR}(F2)+F3 + SB*F3*\mathrm{Burst})))$$

Equ.2c.

$$\mathrm{LR}\,(\text{Residual cues}) = 1/(1+\mathrm{EXP}-(S4*(-C4*\mathrm{LR}(F3)+ SL*\mathrm{Language})))$$

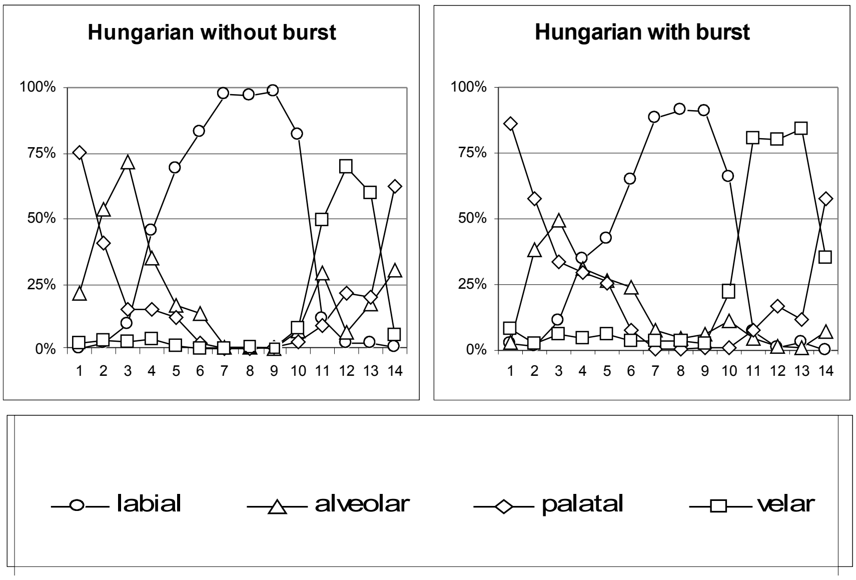
The NLR model includes 8 parameters. The following 8 parameter Logistic Regression model was used as a baseline for assessing the performances of the NLR model:

Equ.3.

$$\text{Labeling response} = 1/(1+\mathrm{EXP}-(C+ S2*F2 + S3*F3 + S4*\mathrm{Burst} + S5*\mathrm{Language} + S5*F2*\mathrm{Burst} + S6\ F3*\mathrm{Burst} + S7*F2*F3))).$$

.4 Results

The labeling curves for the stimuli with or without burst in French and Hungarian are presented in Figure 2. Although there are obvious differences between the labeling curves for the stimuli with vs. without burst, the location of the boundaries (i.e. the stimuli collecting an equal number of responses for two adjacent categories) are only marginally affected. In French, there are three boundaries corresponding to (from right to left in Figure 2) the alveolar/labial, the labial/velar and the velar/alveolar distinctions. Interestingly, there is a secondary peak of velar responses around the alveolar/labial boundary, mainly for the stimuli without burst. In Hungarian, there are four boundaries corresponding to (from right to left in Figure 2) the palatal/alveolar; alveolar/labial, the labial/velar and the velar/alveolar distinctions. The distinctions between palatals, alveolars and velars are not very clearcut. However, the Hungarian palatal and alveolar functions, when taken together, correspond fairly well to the French alveolar function. For these reasons, we decided to merge the palatal and alveolar responses into a single alveolar-palatal category in the following analyses.



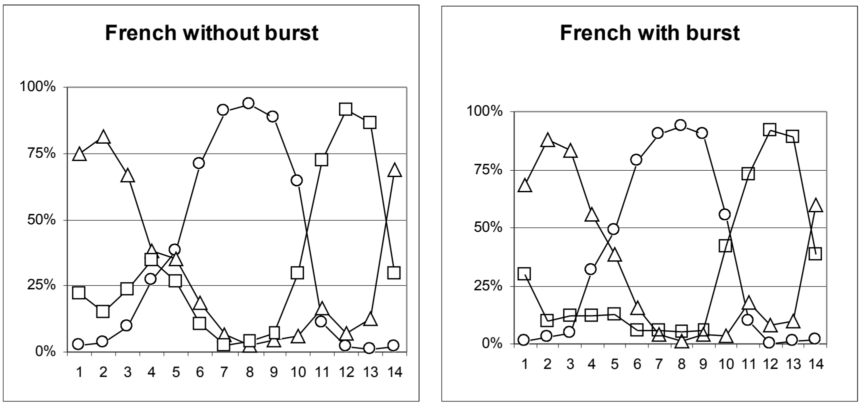


Figure 2. Labeling curves for the stimuli with or without burst in Hungarian and French.

The data were fitted with Nonlinear Regressions (NLR) run on a hierarchical model (Equ.1). A separate NLR was run for each contrast, i.e. labial/velar, velar/alveo-palatal and alveo-palatal/labial. The performances of the NLR models (Equ.1) were compared to those of the simple Logistic Regressions (LR, Equ.2) with the same number of parameters. The percentage of explained variance amounted to 63.5 % with NLR vs. 61.8 % with LR for the labial/velar contrast, to 40 % with NLR vs. 38 % with LR for the velar/alveo-palatal contrast, to 63.9 % with NLR vs. 60.4 % with LR for the alveo-palatal/labial contrast. The NLR models fitted the data better than simple Logistic Regressions although differences are fairly small.

NLR was used for testing the effect of language on place identification as well as specific hypotheses on the location of the place boundaries in the F2-

F3 onset transition space. As explained in the Introduction, we expected that the place contrasts which are common to both languages display the same perceptual boundaries. We also wanted to confirm previous analyses conducted separately on the data collected for each language and which showed that the place boundaries correspond to natural boundaries or to some coupling between natural boundaries. Specifically, we expected that the labial/velar boundary would correspond to a flat F2 transition, that the velar/alveolar-palatal boundary would correspond to a flat F3 transition and that the alveolar-palatal/labial boundary would correspond to a tradeoff between a rising F2 transition and a falling F3 transition.

The boundary estimations are given in Table 1. For the labial/ velar contrast, the model included a F2 component nested in a residual cues component ([LR(Residual Cues, LR(F2))] model). The effect of F3 was not significant. The Burst and Language biases were not significant. The effects of F2, the Burst x F2, Language x F2 (all p<.001) and Burst x language x F2 (p<.05) interactions were significant. The labial/ velar boundary correponds to an almost flat F2 transition in both languages, both for the stimuli with and without bursts (Table 1). For the velar/alveolar-palatal contrast, the model included a F3 component nested in a residual cues component, ([LR(Residual Cues, LR(F3))] model). The effect of F2 was not significant. The effects of F3, Burst x F3 interaction, Burst bias, Language bias and Burst x Language bias were significant (all p<.001, except Burst x F3 interaction, p<.01). The velar/alveolar-palatal boundary correponds to and almost flat F3 transition for the stimuli without burst and a slightly falling F3 transition for the stimuli with burst (Table 1). For the alveolar-palatal/ labial contrast, the model included a F2 component nested in a F3 component, ([LR(F3, LR(F2))] model). The effects of the Residual Cues as well as burst and language biases were not significant. The effects of F2 and F3, as well as the Burst x F2 and Burst x F3 interaction were significant (all p<.001). The tradeoffs between F2 and F3 transition onset values are presented in Table 3, per language and burst condition. A rising F2 transition is compensated by a falling F3 transition in both languages and for both burst conditions, indicating that the the alveolar-palatal/ labial boundary corresponds to a trade-off between a rising F2 and a falling F3 transition.

		Without burst		With burst	
Contrast		Hungarian	French	Hungarian	French
labial/ velar F2 onset	Observed NLR CI limits	49.3 49.4 49.0-50.0	49.1 48.8 48.5-49.4	49.0 49.1 48.7- 49.7	49.1 48.8 48.4-49.4
velar/ alveolar- palatal F3 onset	Observed NLR CI limits	52.2 52.2 51.7-52.5	52.6 52.6 52.5-52.8	52.7 52.7 52.6-52.8	52.8 52.9 52.8-52.9
alveolar- palatal/ labial F3 extent/ F2 extent	Observed NLR CI limits	1.1 1.0 0.5-1.4	0.8 1.0 0.5-1.4	0.6 0.9 0.5-1.1	0.8 0.9 0.5-1.1

Table 1. Values of formant transitions at the perceptual boundary for the three place contrasts, for each burst condition (without vs. with), and for each language (Hungarian vs. French). Each data cell gives the observed values, NLR estimations and 95% CI limits. For the labial/ velar contrast, boundary values are indexed by the onset frequency of the F2 transition in Bark. The boundary values are fairly close to the flat boundary transition (49.0 Bark, 1500 Hz F2) in both languages, both for the stimuli with and without bursts. For the velar/ alveolar-palatal contrast, boundary values are indexed by the onset frequency of the F3 transition in Bark. The boundary values are close to the flat boundary transition (52.4 Bark, 2500 Hz F3) in both languages, both for the stimuli with and without bursts. For the alveolar-palatal/ labial contrast, boundary values are indexed by ratio of the extent of the F3 vs. the extent F2 transition in Bark. The F3/F2 transition extent ratio is fairly close to 1, except for the Hungarian data with stimuli with burst, and never significantly different from 1. This suggests that the alveolar-palatal/ labial boundary corresponds to a trade-off between a rising F2 and a falling F3 transition in both languages.

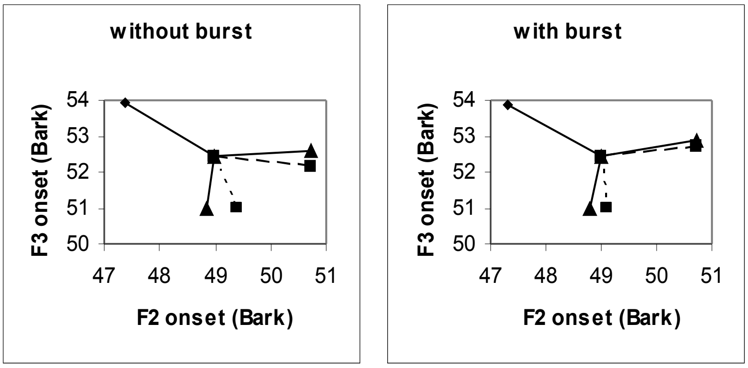


Figure 3. NLR estimations of the territorial maps per burst condition with both French boundaries (plain lines) and Hungarian boundaries (dotted lines). The labial/alveolar-palatal boundaries of the two languages overlap.

Territorial maps of the place categories in the F2-F3 onset frequencies are presented in Figure 3. These maps were obtained by calculating the boundaries between categories from the outputs the Non Linear Regressions (Equ.1). For both the stimuli with and without burst, the velar region corresponds to the lower right quadrant with boundaries corresponding to fairly flat F2 and F3 transitions (see Table 1 for details). The labial/ alveolar-palatal boundary corresponds to the tradeoff between a rising F2 and a falling F3 transition. There is some trendency for the velar region to be narrower in Hungarian but differences between languages are fairly small. **CHRISTIAN : DO YOU WANT TO ADD SOMETHING HERE OR BETTER IN THE DISCUSSION ? YOU SAID: Maybe some words on the theoretical status of the velar boundary?

.5 Discussion

Stability of place boundaries across languages.

The present results show that transitional features are used in much the same way as in both Hungarian and French. Strikingly, the contrasts which are common in both languages use almost the same perceptual boundaries, especially for the stimuli with burst. Further, the boundaries common to both languages are not selected at random but correspond to qualitative changes in the direction of frequency transitions.

On the enrootment of place perception in natural boundaries.

The place boundaries evidenced in the present study are clearly related to natural settings. The labial/velar distinction is based on rising vs. falling direction of the F2 transition which obviously corresponds to a natural boundary. Similarly, the alveo-palatal/velar distinction is based on the direction of the F3 transition. These represent two clear examples of direct implementation of natural boundaries in the phonological framework.

Both infants and adults display a natural sensitivity for perceiving differences between rising and falling transitions in non-speech sounds (Aslin, 1989). This suggests that flat transition boundaries evidenced in the present study correspond to basic psychoacoustic limitations in the processing of frequency transitions. The relevance of these boundaries for speech perception in two languages with quite different place of articulation settings, i.e. French and Hungarian, clearly demonstrate the role of predispositions in speech development. However, it is also clear that predispositions for speech perception are not always directly suited for phonological purposes. Evidently, the distinction between bilabial and alveolar-palatal stops does not directly depend on either the sole F2 or the F3 transition. However, the bilabial / alveolar-palatal contrast calls on a tradeoff between the two transitions, a rising F2 being compensated by a falling F3 to yield a globally flat F2-F3 compound. This tradeoff does not depend on the specificities of place production in the two different languages. Rather, it corresponds to still another qualitative difference between speech sounds: the difference between globally rising vs. globally falling F2 and F3 transitions. Though the difference in the global direction of the transitions is more complex than those between the individual directions of F2 and F3 transitions, they all correspond to qualitative changes in transition direction. Such qualitative changes are higly specific and the present results suggest that they impose strong constraints on the location of place boundaries.

Finally, the present data give some arguments against dispersion models. Both the
Adaptative Dispersion Theory (ADT: Lindblom, 1986) and the D** F** Theory (
DFT: for vowels, see Schwartz, **; for a tentative expansion of DFT to consonants, see Abry, 2003) claim that language tend to optimally divide the acoustic space between phonological categories. Altough these theories were up to now only tested on productive grounds, no doubt they would predict an equal sharing of the acoustic space between peceptual categories. But this is not at all what we see in the present data. The velar region is much smaller than the labial and the alveolar-palatal ones. As we have seen, this is due to the lay-out of natural boundaries across the acoustic space: the limits between categorical regions correspond either to flat F2 or F3 transitions or to a trade-off between these two natural boundaries. Rather than being determined by the optimal sharing of the acoustic space, the categorical regions are determined by perceptual contraints.

CHISTIAN: STILL DOUBTFUL ? YOU SAID: Still Not sure about that

Transitions vs. burst as vectors of place perception.

The results show that the perception of the fourfold place of articulation contrasts in Hungarian are partially based on the direction of F2 and F3 formant transitions. However, these cues are clearly not sufficient for supporting the alveolar/ palatal contrast.

If one reflects the results obtained in the light of design issues for segmental speech perception, it is useful to be aware of the taxonomic dichotomy distinguishing between *deleted cue* and *conflicting cue* experiments. In a deleted cue experiment, one (or more) acoustic cues (or whole acoustic features) are removed and not available for the listener (see Smits, 1996). Perception experiments in the deleted cue paradigm measure the *necessity* of the deleted cues and the *sufficiency* of the cues manipulated in the stimulus design. In contrast, conflicting cue experiments manipulate stimuli such that the different manipulated cues point to different categories. This paradigm is more apt for capturing relative contributions of different cues. The experiments reported here might be viewed as deleted cue experiments in the sense that burst information was entirely correlated with transitions in the experimental stimuli: there was no independent burst manipulation and it only served the purpose of calibration of the stimuli.

However, the burst might also be considered as indexing another *feature*, potentially independent from those supported by F2 and F3 transitions. Our results would then show that the transitional features manipulated are *not* sufficient and that other features are necessary for the addition of the palatal to the three principal places of articulation. To reveal the contribution of these burst-related features one has to use stimuli generated by factorial variation of burst and transitions. There have been several attempts in the past for separating the contributions of burst and transitions to place perception in stop consonants and most of these studies point to the functional equivalence of the two cues across phonetic contexts

(Dorman **, 1977; Carlson **., 1972; Schouten & Pols, 1984; but see Pols & Schouten, 1985). However, these results were collected in languages with only three place categories which makes it rather difficult for evidencing autonomous contributions of the two kinds of cues because both contributed to all the three possible contrasts. Things might turn out differently in a four-category language like Hungarian in which the different place contrasts might rely on different cues: as shown in the present results, transitions are sufficient as long as there is no aveolar/palatal contrast present. It is then possible that the perception of this contrast might rely on burst properties which are independent of the onset frequencies of the formant transitions. Future experiments with stimuli generated by factorial variation of burst and transitions should allow to clarify this point.

Here one has to be aware that the phonemic status of the Hungarian palatal has been the topic of a long-standing debate on whether it should be classified as a true stop or an affricate (for a summary on the phonological treatment of this matter see Siptár & Törkenczy, 2000). Our own recordings of the Hungarian palatal (Geng & Mooshammer, 2004) suggested a high degree of acoustic variability in the realization of this sound comprising clear stop and affricate and even fricative-like realizations with signal portions we interpreted as a residual burst. More thorough and detailed spectral analyses of these data would be required before a conclusive categorization of the observed patterns could be possible.

Some implications for phonological systems.

From a systemic point of view, the present study lends further support to the idea that languages do not use all the possible combinations between two (or several) phonetic features. The data suggest that in languages with four places of articulation categories, such as Hungarian, F2 and F3 transitions are not sufficient for separating these categories and a third feature is necessary. The examination of laryngeal timing contrasts leads to similar conclusions. Although there are two different predispositions for perceiving negative and positive VOT, there are only three rather than four reliable VOT categories in the world's languages and a third phonetic feature (manner of vocal fold vibration) is used in language with four homorganic categories (Lisker & Abramson, 1964).

.

CHRISTIAN ?The fact that the category boundaries stay relatively stable between both languages under consideration while adding an additional place of articulation seems to suggest that there is an upper limit of place categories based on place of articulation alone. This correlates well with the fact that for all obstruents, stops, fricatives and affricates the voiceless cognates are more frequent than the voiced ones. We do not want to overemphasize this point, as there are other well-established mechanisms for the same explanandum:

Stops: Maddieson (2003) refers to missing p and missing B phenomena (where is the formulation??)
For fricatives: conflicting aerodynamic demands of sustaining voicing and frication in terms of airflow

In other words, relating the category density limitations for voiced stops as observed in this paper to the structure of sound inventories is problematic as it hardly seems possible to disentangle effects of the different phonetic mechanisms on frequency data as those available in the UPSID database. The results presented here do not render such speculations as prohibitive either: The density effect observed in this study and the aerodynamic considerations described just above could act in a synergistic fashion in the constitution of the emergent assymmtries observed in the obstruent inventories of the world's languages.

.6 Conclusion

The present findings bear several implications for both speech development and phonological systems. Firstly, the fact that, both in French and in Hungarian**,** perceptual boundaries align along natural boundaries for transition perception or along some combination of these natural boundaries gives further support to the coupling model of speech development. Further, as the alveolar- palatal Hungarian boundary is poorly represented in the transition space, at least one additional phonetic feature seems necessary for perceiving the fourfold place distinctions in Hungarian, thereby giving still a further example of coupling between phonetic features in the build-up of phoneme categories. Secondly, the fact that the boundaries are generated by lawful combinations of perceptual predispositions shows that the latter impose strong constraints on phonological development. An important question for future research will be to understand how these constraints might converge with processes based on distance between categories.

References

Abry, C.

2003 [b]-[d]-[g] as a universal triangle as acoustically optimal as [i]-[a]-[u]. . In M.J.Solé, D.Recaesens & J.Romero (Eds.). *Proc. 15th International Congress on Phonetic Sciences*, 727-730.

Aslin, R.N.

1989 Discrimination of frequency transitions by human infants. Journal of the Acoustical Society of America, 86, 582–590.

Carré, R., & Mrayati, M. (1991). Vowel-vowel trajectories and region modeling. Journal of Phonetics, 19, 433-443.

Dorman, M.F., Studdert-Kennedy, M., & Raphaël, L.S. (1977). Stop consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception and Psychophysics, 22,* 109-122.

Eimas, P.D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception and Psychophysics, 16,* 513-521.

Geng, C., Mády, K., Bogliotti, C., Messaoud-Galusi, S., Medina,V., & Serniclaes, W.
2005 Do palatal consonants correspond to the fourth category in the perceptual F2-F3 space? In: V.Hazan & P. Iverson (Eds.). *ISCA Workshop on Plasticity in Speech Perception, London, June 15-17 2005,* 219-222.

Geng & Mooshammer
2004

Guenther, F., & Bohland, J.
200 Learning sound categories: A neural model and supporting experiments. *Acoust. Sci. & Tech,. 23,* 213-220.

Harnad, S.
1987. *Categorical perception: the groundwork of cognition.* New-York: Cambridge University Press.

Hochberg J. (1981). On cognition in perception: perceptual coupling and unconscious inference. *Cognition, 10,* 127-34.

Hoonhorst, I., Colin, C., Deltenre, P., Radeau, M., & Serniclaes, W. (2006). Emergence of a language specific boundary in perilinguistic infants. *Early Language Development and Disordes, Latsis Colloqium of the University of Geneva, Program and Abstracts*, 45 .

Jakobson, R., Fant, G., & Halle, M.
1952 Preliminaries to Speech Analysis. Cambridge: MIT Press.

CHRISTIAN: DELETE THE FOLLOWING ?
Keating, P.
1993 Phonetic representation of palatalization versus fronting. *UCLA Working Papers in Phonetics*, 85:6–21.

Keating, P. & Lahiri, A.
1993 Fronted velars, palatized velars, and palatals. *Phonetica*, 50:73–101.

Keating, P. A.
1988 Palatals as complex segments: X-ray evidence. *UCLA Working Papers in Phonetics*, 69:77–91.

Keating, P. A.
1991 Coronal places of articulation. In: C. Paradis & J.-F. Prunet (eds.). *Phonetics and Phonology 2: The special status of coronals: Internal and external evidence*, 29–48. San Diego, Calif.: Academic Press.

Koffka, K. (1935). *Principles of Gestalt Psychology.* New York: Hartcourt Brace.

Kuhl, P.K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews*, 5, 831-843.

Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T. Kiritani, S. & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, F13-F21.

1988

Lasky, R.E., Syrdal-Lasky, A. & Klein, R.E. (1975). VOT discrimination by four to six and a half month-old infants from Spanish environment. *Journal of Experimental Child Psychology***, 20,** 215-225.

Liberman, A.M. (1998). Three questions for a theory of speech. *The Journal of the Acoustical Society of America, 103*, 3024.

Lindblom, 1986

Liljencrants & Lindblom, 1972

Lisker, L., & Abramson, A.S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word, 20,* 384-422.

Masin, S.C. (1993). Some philosophical observations on perceptual science. In S. Masin (Ed.), *Foundations of perceptual theory* (pp. 43-73). Amsterdam: Elsevier.

Pegg, J.E. & Werker, J.F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America,* 102, 3742-3753.

Serniclaes, W. (1987). *Etude expérimentale de la perception du trait de voisement des occlusives du français.* Unpublished doctoral thesis. Université Libre de Bruxelles.
Serniclaes, W. (2000). La perception de la parole. In P. Escudier & J.L. Schwartz (Ed.), *La parole : Des modèles cognitifs aux machines communicantes* (pp. 159-190). Paris: Hermes, Science Publications.

Serniclaes, W. & Carré, R. (2002). Contextual effects in the perception of place of articulation: a rotational hypothesis. In J.L.H. Hansen & B. Pellom (Eds.). *7th International Conference on Spoken Language Processing*, 1673-1676.

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

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

Schouten, M.E.H. and Pols, L.C.W. (1984).Identification of intervocalic plosive consonants: the importance of plosive bursts vs. vocalic transition. In Proc. 10th. Int. Cong. Phon. Sci., Utrecht 1983, M.P.R. Van den Broecke and A.Cohen eds., Foris publ.: Dordrecht; Vol.IIb, 464-468.

Siptár, P. & Torkenczy, M.
2000 *The phonology of Hungarian*. Oxford: Univ. Press.

Vihman, M.V.
1996 *Phonological development : The origins of language in the child.* Cambridge (MA): Blackwell.

Werker, J.F. & Tees, R.C. (1984a). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, **7**, 49-63.

Werker, J.F. & Tees, R.C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America, **75**,* 1866-1878.