THE PHONETIC BASIS OF PHONOLOGICAL STRUCTURE: THE ROLE OF AERODYNAMIC FACTORS

Maria-Josep Solé
Universitat Autònoma de Barcelona

ABSTRACT

This paper addresses the question of how aerodynamic factors shape phonological structure, specifically feature co-occurrence restrictions. First we revisit the notion that the principles that determine the shape and behavior of phonology -- 'articulatory economy', 'auditory contrast' and 'articulatory-acoustic stability'-- are based on the physical and physiological properties of the speech production and perception systems. Second, we illustrate how aerodynamic factors may determine which features can and are likely to combine into segments. We report a series of experiments where 'voicing' and 'nasalization' were actively manipulated in fricatives and trills -- two segment types which crucially depend on aerodynamic conditions. We analyze the role of aerodynamic and perceptual factors in determining why some feature combinations, *e.g.*, nasal fricatives and nasal trills, fail to occur, and why other combinations, *e.g.* voiced trills, are universally preferred.

1. INTRODUCTION

This paper investigates the role of aerodynamic factors in shaping phonological structure. Specifically, it discusses how aerodynamic factors, in combination with other constraints of production and perception, determine feature co-occurrence restrictions, that is, why certain combinations of features into segments are likely to occur whereas some others are rare or fail to occur.

This study is part of a research program addressed at characterizing the constraints of the speech production and perception mechanisms that shape phonological structure. The role of the physical and auditory constraints of the speech mechanism in shaping phonological patterns has long been recognized in the phonetic literature (Ohala 1974, 1983, 1993; Liljencrants and Lindblom 1972; Lindblom 1983, 1990; Westbury and Keating 1986; Browman and Goldstein 1986, 1990). This approach is inserted in a biological perspective which sees phonological structure as emerging from functional factors, *i.e*., perceptuo-motor capabilities and constraints, and their optimization into efficient and robust signals. In this view, phonological structure rather than being preformed, arbitrary and governed by principles unique to language is viewed as natural, arising from neuro-motoric, biomechanical, aerodynamic, auditory and perceptual factors not specific to language (Lindblom 1992). Such physical constraints give rise to complex phonological structure that meets the general conditions of pronounceability, distinctiveness and robustness (defined in sections 1.1.1 and 1.1.2 below). Thus general physical and perceptual mechanisms operate on and maintain phonological structure, optimizing the relations between form and function.

In the last decade there has been a growth in interest and research effort directed at deriving phonological form from phonetic substance and providing explanations in terms of physical and auditory principles not specific to language. A number of studies have contributed to our understanding of the phonetic basis of pervasive questions in phonology, such as sound inventories (*e.g*., Liljencrants and Lindblom 1972), possible sound segments or feature co-occurrence restrictions (*e.g*., Ohala, Solé and Ying 1998, Ohala 1983), possible sequences of segments or morpheme structure constraints (*e.g*., Kawasaki 1982), neutralization or loss of feature contrasts (e.g., Steriade 1997), allophony or addition of feature contrasts (*e.g*., Sproat and Fujimura 1993), phonologization of features -- *i.e*., redundant or mechanical features becoming primary -- (*e.g*., Ohala 1993), phonological universals (Westbury and Keating 1986, Solé *in press*), and phonological processes, such as assimilation (*e.g*., Ohala 1990, Recasens 1995) and dissimilation (Ohala 1981).

The relevance of the physical and perceptual properties of speech sounds in phonological structure has only recently begun to be acknowledged in phonological theory. A few researchers have tried to incorporate data on speech perception in formal theories (Steriade 1997), or have incorporated formal phonology into speech perception (Lahiri and Marslen-Wilson 1992). Articulatory Phonology (Browman and Goldstein 1986, 1990) has tried to show how phonological organization may arise from constraints imposed by physical systems, providing a physical foundation to featural phonology. Current phonological theories, such as Optimality Theory, recognize the 'groundedness' of phonological constraints.

Whereas the role of articulatory, auditory and perceptual facts for phonological patterns has been studied in some detail, aerodynamic factors have usually been neglected, with the notable exception of work by Ohala (1976, 1983, 1993).

This paper addresses the phonetic principles that govern the combination of features into segments, or feature co-occurrence restrictions, by exploring the physical structure of speech features, singly and in combination, and their perceptual properties. Specifically, we address the role of aerodynamic and perceptual factors in determining why some feature combinations fail to occur, e.g., nasal fricatives, nasal trills, and why other combinations are universally preferred, e.g., voiced trills.

1.1. The phonetic basis of feature co-occurrence restrictions

Features and combinations of features into segments reflect the linguistic categorization of speech events. Some speech events cannot or are not likely to co-occur. Consequently not all feature values can be freely combined or are as likely to co-occur. Preferred feature combinations are those that meet the conditions of ‘articulatory economy’, ‘auditory distinctiveness’ and ‘stability’. Such conditions (1) reflect and optimize the constraints of the speech production and perceiving mechanisms, and (2) interact to design an efficient communication system.

1.1.1. The notions of 'articulatory economy' and 'auditory distinctiveness'.

The principles that determine the combination of features into possible speech segments and their likelihood have been recognized as deriving from the *physical* (including articulatory and aerodynamic) and *auditory* properties of speech features. Combinations of features obeying constraints of the speech production mechanism (*i.e.*, not involving extra adjustments and/or increased articulatory cost) and having efficient acoustic consequences (*i.e.*, resulting in auditorily salient and distinct signals), or those showing a trade-off between competing demands of perception and articulation will tend to be favored and are more likely to be used cross-linguistically. Combinations of features that fail to meet the physical and auditory demands tend to be disfavored.

The constraints imposed by the speech production and perception systems are at the basis of the notions of 'articulatory economy' (sometimes referred to as 'pronounceability' or 'connectedness') and 'auditory distinctiveness' (or 'contrastivity'), which define the combination of speech features into segments, and their likelihood. For example, the fact that voiced obstruents (*i.e.*, the combination of [+obstruent] and [+voice]) are less common in sound inventories and have a lower frequency in running texts and in lexical items than voiceless obstruents, has been accounted for in *aerodynamic* terms: the rise in pressure in the oral cavity for the obstruent impairs the pressure drop across the glottis required for voicing (Ohala 1983, Westbury and Keating 1986). Additional articulatory manoeuvers in cavity volume, vocal tract compliance, etc., are required to sustain voicing in obstruents, making these combinations biomechanically more costly, and thus less favored.

Poor *auditory* result (as measured from confusion coefficients and evidenced by dephonologization or neutralization) can be adduced to account for the relatively low incidence of voiceless nasals, *i.e.*, [+nasal] combining with [-voice], as the low frequency amplitude modulation for nasals is impaired by voicelessness. Conversely, some combination of features work in synergy and enhance the acoustic-auditory image, *e.g.*, the combined action of [+back] and [+rounded] heighten F2 lowering (Perkell, Matthies and Jordan 1993), which is in line with the universal preference to round back vowels.

Thus, combinations of features which obey the constraints of the speech production system and do not require extra articulatory cost or additional manoeuvers, and those that result in auditorily salient signals tend to be favoured, reflecting the phonetic groundedness of the principles governing the combination of features.

1.1.2. The notion of articulatory-acoustic stability.

This study rests on the assumption that, in order to account for preferred combinations of features into segments, it is necessary to characterize (1) a set of parameters -- physiological, aerodynamic and acoustic/auditory -- and their range of variation, singly and in combination, and (2) a set of categorial values along these parameters which remain relatively *stable* with variations in the other parameters, that is, regions which allow articulatory and/or aerodynamic variation without causing major changes in the acoustic-auditory domain. These categories or 'optimal settings' across the different parameters are the more likely combinations of features into segments (Stevens 1972, 1989, Lindblom 1986, Ohala 1983). Some of the gradient physiological or aerodynamic variation in combining feature values will tend to result in categorial changes along the acoustic parameter due to the quantal nature of speech (Stevens 1972, 1989), and will tend to be disfavored. Thus, the range of allowable variation in production, along a single or different parameters, while maintaining auditory integrity underlies the notions of 'acoustic stability' and 'robustness' (as well as the notion of 'categorical perception') utilized to define features and co-occurrence of features.

The notion of articulatory-acoustic stability bears on the notion of binary feature values. Phonologically, as a reflex of the perceptual system, the speech production mechanism is required to provide *binary* distinctions between, *e.g.*, nasal and oral, or voiced and voiceless segments. Physiologically and aerodynamically, however, the speech mechanism functions in a *gradual* manner such that, when moving towards or away from a target, it will necesarily go through intermediate positions or values. The values above (and below) the perceptual threshold for the feature, say [nasal] or [voice], will necessarily exhibit some physiological and aerodynamic variation without causing perceptual changes (*e.g.*, for oral consonants the velum continues rising after complete velopharyngeal closure has been attained (Moll and Daniloff 1971, Figure 2)). The range of allowable articulatory/aerodynamic variation within which the percept of the feature is not affected will define the stability of the articulatory-acoustic correlation. This stable range will vary, *i.e.*, will be expanded, reduced or shifted, with co-occurring features. Feature combinations which result in narrowly constrained articulatory or aerodynamic conditions are unstable articulatorily and may be easily thrown off. Such articulatorily unstable combinations may easily change into a different percept, and will tend not to be used (thus paralleling 'natural selection').

A number of factors bear on the articulatory stability of features: coarticulation, position, prosodic factors, temporal constraints, relative timing of features, etc. For reasons of focus, we will concentrate on the articulatory and acoustic stability of *simultaneous* feature values.

1.1.3. In summary, the conditions of gestural economy, auditory distinctiveness and articulatory-acoustic stability are important building blocks on which phonological structure rests. Precise and explicit definitions of these concepts -- including quantification of the physical and perceptual phenomena involved -- and verifiable empirical support are required if these principles are to be explanations for phonological organization. In order to contribute to this goal, this paper attempts to quantify the articulatory-acoustic stability of speech features, specifically trilling and frication, singly and in combination with other features, and to characterize their saliency and auditory distinctiveness. An advantage of this approach is the independent motivation of the physical principles and mechanisms that interact to adequately shape phonological structure, allowing a natural and unified account.

In order to characterize the phonetic principles that govern the possibilities of combination of features, we investigated (1) the aerodynamic conditions required for fricatives and trills -- two segment types which crucially depend on aerodynamic factors--, (2) the range of allowable variation in oral pressure before their spectral identity is compromised, and (3) how the aerodynamic conditions are affected by co-occurring features, specifically [nasal] and [voice]. The results throw light on why some feature combinations fail to occur, e.g., nasal fricatives, nasal trills, and why some other combinations are universally preferred, e.g., voiced trills.

2. METHOD

In order to determine how the aerodynamic conditions required for trilling and frication are affected by the features [voice] and [nasal], we analyzed the aerodynamic and acoustic effects on trills and fricatives of actively varying (i) voicing, voiced and voiceless segments were investigated, and (ii) nasality, by venting oral pressure with a pseudo-pharyngeal valve, *i.e.*, simulating the effects of co-occurring nasalization.

Intraoral pressure (Po) and airflow were recorded in two trained phoneticians producing a variety of voiced and voiceless fricatives and apical trills in intervocalic position and as artificially prolonged steady states. Po was intermittently vented with catheters of varying cross-sectional areas (7.9, 17.8, 31.7, and 49.5 mm^2), all 25 cm long, inserted into the mouth via the buccal sulcus and the gap behind the back molars (as described in Solé 1998). The catheters venting the oral pressure were intended to simulate the effects of varying degrees of velopharyngeal opening, *i.e.*, nasalization. White noise was placed on the speakers through earphones at a loudness that was sufficient to mask the high frequency noise of the fricatives spoken. The articulatory and acoustic effects associated to variations in Po were analyzed.

3. AERODYNAMIC CONSTRAINTS ON THE CO-OCCURRENCE OF FRICATION AND NASALITY

The first question addressed was whether fricatives can be simultaneously nasalized, *i.e.*, if the features [fricative] and [nasal] can co-occur in a segment (Ohala, Solé and Ying 1998, Ohala 1988). Some authors claim that fricatives with the constriction location further forward than the velo-pharyngeal port (*i.e.*, from labial to velar fricatives) cannot be simultaneously nasalized (Cohn 1993, Ohala and Ohala 1993). They argue that a lowered velum would allow air to flow out through the nose thus reducing the pressure build up in the oral cavity required to generate audible friction at the oral constriction.

Other authors claim that nasalized fricatives are rare but exist in UMbundu (Schadeberg 1982), Coatzospan Mixtec (Gerfen 1996), and Waffa (Stringer and Hotz 1973). None of these authors provide instrumental evidence showing simultaneous frication and nasalization to support their claim, except Gerfen. His data, however, do not provide unequivocal evidence for simultaneous frication and nasalization (as argued in Ohala 1998).

Formal phonology is of no help in this issue. Mainstream phonological accounts attribute the rarity of nasal fricatives to arbitrary 'antagonistic constraints':

' (...) nasalized liquids, glides and fricatives occurring more rarely. The rarity of such segments can be attributed to an antagonistic constraint NAS/CONT: A nasal must not be continuant' (Pulleyblank 1997: 76).

Rather than attributing the absence or rarity of nasal fricatives to arbitrary 'antagonistic constraints' we attempted to settle the issue empirically by analysing the aerodynamic and acoustic effects on fricatives of venting Po with a pseudo-velopharyngeal valve.

The catheters venting the Po were supposed to simulate the effects of varying degrees of velopharyngeal opening. Amount of nasal leakage, however, is not only determined by the area of the velopharyngeal opening but by the total impedance or resistance to the airflow. Impedance is a function of the area of the aperture, the properties of the channel through which the air passes (length, compliance, surface) and the rate of flow; the smaller the aperture, the longer the channel and the larger the amount of flow, the higher the impedance. We determined the impedance of the catheters at the flowrates used in fricatives, and the impedance at the oral constriction during the production of the various fricatives. These are shown in Figure 1.

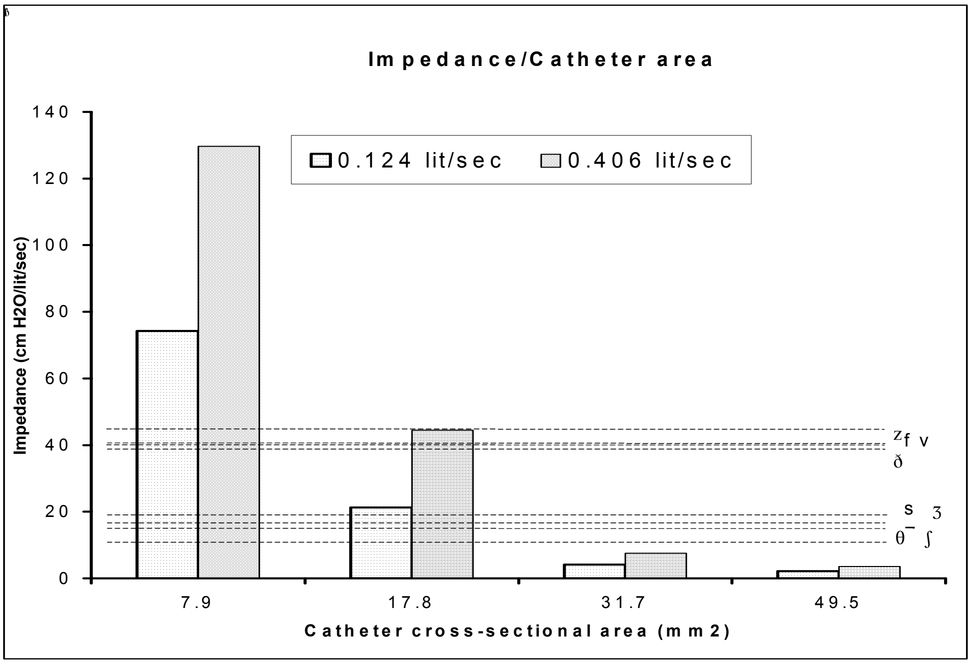

Figure 1. Impedance (ordinate) for catheters of varying cross-sectional area (abscissa) at the flowrates used in fricatives (in lit/sec). Values of measured vocal tract impedance for the various fricatives have been overlayed on the graph (Speaker MJ).

In producing an oral fricative, there can be some opening of the velic valve, but the impedance of this valve has to be relatively high vis-à-vis that of the oral constriction so that the air will escape through the aperture with lower impedance and create frication at the oral constriction. The results indeed showed that venting the Po with catheters with a ***higher*** impedance than that at the oral constriction (catheter area 7.9 mm^2, Fig. 1) produced no detectable alteration of the quality of fricatives (just a slight attenuation of the fricative noise).

Catheters with values for impedance ***similar*** to or ***lower*** than impedance at the oral constriction (catheter areas ≥ 17.8 mm^2, see Figure 1) caused a major reduction in the pressure drop across the oral constriction, and had noticeable auditory effects on fricatives. With a vent of 17.8mm^2 -- with similar impedance to that at the vocal tract-- fricatives lost much of their high-frequency aperiodic energy (*e.g.*, the spectral peak at 6kHz for [s] disappeared and the energy level dropped 20 dB). Sibilant fricatives sounded non-sibilant, in accordance with Behrens and Blumstein (1988) who found an increase in perception of non-sibilant [f, θ] when noise amplitude in sibilant and non-sibilant fricatives [s, ʃ, f, θ] was reduced. Voiced fricatives became frictionless continuants, with increased energy of voicing, (*i.e.*, a lower C/V energy ratio). With catheters with lower impedances (catheter areas ≥ 31.7 mm^2 in Figure1), the airflow exited mainly through the pseudo-velopharyngeal aperture impairing the generation of audible turbulence at the oral constriction. Voiced fricatives were more seriously affected, becoming vowel-like. In voiceless fricatives the intensity of high-frequency noise was substantially reduced. The perceptual validation of the results is currently under way.

The results show that if impedance at the velopharyngeal port is lower than that at the oral constriction the air will escape through the nose (*i.e.*, the fricative will be nasalized), thus reducing the required pressure drop across the oral constriction to generate frication (*i.e.*, frication will be impaired) [1]. Velic openings which do not impair frication (< 17.8 mm^2) would be insufficient to create the percept of nasalization in the fricative or even adjacent vowels. A greater coupling between the oral and the nasal cavity is required for vowels and sonorants to be perceived as 'nasalized'. Maeda (1993) suggests that a robust percept of nasalization requires an opening of 40mm^2 or more for vowels. Consequently, if impedance at the velopharyngeal port is high enough not to affect the fricative quality, the fricative will not sound nasalized.

In summary, to the extent that a fricative is a good fricative perceptually, it cannot be nasalized (without added biomechanical cost, *e.g.*, increased subglottal pressure). Thus, the features frication and nasalization bleed each other aerodynamically and do not combine into a sufficiently discriminable percept. In other words, along the independent physical parameters of frication and nasalization there are categorial values which show stable perceptual properties, *i.e.*, a certain range within the continuum where a reliable identification of frication (or nasalization), say 80%, can be obtained. The two ranges of reliability for frication and nasality, however, do not overlap, *i.e.*, there is not a range of values for both frication and nasalization where you may get 80% identification for both features.

This illustrates how aerodynamic and perceptual constraints account for the lack of nasalized fricatives. Thus, the putative 'nasalized fricatives' are better described as nasalized approximants.

4. CO-OCCURRENCE RESTRICTIONS OF TRILLS

The second question concerns the co-occurrence restrictions of tongue-tip trills and how they can be made to follow from their aerodynamic and perceptual properties.

4.1 Universal co-occurrence of trilling and voice.

Solé (1998) addressed the preference for voiced over voiceless tongue-tip trills -- i.e., of the feature [trill] combining almost exclusively with [+voice]. The statistical preference for voiced over voiceless trills is evidenced in phonological inventories and diachronic variation. Trills are mostly voiced in the languages of the world (98.5%, Maddieson 1984). Voiceless trills tend to disappear historically, merging with /r/s or fricatives, mostly /h/. An example is Proto-Tai *hr-, a voiceless tongue tip trill, which developed into r- (or its equivalent, *i.e.*, /l/, /ð/, /ɣ/) in Northern dialects, and into /h/ in Central and South Western dialects (Li 1977: 142, 148 ff). Another example is Old English /hr/ in *hringan*, *hrēod*, which became /r/ in Middle English, *ring(en)*, *reed*.

Maddieson (1984: 78) suggests that 'perhaps there is some factor in the aerodynamic conditions required for trilling which leads to preference of voicing because of the associated reduction in airflow'. We set out to test this hypothesis by investigating the aerodynamic conditions required for tongue-tip trilling and their range of variation, and how these conditions were affected by co-occurring [voice].

4.1.1. Aerodynamic properties of [trill] [±voice].

Solé, Ohala and Ying (1998), in the experiment described in section 2 where voicing was varied in intervocalic and steady state trills, found that the higher rate of flow through the open glottis for voiceless trills, vis-à-vis voiced trills was responsible for the following differences: (a) Voiceless trills show a higher Po (as illustrated in Fig. 2) and a larger rate of flow through the oral constriction than voiced trills, which results in a higher particle velocity and the creation of turbulence or friction noise across the lingual constriction. (b) A larger open to closed period ratio in tongue tip vibration was found for voiceless as opposed to voiced trills (1.96 vs 1.29). Fig. 2 illustrates the shorter closure period (indicated by a rising Po) for voiceless than for voiced trills whereas the opening period is comparable in both. The proportionally longer open period in voiceless trills results in longer periods of released (turbulent) energy, vis-à-vis voiced. (c) Voiceless trills tend to show frication during the closure (as shown in the waveform in Fig. 2), reflecting failure to achieve full palato-lingual closure.

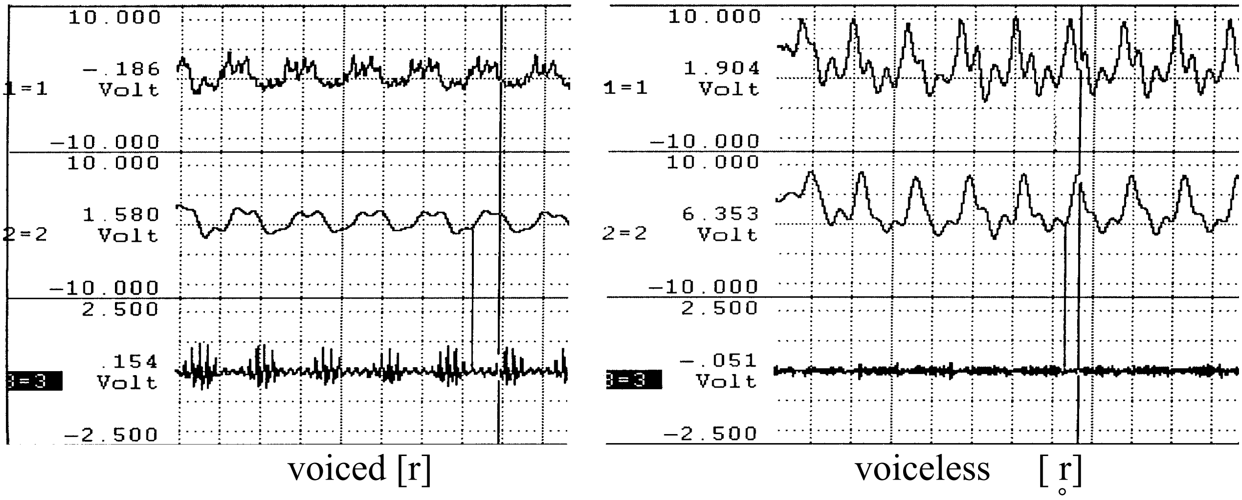

FIGURE 2. Unfiltered Po (channel 1), low-pass filtered Po (channel 2) and audio signal (channel 3), in volts, for sustained voiced and voiceless trills. Channels 1 and 2 show a higher peak Po for voiceless than for voiced trills. The tongue tip contact period, showing a rise in Po, is indicated between lines in the filtered Po trace. Voiceless trills exhibit a shorter closure period than voiced trills.

It was found that voiced trills involve very precise aerodynamic conditions in order to sustain trilling and voicing. Po needs to be high enough to produce tongue-tip vibration (≥ 4cmH20) and low enough not to impair the transglottal flow required for voicing (ΔP≥ 2-3cmH20).

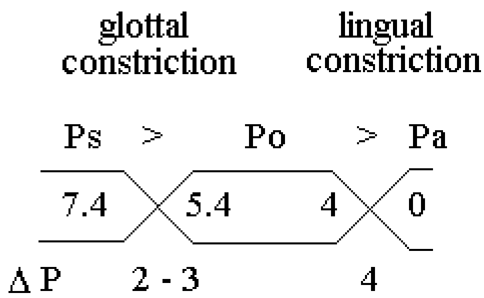

FIGURE 3. Estimated range of Po for voiced trills for speaker MJ.

According to our estimates, for subject MJ oral pressure could vary between a very narrow range of 5.4 – 4 cmH2O in order to sustain voicing and trilling (Ps= 7.4cmH2O; Ps-$\Delta P_{transglottal}$= 5.4cmH2O; $\Delta P_{translingual}$= 4cmH2O), as schematically shown in Fig. 3. The estimation of the range of Po variation for speaker JJ was between 11-8cmH2O. A higher Po would impair transglottal flow and lead to devoicing, a lower Po would endanger translingual flow and lead to cessation of trilling. Trill devoicing and non-trilled variants are found allophonically in a number a languages. Thus, the Po range for voiced trills is very narrow and unforgiving, and small pressure variations may endanger the spectral integrity of the segment. Such severe aerodynamic constraints are not present in voiceless trills.

4.1.2. Variations in oral pressure in voiced and voiceless trills.

In order to determine the range of allowable variation in intraoral pressure in the production of trills, the backpressure during trills was intermittently bled with catheters of different cross-sectional area. Fig. 4 presents the reduction in oral pressure associated with venting the backpressure with catheters of different areas. When the Po dropped below a certain threshold, trills were extinguished resulting in a fricative (a non-sibilant voiced alveolar fricative in the case of voiced trills, and a [h] sound for voiceless trills) or an approximant with catheters of larger areas (≥ 31.7 mm2). The threshold for trilling was determined empirically by measuring the pressure values at which tongue-tip vibration *ceased* on vent. Thus, for example, for speaker MJ a vent of 7.9mm^2 reduced the Po in voiced trills below 4cmH2O and extinguished sustained voiced trilling. The same vent aperture did not extinguish voiced trilling for speaker JJ who exhibited inherently higher pressure values. It is not possible to report absolute pressure values at which trilling was impaired, since depending on initial conditions (articulator tension, mass, cavity volume, articulatory position, compliance, etc.) and speaker, the minimum oral pressure required for tongue-tip vibration varied. It was found that a pressure drop of 2.5-3.5 cmH20 impaired sustained trilling in voiced segments. A larger pressure drop, 5 cmH20, was needed to impair voiceless trills. Thus, voiceless trills are more resistant to variations in Po due to (a) a higher Po which allows a larger reduction in pressure before reaching the minimum pressure drop across the lingual constriction required for trilling, (b) direct access to Ps to replace vented airflow.

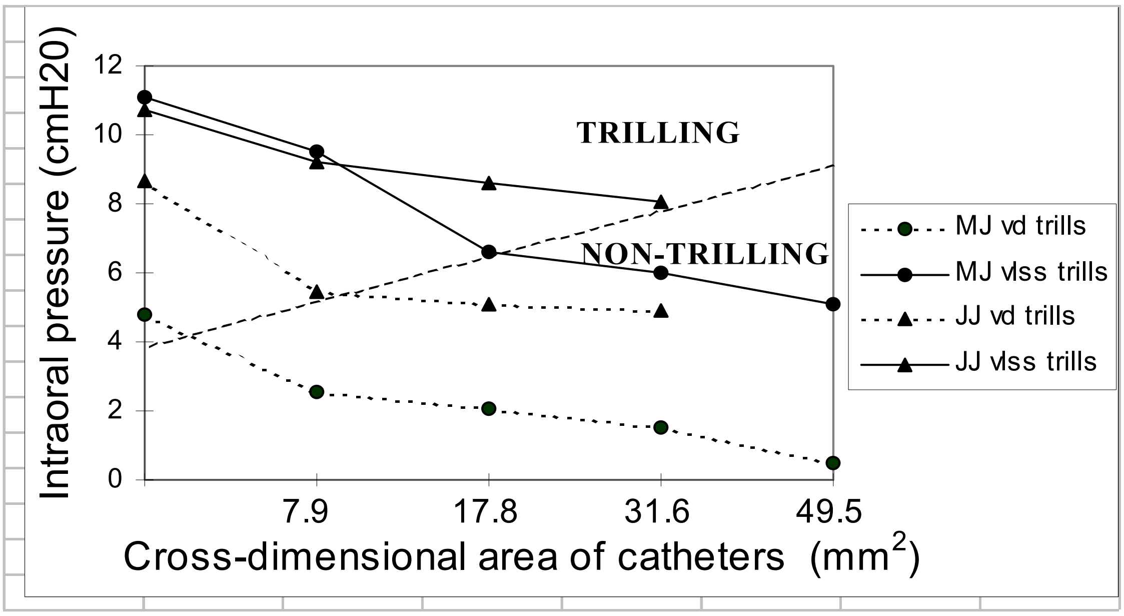

FIGURE 4. Reduction in intraoral pressure (y axis) per catheter cross-sectional area (x axis) for steady state voiced (dotted line) and voiceless (solid line) trills produced by speakers MJ and JJ. The threshold of trilling is indicated by a dashed line.

Comparison of Figures 4 and 5 shows that voiced trills were extinguished not only when vented with catheters with values for impedance similar to or lower than impedance at the oral constriction (catheter areas ≥ 31.7mm^2), as for voiceless trills, but also when vented with catheters with higher impedances (areas 7.9 and 17,8mm^2) -- most probably due to bleeding the Po below the threshold necessary for trilling--, thus reflecting the uncompromising aerodynamic requirements of voiced trills.

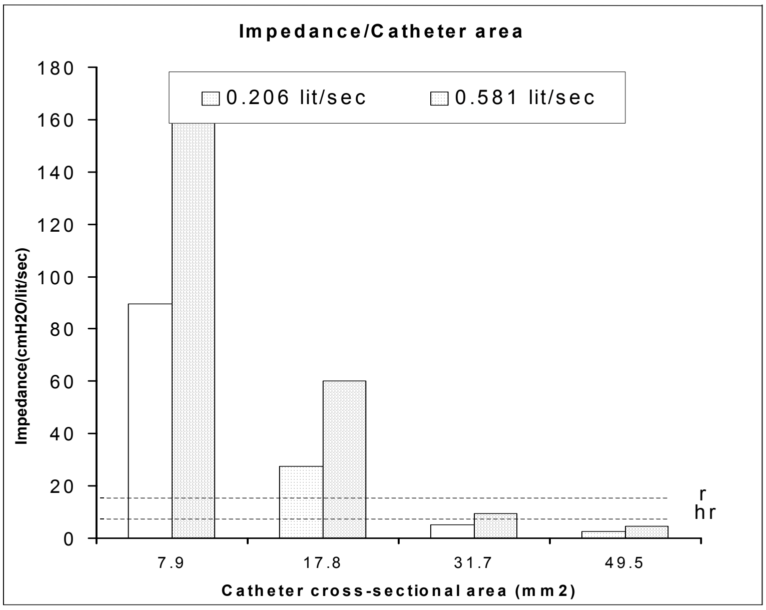

Figure 5. Impedance (ordinate) for catheters of varying cross-sectional area (abscissa) at the flowrates used in trills (in lit/sec). Values of measured vocal tract impedance for the voiced and voiceless trills have been overlayed on the graph (Speaker MJ).

4.1.3. The results show that the preference for voiced over voiceless tongue-tip trills -- *i.e.*, of the feature [trill] co-occurring with [+voice] -- cannot be accounted for exclusively in aerodynamic terms. According to the 'stability of the output' criterion presented in 1.1.2 above, *voiceless* trills should be preferred since (1) they are articulatorily more robust in varying aerodynamic conditions (gradual reductions in Po extinguished voiceless trills later than voiced trills), and (2) they have less constrained aerodynamic requirements (in voiced trills Po may vary over a very narrow range in order to sustain voicing and trilling). Auditorily, however, voiceless trills are poorly differentiated from fricatives, as shown in Figure 6, where contrastive voiced and voiceless trills for Lai-chin are shown. The aerodynamic and articulatory characteristics of voiceless trills, vis-à-vis voiced trills (glottal friction; higher Po and higher particle velocity across the oral constriction, larger ratio open to closed period, and failure to achieve full palato-lingual closure) contribute to turbulent energy throughout the sound which makes them auditorily fricative-like.

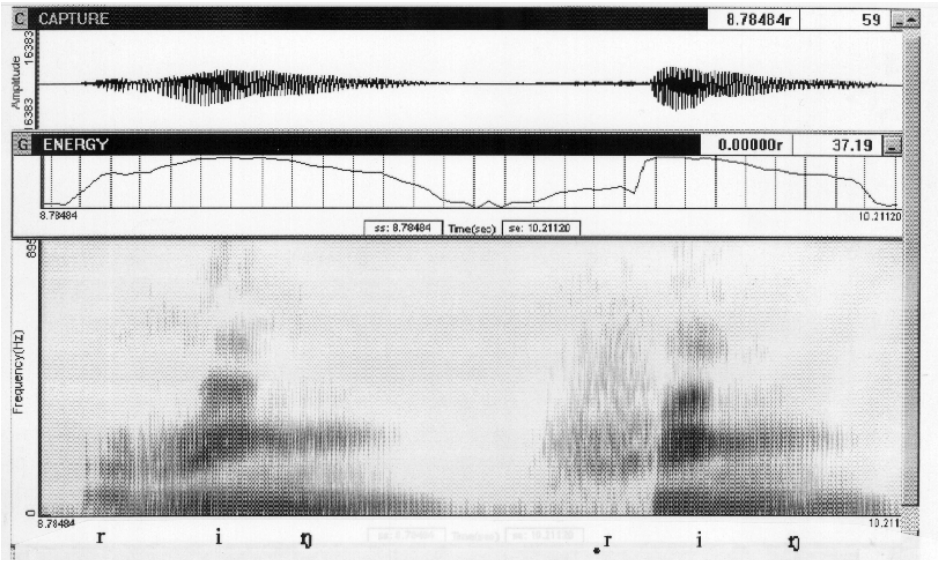

FIGURE 6. Voiced and voiceless trills in Lai-chin ([riŋ] ‘loud’ and [r̥iŋ] ‘green’).

The reduction in transglottal airflow associated to voiced trills makes them less preferable aerodynamically -- contrary to Maddieson's (1984) claim -- but contributes to a regular alternation of bursts of periodic energy. Such spectro-temporal discontinuities result in an auditorily distinct signal. In Figure 6 voiced trills show one clear contact followed by a burst of energy and subsequent decreases in intensity due to the vibration of the tongue tip approaching a closure. Thus, voiced trills exhibit severely constrained aerodynamic requirements and little articulatory stability vis-à-vis voiceless trills, but result in a clearly modulated signal, clearly distinct form other speech segments. The co-occurrence of trilling and voicing can be seen as the natural byproduct of 'auditory dictinctiveness'.

The preference for trilling to co-occur with voicing in phonological systems seems to reflect a trade-off between articulatory stability (*i.e.*, preserving trilling in a narrow range of aerodynamic conditions) and acoustic/auditory salience (*i.e.*, distinct signal modulation).

4.2. Lack of co-occurrence of trilling and nasality.

A feature co-occurrence restriction that is dictated exclusively by aerodynamic factors is the absence of nasal trills. An open velo-pharyngeal port for nasality would bleed the intraoral pressure required to make the relaxed tongue-tip vibrate for trills. Fig. 4 shows that when Po was vented with catheters of varying areas, simulating co-occurring nasalization, tongue-tip trilling ceased when Po dropped below a certain threshold. Lingual trills require a minimum intraoral pressure of 4cmH20 to sustain tongue tip vibration (and even higher values to initiate it), and voiced trilling was extinguished when the Po was reduced by 2.5 cmH2O (vent areas ≥ 7.9 or 17.8 mm^2, depending on absolute Po values). Thus, trilling cannot combine with an open velopharyngeal port that reduces the Po by 2.5 cmH20 or more. As argued for fricatives in section 3, the small velopharyngeal openings which do not impair trilling would be insufficient to create a percept of nasalization. Thus, aerodynamic factors explain the lack of nasal trills.

5. CONCLUSIONS

This study provided empirical evidence on how aerodynamic factors in combination with constraints of speech perception may determine phonological structure, specifically feature co-occurrence restrictions. We analyzed the aerodynamic conditions required for frication and trilling and how they were affected by co-occurring [voice] and [nasality]. It was found that sufficient velopharyngeal opening to create the percept of nasalization bled the high oral pressure necessary to create audible frication. Thus aerodynamic principles account for the difficulty in producing phonetically nasalized fricatives. The more costly aerodynamic conditions required to produce perceptually robust nasalized fricatives, or nasalized trills for that matter, *e.g.*, increased subglottal pressure, do not result in correspondingly salient auditory effects, that is, nasalized fricatives are auditorily close to fricatives and voiceless nasals. In summary, the fact that [nasal] and [fricative] do not combine in segments can be attributed to aerodynamic constraints and lack of optimization of production and perceptual factors.

Voiced trills showed highly constrained aerodynamic requirements and were less stable articulatorily than voiceless trills, but exhibited a clearly modulated and distinct signal, whereas voiceless trills are auditorily similar to fricatives. The common combination of trilling and [+voice] in phonological systems thus reflects a trade-off between competing demands of perception and production: preserving maximum auditory distinctiveness in a narrow range of aerodynamic conditions.

The view that phonological structure emerges from phonetic principles such as articulatory-acoustic stability, auditory contrast and articulatory economy has a number of methodological and theoretical advantages. First, these principles can be explicitly defined in terms of physical and perceptual phenomena. Second, the proposed phonetic principles can be measured and quantified thus providing us with great predictive and explanatory power. For example, once a relationship between certain feature values is discovered, like the minimum Po required to sustain trilling and the effect of velopharyngeal opening on Po, we can predict that, if the initial conditions are not changed, trilling and nasalization cannot co-occur. Third, the phonetic principles and mechanisms offered as explanations have been independently motivated and allow a 'natural', phonetically motivated account of phonological structure.

We have illustrated that aerodynamic constraints and perceptual constraints interact to adequately shape phonological structure, optimizing requirements of perception and articulation, and ensuring robustness. In Darwinian terms, we have shown that complex phonological structure can emerge from simpler (phonetic) mechanisms; and feature co-occurrence restrictions seem to reflect the darwinian drift towards patterns that almost seem designed for their environment -- efficient speech production and perception.

ACKNOWLEDGMENTS

- This research was supported by DGICYT grant PB 96-1158 to the Universitat Autònoma de Barcelona, Spain.

NOTES

1. Phonological reflexes of velic opening bleeding the pressure build up needed for fricatives are found in on-line speech, English *isn't*, *doesn't* pronounced [ɪnnt], [dʌnnt] (Gimson 1962); diachronically, /s>j/ before a nasal, Latin *mans(io)nata > mas'nata > Catalan *mainada* 'kids' (Badia Margarit 1951), and dialectally, Standard Catalan *esma* 'judgment' (< Latin *aestimare*) pronounced ['ejmə] Conflent, ['əjmə] Mallorca, ['ęjmę] Maó dialect (Alcover and Moll 1972). Similarly, synchronic and diachronic epenthetic stops between nasals and fricatives/stops, *e.g.*, English *incidence* [nts], *glimpse* [mps](< OE *glimsian*); Catalan *ensabonar* 'to lather' [nts], reflect an early closing of the velum to allow sufficient pressure build up for the fricative/stop.

REFERENCES

BADIA MARGARIT, A.M. 1951. *Gramàtica Històrica Catalana*. 3rd ed. Valencia: Biblioteca d'Estudis i Investigacions. Tres i Quatre.

BEHRENS, S. and S.E. BLUMSTEIN .1988. On the role of amplitude of the fricative noise in the perception of place of articulation in voiceless fricative consonants. *Journal of the Acoustical Society of America* 84, 861-867.

BROWMAN, C.P. and L.M. GOLDSTEIN, L. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3, pp. 219-252.

-----. 1990. Tiers in Articulatory Phonology. In J. Kingston and M. E. Beckman (eds.), *Papers in Laboratory Phonology I* (pp. 341-376). Cambridge. Cambridge University Press.

COHN, A.C. 1993. The status of nasalized continuants. In M.K. Huffman & R.A. Krakow (eds), *Nasals, nasalization and the velum* (329-367). San Diego, CA: Academic Press.

GERFEN, H.J., Jr. 1996. Topics in the phonology and phonetics of Catzospan Mixtec. University of Arizona, doctoral dissertation.

GIMSON, A.C. 1962. *An introduction to the pronunciation of English*. 2nd ed. London: Arnold.

KAWASAKI, H. 1982. An acoustical basis for universal constraints on sound sequences. University of California, Berkeley, doctoral dissertation.

LADEFOGED, P. and I. MADDIESON. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

LAHIRI, A. and W. MARSLEN-WILSON. 1992. Lexical processing and phonological representation. In G. Docherty and R. Ladd (eds), *Papers in Laboratory PhonologyII: Gestures, segment, prosody*. Cambridge: Cambridge University Press, pp.229-254.

LI, F. K.1997. *A handbook of comparative Thai*. Hawaii: The University Press ofHawaii.

LILJENCRANTS, J. and B. LINDBLOM. 1972. Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* 48, 839-862.

LINDBLOM, B. 1983. Economy of speech gestures. In P.F. MacNeilage (ed), *The Production of Speech* (pp. 217-246). New York: Springer-Verlag.

-----. 1986. Phonetic universals in vowel systems. In J.J. Ohala and J.J. Jaeger (eds), *Experimental Phonology,* (pp.13-44). Orlando, FL: Academic Press.

-----. 1990. Explaining phonetic variation: a sketch of the H&H theory. In W.J. Hardcastle and A. Marchal (eds), *Speech Production and Speech Modelling* (pp.403-439). Dordrecht: Kluwer.

-----. 1992. Phonological units as adaptive emergents of lexical development. In C.A. Ferguson, L.Menn and C. Stoel-Gammon (eds), *Phonological development: models, research, implications*. Parkton, MD: York Press.

MADDIESON, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

MAEDA, S. 1993. Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In M.K. Huffman & R.A. Krakow (eds), *Nasals, nasalization and the velum* (147-167). San Diego, CA: Academic Press.

MOLL, K. and DANILOFF, R. 1971. Investigation of the timing of velar movements during speech. *Journal of the Acoustical Society of America* 50, 678-684.

OHALA, J.J. 1974. Phonetic explanantions in phonology. In A. Bruck, R. Fox & M. LaGaly (eds.), Papers from the parasession on natural phonology. Chicago: Chicago Linguistics Society, 251-274.

-----. 1976. A model of speech aerodynamics. *Report of the Phonology Laboratory* (Berkeley), 1, 9-107.

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

-----. 1983. The origin of sound patterns in vocal tract constraints. In P.F. MacNeilage (ed), *The Production of Speech* (189-216). New York: Springer-Verlag.

-----. 1990. The phonetics and phonology of aspects of assimilation. In J. Kingston & M.E. Beckman (eds), *Papers in Laboratory Phonology I: Betwen the grammar and physics of speech* (pp. 258-275). Cambridge, U.K.: Cambridge University Press.

-----. 1993. The phonetics of sound change. In Ch.Jones (ed) *Historical Linguistics: Problems and perspectives* (pp.237-278). London: Longman

-----. 1998. Can nasalized fricatives exist? Paper presented at the 24th Meeting of the Berkeley Linguistics Society. University of California, Berkeley.

OHALA, J.J. and OHALA, M. 1993. The phonetics of nasal phonology: theorems and data. In M.K. Huffman & R.A. Krakow (eds), *Nasals, nasalization and the velum* (225-249). San Diego, CA: Academic Press.

OHALA, J.J., M.J. SOLÉ, and G. YING. 1998. The controversy of nasalized fricatives. *Proceedings of the 135th Meeting of the ICA/ASA.* Seattle, Washington, pp. 2921-2922 .

PERKELL, J.S., M.L. MATTHIES, M.A. SVIRSKY, and M.I. JORDAN. 1993. Trading relations between tongue-body raising and lip-rounding in production of the vowel /u/. *Journal of the Acoustical Society of America* 93. 2948-2961.

PULLEYBLANK, D. 1997. Optimality theory and features. In D. Archangeli and D.Terence Langendoen (eds), *Optimality Theory. An overview* (pp.59-101). Oxford: Blackwell.

RECASENS, D., FONTDEVILA, J. & M.D. PALLARÈS. 1995. A production and perceptual account of palatalization. In B. Connell & A. Arvaniti (eds), *Papers in Laboratory Phonology IV: Phonology and phonetic evidence.* Cambridge, U.K.: Cambridge University Press.

SCHADEBERG, T.C. 1982. Nasalization in UMbundu. *Journal of African languages and Linguistics* 4, 109-132.

SOLÉ, M.J. 1998. Phonological Universals: Trilling, voicing and frication. *Proceedings of the Berkeley Linguistics Society* 24. University of California, Berkeley, pp. 403-416.

SOLÉ, M.J., OHALA, J.J. and G. Ying. 1998. Aerodynamic characteristics of trills. *Proceedings of the 135th Meeting of the ICA/ASA.* Seattle, Washington, pp. 2923-2924 .

SPROAT, R.W. and FUJIMURA, O. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21, 291-311.

STERIADE, D. 1997. Phonetics in phonology: the case of laryngeal neutralization. UCLA, Ms.

STEVENS, K.N.1972. The quantal nature of speech: Evidence from articuatory-acoustic data. In P.B. Denes and E.E. David Jr. (eds), *Human Communication, A Unified View* (51-66). New York: McGraw-Hill.

-----. 1989. On the quantal nature of speech. *Journal of Phonetics* 17. 3-46.

STRINGER, M. and J. HOTZ 1973. Waffa phonemes. In H. McKaughan (ed), *The language of the Eastern Family of the East New Guinea Highland Stock*. University of Washington Press, Seattle, 523-529.

WESTBURY, J.R., and P.A. KEATING. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145-166.