Zoltán Kiss

# Markedness, graduality and closedness in phonotactics—a phonetically grounded approach*

## 1 Introduction

This paper attempts to present a fresh view of consonantal phonotactics by employing a functionalist, non-representational approach. The central claim, which is becoming more and more prevalent in the phonological literature these days is that phonological processes can be explained if various functional arguments are made use of. The most important of these arguments is that phonological phenomena are influenced by the phonetic factors of sound perception and production. The paper wishes to argue that even such static phonological events as the distribution of sound segments can be satisfactorily explained provided that functionalist principles are considered.

The paper is structured as follows. The following section (§**2**) provides a short historical overview on the role of phonetics in phonology. §**3** sums up the most important issues in a functional approach to phonology. The clarifying of the meaning of markedness as being used in this paper is indispensable, and so this is done in §**4**. §**5** introduces the most important functional principle of the paper, Phonotactic Closedness (cf. Rebrus & Trón 2002, 2004); its workings in Hungarian and English are also described in this section. §**6** sums up the most important findings.

## 2 The role of phonetics in mainstream phonology —a brief historical overview

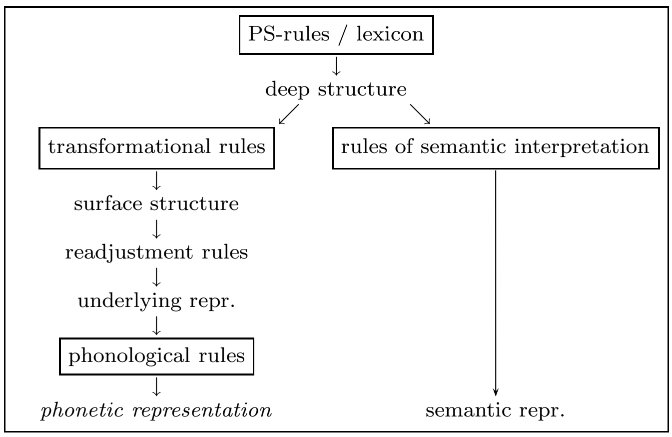
If one considers the history of phonology, it seems clear that the role of phonetics in describing and explaining various phonological patterns has always been minimal and secondary until today. In structuralist phonology, for instance, the main—procedural—role of phonetics was to provide an articulation-based *description* and *classification* of segments (cf. the *natural classes* of later generative phonology). The notion "phonetic similarity" was, for example, evoked in identifying the allophonic/phonemic status of segments: two sounds were supposed to be allophones of a phoneme provided that they were phonetically similar, besides also standing in complementary distribution (the case of English **k** and **p**$^{\mathbf{h}}$, as well as **ŋ** and **h** are widely known and often cited examples; cf. e.g., Lass 1984 : 18).

* Central parts of this paper have been inspired by conversations with Péter Rebrus, Miklós Törkenczy, Zsuzsi Bárkányi and Anna Kürti. I would also like to thank Péter Rebrus and Miklós Törkency for reading and commenting earlier drafts of this paper. As always, the remaining mistakes are all mine.

The bias for articulatory phonetics on the part of phonologists is understandable: no special training is required to examine and describe the various movements of the articulators, they are easily observable. A serious drawback of the articulatory bias is, nevertheless, that it only considers the role of the speaker, and ignores that of the listener — after all, it is the speech signal ("the acoustic experience") that is shared by both the speaker and the listener. It was not until the development of the sound spectrograph that the role of the *listener* was put in the foreground. A pioneering step in this respect was the publishing of *Preliminaries to speech analysis: the distinctive features and their correlates* (Jakobson et al. 1952). In it, the atoms of phonology — the distinctive features — gained a central theoretical role: it is features that phonological processes access, for example. Crucially, the Jakobsonian features were defined primarily in *acoustic*, rather than in articulatory terms.[1]

Most probably because of the novelty of the technology,[2] acoustic phonetics was thrust into the background again, and phonologists returned to the traditional and "more reliable" articulatory stance. This is fairly evident in mainstream generative phonology (as put forth in Chomky & Halle 1968). Speech articulation played the following main roles in mainstream "classical" generative phonology: (i) distinctive features were primarily defined in articulatory terms; (ii) phonetics was evoked in deciding the "naturalness" of phonological representations and rules;[3] and (iii) it was the phonetic module that was responsible for *interpreting/implementing* the output of phonological rules (the surface/phonetic representation). Consider (1), which presents Mohanan's (1995 : 27) reconstruction of the generative grammar model:

(1)




[1] There were some features, actually, that were not given an articulatory definition at all; for example, "Compact/Diffuse", "Grave/Acute", "Tense/Lax", etc.

[2] On this, see Joos's comments, cited in Hume & Johnson (2001 : 1).

[3] A rule was said to be natural if it was phonetically motivated; consider the "Naturalness Condition" of Postal (1968) and Hooper (1976).

In this model, thus, the relation of phonology and phonetics is strictly unidirectional and interpretational, phonetic principles can never have a *direct* effect on phonology,[4] phonetics is deemed to interpret phonology's output: certain phonetic laws are supposed to direct the way how phonological representations are to be mapped onto articulatory instructions (how it should work in practice was seldom made explicit, though). As Steriade puts it, "[t]he downward arrow connecting phonology to phonetics means that the decision to have a contrast and have it in a specific position is taken in phonology. [...] Phonetic implementation has to live with prior decisions taken in the phonology" (1997 : 3).

A welcome effect of the Jakobsonian approach to distinctive features in phonology has been, however, that features are often assumed to have both acoustic and articulatory definitions (even if speech articulation is usually given the precedence). One noteworthy approach where the atoms of phonology are not primarily described in articulatory fashion is the privative element theory of Harris & Lindsey (1995). In this theory, the elements that make up sound segments are specified in acoustic terms, although the authors add that

> "elements should [not] be construed as acoustic (or articulatory) events. They are properly understood as cognitive objects which perform the grammatical function of coding lexical contrasts. Nevertheless, [...] we consider their phonetic implementation as involving in the first instance mapping onto sound patterns in the acoustic signal. Viewed in these terms, articulation and perception are parasitic on this mapping relation. That is, elements are internally represented pattern templates by reference to which listeners decode auditory input and listeners orchestrate and monitor their articulations." (ibid. : 50f)

Again, the exact workings of these mapping mechanisms have not been made explicit, and crucially, phonetics is still designated to solely play the role of implementing phonological representations.

Assuming that phonetics can also influence phonology, two important research questions are the following: what phonetic principles can affect phonology and how can this influence be modelled in the phonology? It is these issues that we turn to in the next section.

## 3 Functionalist phonology

There is more and more evidence that "language-external" factors, such as the physics of *speech perception* and *speech articulation* affect phonological systems, that they are part of phonological competence. It must be emphasized that it is the first time in the history of phonology that speech perception has been given any formal role in modelling sound systems. As Hume & Johnson (2001 : 2) claim, this is the result of two main factors: (i) rapid technological advancements in laboratory phonetics and phonology, and (ii) the development of Optimality Theory, which—according to the authors—provides a suitable framework "[to state] perceptually

[4] Or only inasmuch as they are "built" into phonology—in the shape of articulatorily defined features, phonetically motivated (i.e., "natural") rules, as well as markedness theory; see Chomky & Halle (1968 : 400ff) and Lass (1984 : 195ff).

grounded constraints which interact dynamically with constraints motivated by other principles." Probably a major development that underlies the importance of a functionalist approach like this is that independently verifiable universal physical factors can be evoked in explaining sound patterning. One advantage of this is that the principles or constraints set up to account for phonological patterning are not solely abstract representational constructs without physical (phonetic) grounding. For example, the fact that word-final stops tend to lenite (as in say, English *bet* **beʔ**) may be given a non-grounded, purely representational explanation, according to which the word-final position is a "weak" position, because — it could be argued — the stop there is followed by an "empty nucleus" which does not have enough "power" to license all the elements a stop is supposed to be made up of, hence it loses all its elements but its occlusion (thus **t** → **ʔ**).[5] It will be shown below that the word-final position is indeed weak, and a stop is not to be licensed there, but the explanation can be grounded in phonetic, more closely, perceptual factors: in that position the contrast in question is not well cued perception-wise (it is difficult to perceive), hence the neutralization (lenition) into the glottal stop.[6] Where the perception-based approach seems superior to the purely representational one is perhaps its less abstract nature, namely that certain "hardcore" physical aspects can be linked to the setting up of phonological principles responsible for sound patterning: the theory—at least in this respect—is thus less metaphorical, less *ad hoc*. Besides, more crucially, as Côté (2000 : 13) also argues, the perception-based account of phonotactic patterns (and other phonological patterning in general) is superior to representational accounts (specifically to the syllable or other prosody-based models) because the latter can be proven to be "insufficient, inadequate, [and] unnecessary." As it will turn out later, the fact that stops in word-final position usually lenite is *not accidental*—provided that the phonetic properties of that context are considered: a *direct relationship* between the event and the context can thus be established, the theory is more explanatory this way.

There are several important questions that arise in connection with a functionalist approach to phonology; I have listed some of them below (cf. Hume & Johnson 2001 : 2):

(2) a. To what extent do speech perception and articulation influence phonological systems?

b. To what extent do phonological systems influence speech perception and articulation?

c. Where do speech perception and articulation phenomena belong in relation to a formal description of the sound structure of language?

[5] This account of lenition (i.e., making use of empty constituents/positions) — admittedly presented here very sketchily — is prevalent in some offsprings of Government Phonology; see, among numerous others, Cyran (2003). Early treatments of licensing (and the lack of it in the coda, for example) include Itô's (1986).

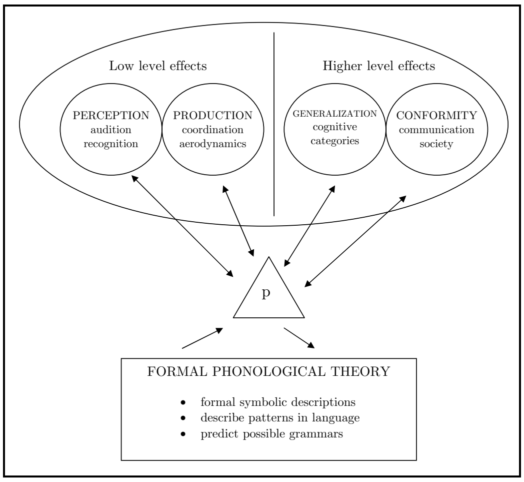
[6] It is true, though, that it is only the fact of neutralization that is predicted this way, but not the actual *output* of it (the glottal stop, in the example).

d. What other cognitive forces (non-phonetic, language-external ones) affect phonological patterning? Which aspects of phonology can be given a phonetically-grounded account, and which cannot be?

e. How can the interaction of functional principles and phonological patterning be represented in a formal model?

According to the functionalist model proposed in Hume & Johnson (2001 : 8), the interplay of language-external forces and phonology can be represented as in (3):

(3)




Without giving a detailed description of the model (the reader is referred to Hume & Johnson's article for that), let me just mention those aspects that are important for the present paper. First of all, perception and articulation (speech production) both have a direct effect on the phonological cognitive representation (represented by the *p* in the triangle). There are also "higher level" effects, generalization and conformity, which include various non-phonetic, yet language-external principles that apparently work in the shaping of sound patterns. These include the general ability of categorization (segmentation, contrast creation), analogy creation, the influence of the frequency of items and patterns in the lexicon, and, as I will claim below, the principle of *closedness* (or "non-interruption").

## 4 Contrast, segmental markedness and perception

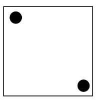
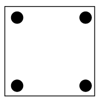
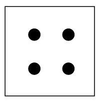
As it has long been established by functionalist accounts, phonological systems of languages are claimed to be shaped by the interaction of the following (partially conflicting) factors:[7]

[7] Cf. Zipf (1949); Liljencrants & Lindblom (1972); Diver (1979); Flemming (1996, 2001); Rebrus & Trón (2002).

(4) a. contrast creation;
  b. maximizing the number of contrasts;
  c. maximizing the perceptual distinctiveness of contrasts;
  d. minimizing the articulatory effort.

The first of these principles is responsible for the creation of contrastive cognitive categories; by maximizing the number of contrastive categories (4b), the expressiveness of communication is enhanced by building up a substantial lexicon of categories. Principle (4c) accounts for the salience of the distinct basic categorical elements—according to it, categories must have acoustic properties that make them maximally salient from each other perceptually. The last principle secures that the actual implementation (articulation) of the categories is to be carried out using as little energy as possible.

As Flemming (2001 : 5) shows, principle (4c) is inherently in conflict with both principles b. and d. Provided that in the two-dimensional phonological space (see the (5a)), there are two distinct categories (so there is only one contrast), and the two are perceptually well distinguishable (they occupy the opposing corners of this space—thus satisfying principle (4c) this way), principle (4b) is trivially violated, as well as (4d), as the two categories are far from each other in articulatory terms, too. If we try to satisfy principle (4b) by increasing the number of contrast, cf. (5b), the requirement for perceptual distinctiveness is violated, as some categories will necessarily be closer to each other. In (5b), articulation is still energy-consuming; a way to minimize articulatory effort is to bring the categories closer to each other (they are thus produced at a similar place, for example); however, this sacrifices their perceptual salience (5c).

(5) a.  b.  c. 

It seems clear then that in a functional theory of contrast, some weighing of potentially conflicting principles is inevitable and the weighing may well be language-specific.[8]

Segmental markedness as described in most works on phonology is usually defined in absolute, universal and context-independent terms. In frameworks like those, a (contrastive) segment is said to be marked if it occurs in a relatively small number of languages. A typical example for this approach is Maddieson (1984). For example, since all languages have stops (as opposed to, say, liquids), stops are universally and typologically unmarked. Statements like these form the basis of implicational universals, like, for example, that the presence of a liquid in a language necessarily implies the existence of a stop, too. However, as I will argue below, segmental markedness is more meaningful if it is defined in terms of *relative contrast*, *context* and *perceptual factors*.

[8] This is perhaps why functional phonological approaches are usually shaped in Optimality Theoretic terms.

### 4.1 Segmental markedness is relational

In absolute terms, the vowel **ɯ** for instance is marked, because non-low back vowels are generally rounded (93.5% of the languages in Maddieson's (1984 : 124) database); also, within a language, if it has a contrastive unrounded back **ɯ**, it must generally have its rounded pair **u**, too; the reverse, however, does not usually stand.[9] The perceptual account of the universal markedness of **ɯ** can be briefly summarized as follows. It is a well-know fact that if a langauge has five contrastive vowels, they are **i, e, a, o** and **u**. This is said to be an optimal system because it fills the available phonological/acoustic space the most optimally. Considering the horizontal dimension, we can say that the front – back contrast is along the line of the vowels' formant 2 values (**i** has the highest F2, **u** the lowest). It is also a well-established phonetic fact that rounding lowers F2, and so a rounded high vowel is maximally distinct from its front unrounded pair in F2.

Obviously then, the occurrence of an unrounded back vowel (or a rounded front vowel) in this system makes it suboptimal. What must be emphasized though is that the suboptimality of the hypothetical {**i e a o ɯ u**} system is only due to the perceptual markedness of **ɯ** *with respect to* **u** because their F2 values will be very similar. If we relate **ɯ** to **i**, their F2 values will be on the two ends of the F2 scale, and this way then **ɯ** will not be marked since **i** and **ɯ** are perceptually distinct. It is thus not **ɯ** in itself that is marked but its *contrast* with **u**; as Flemming (2001 : 3) puts it, "[the] markedness of sounds is indeed dependent on the contrasts that they enter into."[10]

### 4.2 Segmental markedness is contextual

A contrast may well be perceptually unmarked in a given context, yet the same contrast is marked in another. In other words, segmental markedness must also be related to the context it occurs in: certain positions favour segmental contrast because in those particular contexts the contrast is well-cued, while in others the same contrast is less salient. This idea is expressed in Steriade's *Licensing by Cue* principle:

(6) *Licensing by Cue* (Steriade 1999 : 4)
The likelihood that distinctive values of the feature $F$ will occur in a given context is a function of the relative perceptability of the $F$-contrast in that context.

Let us briefly consider the salience of the voicing of stops in various environments (based on Steriade 1997), using hypothetical examples:

[9] Japanese is exceptional in this respect with an {**i e a o ɯ**} vowel inventory. Here, effort minimization is preferred over maximal perceptibility.

[10] Flemming (*ibid.*) also shows that a segment that is universally/typologically marked may well be unmarked within a system which does not make use of a particular contrast. For example, in the back – front dimension, high central **ɨ** is universally marked, but in languages that do not contrast back – front vowels (the so-called "vertical" vowel systems, like Kabardian, Marshallese), the vowels that actually occur have a central quality (like **ɨ** does). Crucially, no "vertical" languages exist with a {**i e a**} or {**u o a**} inventory.

(7) Perception cues for the voicing of stops in various environments:

a. (i) $V_1$__$V_2$: ***apa, aba***; (ii) $V_1$__son: ***apra, abra***
cues: voicing of closure; length of closure; length of $V_1$; F1 of $V_1$; length/strength of release; VOT value; F0 and F1 of $V_2$

b. (i) #__: ***pa, ba, pra, bra***; (ii) obstr__son: ***aspa, asba, aspra, asbra***
cues: voicing of closure; length of closure; length/strength of release; VOT value; F0 and F1 of $V_2$

c. V__#: ***ap, ab***; cues: voicing of closure; length of closure; length of V; length/strength of release

d. V__obstr: ***apsa, absa***; cues: voicing of closure; length of closure; length of $V_1$; F1 values of $V_1$

e. obstr__obstr: ***aspta, asbta***; cues: voicing of closure; length of closure

f. obstr__#: ***asp, asb***; cues: voicing of closure; length of closure

g. #__obstr: ***psa, bsa***; cues: voicing of closure; length of closure

(7a) is the context which provides the most cues for the contrast in question; as we go down in this list to (7e–g), the number of the cues is less and less. In this sense then, the hypothetical contrast of *apa – aba* is less marked (i.e., less difficult to perceive) than that of *psa – bsa.* According to the principle of Licensing by Cue, the *psa – bsa* contrast is not likely to occur; it is in fact in the badly cued contexts where we expect the neutralization of the contrast. This state of affairs have two important consequences. The first is that phonotactic patterns can be related to perceptual markedness. Still remaining with our hypothetical example, the fact that in a language there are no forms with a word-initial **bs** cluster (there are only word-initial **ps** clusters) is a direct upshot of the fact that **b** in this position is not salient perceptually—hence the neutralization of the **p** – **b** contrast.

The other important result of this approach is that markedness is based on context. Specific categories need specific positions to be perceptually salient. The place contrast of stops, for example, is best perceived when the stop is before a vowel, but less salient before another stop. Retroflexion, however, is best perceived if the retroflex stop *follows* a vowel; in postvocalic position, the contrast between retroflex stops tends to be neutralized (cf. Steriade 1999). Phonological patterning is thus sensitive to various dimensions: one category (contrast) in one position may be perceptually unmarked, but the same contrast may well be marked when considering another dimension (such as position).

## 5 Phonotactic Closedness

The list in (7) can thus be translated into a perceptual difficulty (markedness) scale of a given segmental contrast (**p** – **b**). Markedness scales like (7), together with the principle of Licensing by Cue, predict what contrast in what environment is likely (unmarked) and in what context it is likely to be neutralized. Importantly, these scales *predict the typology of phonotactic patterns found in languages*: which patterns are possible and which are most unlikely. The difficulty scale based on (7) is shown in (8), where "$A > B$" means that $A$ is a more marked/difficult position perceptually for the given contrast than $B$, because it provides less/worse perception cues:[11]

(8) *Perceptual difficulty scale for the voicing contrast of obstruents*:
{O__O , O__#, #__O} > V__O > V__# > {#__, O__R} > V__R

As Steriade (1997 : 17f) shows, one type of voicing neutralization pattern (represented by Polish, Lithuanian, Sanskrit, etc.) corresponds with the scale in (8):

(9) One voicing neutralization pattern:

a. The voicing of obstruents is neutralized word finally (only a voiceless obstruent can occur).
Lith.: *daug* **dauk** 'much', *kad* **kat** 'that'

b. The voicing of obstruents is neutralized before obstruents (there is regressive voicing assimilation).
Lith.: *atgal* **-dg-** 'back', *degti* **-kt-** 'burn-inf.'

c. Obstruents are distinctively voiced before sonorants (vowels/son. Cs).
Lith.: *aukle* **-kl-** 'governess', *auglingas* **-gl-** 'fruitful', *silpnas* **-pn-** 'weak', *skobnis* **-bn-** 'morning'

The table in (10) displays examples for the patterns of the voicing neutralization of stops (taken from Steriade 1997 : 9):

(10)

| | #__O, O__# | R__O | R__# | #__R | R__R |
|---|---|---|---|---|---|
| Totontepec Mixe | | | | | + |
| Lithuanian | | | | + | + |
| French | | | + | + | + |
| Shilha | | + | + | + | + |
| Khasi | + | + | + | + | + |

fewer/weaker cues ⟷ more/stronger cues
(more marked environment) (less marked env.)

The + indicates that the contrast is available in the given language in the specific environment. The importance of the table above lies in its empty cells: as Steriade says, "no language surveyed maintains the voicing contrast in a [perceptually] less informative context, *unless it also does so in the more informative contexts*" (*ibid.*; emphasis mine). Thus, for example, no language neutralizes the voicing of stops word finally after a vowel without also neutralizing medially in the V__obstruent context.

[11] O = any obstruent; R = any sonorant.

Difficulty hierarchies like (8) can therefore be claimed to set the boundaries of phonological systems, more specifically, that of phonotactic patterns. They delimit what segment combinations can occur in which positions. It can be argued that if a contrast occurs in a given context, then the same contrast will necessarily occur in another context *which provides better perception cues for the contrast*; in simple terms: the existence of the "more difficult" implies the existence of the "less difficult". This idea is phrased in the principle of Phonotactic Closedness:

(11) *Phonotactic Closedness* (cf. Rebrus & Trón 2002 : 21)

If a given contrast occurs in a perceptually marked environment (one providing few/weak cues), it will also occur in a perceptually less marked environment (with more/better cues). Therefore, the set of segmental contrasts is closed with respect to positional markedness, towards the unmarked cases: the more marked implies the presence of the less marked.

Phonotactic Closedness predicts systems like (12a), but no systems like (12b):[12]

(12) a.

| | T–D |
|---|---|
| #__O (*psa–bsa*) | □ |
| O__# (*spa–sba*) | □ |
| R__O (*apta–abta*) | • |
| R__# (*ap–ab*) | • |
| __R (*pa–ba*) | • |
| R__R (*apa–aba*) | • |

b.

| | T–D |
|---|---|
| #__O (*psa–bsa*) | □ |
| O__# (*spa–sba*) | □ |
| R__O (*apta–abta*) | • |
| R__# (*ap–ab*) | □ |
| __R (*pa–ba*) | • |
| R__R (*apa–aba*) | • |

The figures in (12) illustrate the voicing contrast of stops in specific environments; the environments are hierarchically ordered in terms of perceptual difficulty (cf. (7) and (8)), R__R being the least marked environment for the voicing contrast. The filled bullet indicates that the contrast in question is attested, while the empty square shows that the contrast is missing (thus only the unmarked occurs, the voiceless stop). (12b) violates Phonotactic Closedness because there is **p** – **b** contrast in R__O, but no contrast in R__#—a better cued environment than the other. Closedness predicts that there cannot be gaps in hierarchies like (12). Notice that closedness is a consequence of the functional principles introduced in (4): the phonotactic space is filled with elements in a way that they prefer to occur in a perceptually favourable context so that the identification of the contrast may be easier ("maximize perceptibility").

The following sections now turn to the testing of the claims introduced here concerning the consonantal phonotactics of two languages, Hungarian and English.

## 5.1 Closedness in Hungarian

The workings of the Closedness principle in the phonotactic patterns of Hungarian monomorphemic stems are illustrated in Rebrus & Trón (2002 : 22ff). Let us first

[12] T = any voiceless stop; D = any voiced stop.

consider the place contrast of voiceless stops in the following contexts: (i) prevocalic, and before coronal (ii) liquids, (iii) nasals, (iv) voiceless fricatives, (v) voiceless stops. The perceptual difficulty hierarchy of these contexts for the relevant place contrasts is (13a). According to Rebrus & Trón (2002), the perceptual difficulty scale for the voiceless non-coronal stops before coronal consonants is (13b):[13]

(13) a. __stop > __fric. > __nasal > __liquid > __vowel

b. **c** > **p** > **k**

These two dimensions (place contrast and its environment) define a phonotactic space filled by the following forms in Hungarian:[14]

(14) Hungarian $V_1C_1C_2V_2$; $C_1C_2$: voiceless, $C_1$: non-coronal stop, $C_2$: coronal (*ibid.*: 23)

| | | | |
|---|---|---|---|
| stop **t** | a**kt**a | kri**pt**a | *__ct__ |
| fricative **s** | bu**ks**a | ka**ps**ula | *__cs__ |
| nasal **n** | a**kn**a | sra**pn**el | *__cn__ |
| liq. **l** | cé**kl**a | pa**pl**an | %tro**cl**i |
| V | ba**k**a | ré**p**a | ku**c**a |
| | **k** | **p** | **c** |

Notice that the existing forms fill in the space as required by Phonotactic Closedness: there are no gaps within the phonotactic space, the forms at the extreme areas imply the presence of the ones below them. The non-existing forms are marked on both dimensions: they occupy the outermost positions in the difficulty scales.

As it was mentioned, markedness is relational and contextual. If we change the context of the given contrast, the markedness of the a given category (like the place of voiceless stops) may well change. This is the case if we consider the place contrast of voiceless non-labial stops before labials. The markedness scale of the places is now different:

(15) a. **t** > **k** > **c**

b. Hungarian $V_1C_1C_2V_2$; $C_1C_2$: voiceless, $C_1$: non-labial stop, $C_2$: labial

| | | | |
|---|---|---|---|
| stop **p** | pi**cp**ang | *__kp__ | *__tp__ |
| fricative **f** | fi**cf**ene | bu**kf**enc | (hé**tf**ő) |
| nasal **m** | fi**cm**a | la**km**usz | ri**tm**us |
| approx. **v** | ka**cv**asz | le**kv**ár | pi**tv**ar |
| V | ku**c**a | ba**k**a | sa**t**u |
| | **c** | **k** | **t** |

[13] **c** is used in its IPA value (voiceless palatal stop).

[14] The glosses of the items in (14) are: *akta* 'document', *akna* 'shaft' *baka* 'soldier', *buksza* 'purse', *cékla* 'beet', *kapszula* 'capsule', *kripta* 'crypt', *kutya* 'dog', *paplan* 'duvet', *répa* 'carrot', *srapnel* 'srapnel', *trotyli* 'tramp'.

Before labials, it is now **c** which is the least marked (**c** was the most marked before coronals above).[15] The forms in this dimension of the phonotactic space confirm the scale in (15).[16] Again, crucially, Phonotactic Closedness is not violated: the gaps only occur at the edges of the scales.

## 5.2 English

This section focuses on the phonotactic space of intervocalic two-member consonant clusters of English monomorphemic words. The data was collected from a searchable electronic database[17] of about 70,000 English words. The charts below also include the number of the words in which the clusters in question occur,therefore, they indicate the approximate[18] lexical (type) frequency of the clusters. Since—as opposed to Hungarian—stress plays an important role in this language (cf. for example the neutralization of vowel contrast in an unstressed syllable), the clusters have been distinguished whether they occur before or after a stressed vowel.

The first diagram (16) shows the occurrence of voiceless non-coronal stops before coronal consonants (of which the obstruents are voiceless, too); in the first chart, it is the vowel following the cluster that is primary stressed (indicated by the accent), while in the second it is the first vowel that bears the stress. The markedness hierarchy (**p** > **k**) follows Rebrus & Trón (2002)'s assumptions (cf. 13).

(16) English $V_1C_1C_2\acute{V}_2$ – $\acute{V}_1C_1C_2V_2$: $C_1C_2$: v'less, $C_1$: non-cor. stop, $C_2$: cor.

| | **k** | **p** |
|---|---|---|
| stop **t** | di**kt**áte$_{83}$ | ca**pt**ívity$_{24}$ |
| fricative **s** | e**ks**íte$_{139}$ | u**ps**ét$_{13}$ |
| nasal **n** | te**kn**íque$_{17}$ | hy**pn**ósis$_{2}$ |
| liquid **l** | a**kl**áim$_{101}$ | a**pl**ý$_{121}$ |
| V́ | a**k**ústom$_{1339}$ | a**p**óint$_{1133}$ |

| | **k** | **p** |
|---|---|---|
| stop **t** | cá**kt**us$_{336}$ | á**pt**itude$_{192}$ |
| fricative **s** | á**ks**ent$_{494}$ | ellí**ps**is$_{111}$ |
| nasal **n** | á**kn**e$_{67}$ | shrá**pn**el$_{21}$ |
| liquid **l** | á**kl**imate$_{320}$ | múlti**pl**y$_{256}$ |
| V | có**k**oa$_{2481}$ | có**p**y$_{1365}$ |

As (16) shows, the contrast between **p** and **k** is maintained in all positions: the phonotactic space is totally filled by existing words, even at the edges. As it is suggested by Rebrus & Trón (2002 : 24), the number of the words displaying the given phonotactic pattern increases with their markedness status monotonously: if one member of the opposition occurring in a given environment is less marked than the other, then it is supposed to be more frequent, too. This seems to be the case in (16), as well: in each environment it is the more marked **p** that is less frequent. As far as the markedness of the contexts is concerned, however, the numbers at

[15] The phonetic grounding of this hierarchy is only assumed in Rebrus & Trón (2002), as they do not provide phonetic evidence.

[16] The glosses of the items in (15b) are: *baka* 'soldier', *bukfenc* 'somersault', *fityfene* 'devil', *fityma* 'foreskin', *hétfő* 'Monday', *katyvasz* 'mixture', *kutya* 'dog', *lakmusz* 'litmus', *lekvár* 'jam', *pitvar* 'porch', *pitypang* 'dandelion', *ritmus* 'rhythm', *satu* 'vice'. *Hétfő* 'Monday' is the only stem with a **tf** cluster; it could well be considered a compound, too: *hét* 'week' + *fő* 'head'.

[17] This online dictionary is currently available at http://seas3.elte.hu/szigetva/etcetera.

[18] These numbers are probably far from being accurate; but they nevertheless exhibit important tendencies in the frequencies of the clusters.

least suggest that for this contrast (**k** – **p**) the pre-nasal position is more marked than any of the others. At this point, I am not aware of experiments involving the perceptibility of stops before nasals (especially, ones that show that stops are less salient in this context than in others), and so this issue needs further investigation. If it turns out that the pre-nasal position badly cues stop[19] opposition, the perceptibility hierarchy of context for this contrast must be revised (to something like __**n** > __**t** > __**s** > __**l** > V).

Let us turn our attention to the **g** – **b** contrast in the same environments as in (16) (except that now the obstruents following the two segments are *voiced*):

(17) English $V_1C_1C_2\acute{V}_2 - \acute{V}_1C_1C_2V_2$: $C_1C_2$: v'ed, $C_1$: non-cor. stop, $C_2$: cor.

| | g | b |
|---|---|---|
| stop **d** | ***gd** | a**bd**úct$_{7}$ |
| fricative **z** | e**gz**áct$_{148}$ | a**bz**ólve$_{23}$ |
| nasal **n** | i**gn**íte$_{75}$ | o**bn**óxious$_{7}$ |
| liquid **l** | ne**gl**éct$_{45}$ | o**bl**íge$_{37}$ |
| V́ | ci**g**ár$_{244}$ | a**b**óde$_{695}$ |

| | g | b |
|---|---|---|
| stop **d** | amý**gd**aloid$_{2}$ | á**bd**omen$_{17}$ |
| fricative **z** | é**gz**altation$_{6}$ | ó**bz**ervation$_{2}$ |
| nasal **n** | pré**gn**ant$_{164}$ | á**bn**egation$_{13}$ |
| liquid **l** | ú**gl**y$_{104}$ | bí**bl**ical$_{207}$ |
| V | é**g**o$_{1015}$ | ló**b**y$_{1490}$ |

The first thing that is apparent in the first chart of (17) is that it contains a gap in a position that violates Phonotactic Closedness: the less marked **g** is missing before **d** even though the more marked **b** does occur there (although only in seven words). There can be two approaches to resolve this problem. The first one is a little radical: it may well be the case that the markedness of the two segments (**g** and **b**) is to be reversed to **g** > **b**. This would necessarily place the gap in its "right" position: the *marked* segment would now occur in the *marked* context. As the frequency of the two segments also suggests, especially when they are before a vowel, the reordering of the two segments with respect to their markedness could be justified. According to Maddieson (1984 : 36), "[among languages that have voiced stops], **g** is more likely to be missing than **b** or [the coronals];" in other words, the universal markedness of voiced stops is **g** > **D** > **b**.[20] This markedness hierarchy is grounded in articulatory phonetics in Hayes & Steriade (2003 : 12ff). According to them, the aerodynamics of voicing requires that there be an active oral tract expansion (e.g., by advancing the tongue root or lowering the larynx) to maintain a continuous airflow so that the vocal folds may be able to vibrate during the production of a voiced stop.[21] If the dimension of place is also brought into the picture, it turns out that to maintain voicing for velar stops is more difficult than for non-velars: the production of bilabials necessarily creates a larger cavity in the mouth, "which allows the cavity to continue for a longer time to expand passively in response to airflow" (Hayes & Steriade 2003 : 12).

[19] At least that of the *voiceless* ones; see (17) below.

[20] **D** represents any voiced dental or alveolar consonant; Maddieson (*ibid.* : 35) claims that there are 199 languages with **b**, 195 with **D** and 175 with **g**. There are six languages whose only voiced stop is **b**, for instance, and only two which only contain a **D**. There are only 3 languages with **g** but without **b**, two of these also lack **D**.

[21] This is also a reason why *long* voiced stops are typologically marked: their production is in this sense more difficult to sustain than that of voiceless geminate stops.

(18) a. English $V_1C_1C_2V_2$ $C_1C_2$: v'ed, $C_1$: non-cor. stop, $C_2$: cor.

| | | |
|---|---|---|
| stop **d** | **gd**$_{2}$ | **bd**$_{24}$ |
| fricative **z** | **gz**$_{154}$ | **bz**$_{25}$ |
| nasal **n** | **gn**$_{239}$ | **bn**$_{20}$ |
| liquid **l** | **gl**$_{149}$ | **bl**$_{244}$ |
| V́ | **g**$_{1259}$ | **b**$_{2185}$ |
| | **g** | **b** |

b. English $V_1C_1C_2V_2$ $C_1C_2$: v'ed, $C_1$: non-cor. stop, $C_2$: cor.

| | | |
|---|---|---|
| stop **d** | **bd**$_{24}$ | **gd**$_{2}$ |
| fricative **z** | **bz**$_{25}$ | **gz**$_{154}$ |
| nasal **n** | **bn**$_{20}$ | **gn**$_{239}$ |
| liquid **l** | **bl**$_{244}$ | **gl**$_{149}$ |
| V́ | **b**$_{2185}$ | **g**$_{1259}$ |
| | **b** | **g** |

The other choice that is suggested by (17) is that perhaps stress does not play a role (i.e., it is not an active dimension)—at least not in the phonotactics of CC clusters. Because if we do not separate the two cases, in other words, we collapse the two charts, the gap disappears (see (18a)).

If we reverse the relative difficulty markedness of **b** and **g** (as suggested above), then we get the following chart in (18b). If the aerodynamics argument is valid, then the ordering has to be changed accordingly, as it is done in (18b). Notice that two environments are still problematic if we wish to maintain that the frequency of a cluster is parallel with its markedness: there are around six times more **gz** clusters than **bz**, and 12 times more **gn** clusters than **bn** (even if we disregard the stressing difference of the following/preceding vowel). It seems at this point that frequency is merely an indication of markedness but Rebrus & Trón (2002)'s claim about the relationship of frequency and markedness cannot be maintained. The frequency numbers clearly *indicate* that at the origin (the most unmarked area), there are always more items exhibiting the relevant cluster than at the edges (compare the VCV position with VC**d**V position, for example): the "density", as it were, of the phonotactic space is thus always heavier at the origin than at the outskirts.

The following dimension of the phonotactic space of English (19) shows voiceless non-labial consonants before labials (and vowels); notice that the voiceless velar stop is less marked than the coronal before labials (cf. (15)):

(19) English $V_1C_1C_2\acute{V}_2$ – $\acute{V}_1C_1C_2V_2$: $C_1C_2$: v'less, $C_1$: non-lab. stop, $C_2$: lab.

| | | |
|---|---|---|
| stop **p** | ***kp** | ***tp** |
| fricative **f** | ***kf** | ***tf** |
| nasal **m** | ***km** | ***tm** |
| approx. **w** | a**kw**íre$_{137}$ | be**tw**éen$_{6}$ |
| V́ | a**k**ústom$_{1339}$ | a**t**íre$_{654}$ |
| | **k** | **t** |

| | | |
|---|---|---|
| stop **p** | ***kp** | ***tp** |
| fricative **f** | (bréa**kf**ast$_{1}$) | ***tf** |
| nasal **m** | á**km**e$_{2}$ | ú**tm**ost$_{8}$ |
| approx. **w** | lí**kw**id$_{231}$ | reper**tw**íre$_{8}$ |
| V | có**k**oa$_{2481}$ | cí**t**adel$_{4700}$ |
| | **k** | **t** |

It is interesting that there are always more clusters if it is the *first* vowel that is stressed (compare, for example V**t**V́ (654 items) with V́**t**V (4700 items)). More importantly, at least for the topic of the present paper, Phonotactic Closedness is not violated.[22] If we consider the same situation, but this time with the contrast of the *voiced* non-labials (**g** – **d**), the picture is apparently problematic again:

[22] The frequencies of the clusters nevertheless are indicative of splitting the environment VC(C)V into VC(C)V́ and V́C(C)V. Whether the stressing of the first vowel makes the markedness hierarchy different for **k** and **t** (namely that if the first vowel is stressed, then

(20) English $V_1C_1C_2\acute{V}_2$ – $\acute{V}_1C_1C_2V_2$: $C_1C_2$: voiced, $C_1$: non-lab. stop, $C_2$: lab.

| | **g** | **d** |
|---|---|---|
| stop **b** | ***gb** | ***db** |
| fricative **v** | ***gv** | a**dv**ánce$_{69}$ |
| nasal **m** | do**gm**átic$_{22}$ | a**dm**íre$_{39}$ |
| approx. **w** | i**gw**ána$_{6}$ | ***dw** |
| V́ | ci**g**ár$_{244}$ | a**d**ápt$_{4700}$ |

| | **g** | **d** |
|---|---|---|
| stop **b** | rú**gb**y$_{1}$ | ***db** |
| fricative **v** | ***gv** | á**dv**erb$_{31}$ |
| nasal **m** | stí**gm**a$_{56}$ | á**dm**iral$_{10}$ |
| approx. **w** | (wí**gw**am$_{1}$) | bú**dw**ir$_{2}$ |
| V | é**g**o$_{1015}$ | é**d**it$_{1385}$ |

The problem concerns the lack of **gv** clusters in English. Provided that the perceptual hierarchy scale for the contrast of **g** – **d** is what is indicated in (20), Phonotactic Closedness is not satisfied (even if we collapse the two charts into one, thus disregarding the stress difference). The frequencies, again, may well motivate the reordering of the markedness of the two voiced stops into **g** > **d**,[23] if we do this, as well as collapse the two relevant charts, we get this:

(21) English $V_1C_1C_2V_2$: $C_1C_2$: voiced, $C_1$: non-labial stop, $C_2$: labial, revised

| | **d** | **g** |
|---|---|---|
| stop **b** | ***db** | **gb**$_{1}$ |
| fricative **v** | **dv**$_{100}$ | ***gv** |
| nasal **m** | **dm**$_{49}$ | **gm**$_{78}$ |
| approx. **w** | **dw**$_{2}$ | **gw**$_{7}$ |
| V́ | **d**$_{6085}$ | **g**$_{1259}$ |

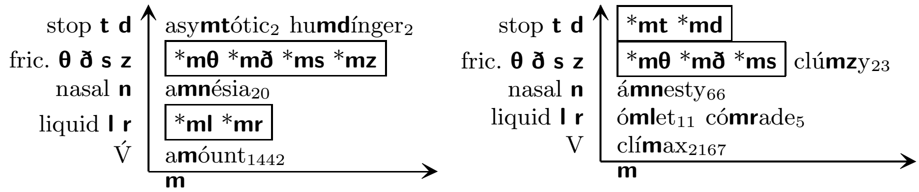
This time the lack of **db** clusters raises problems for Closedness. However, the only item with **gb** is *rugby*, which comes from the corresponding town's name, and according to many authors, proper names have a separate phonotactics, which is usually more lenient than that of non-proper names. If we remove this item from the current phonotactic layer (that of non-proper names), then there will be no gap, and so Closedness is not violated. It is obvious then that the dimension of separate phonotactic layers need also be considered; how it is to be done is, again, a matter of future research.[24] The relative markedness of **dw** and **gw** is also conspicuous in (21). Possibly all of these words could be considered non-core vocabulary (cf. *boudoir* **'bu:dwɑ:**, *iguana* **ɪ'gwɑ:nə**, *wigwam* **'wɪgwæm**). What is also curious is that the number of **gw** clusters raises provided they follow the velar nasal **ŋ** (e.g.,

**t** is *less* marked than **k**) is definitely worth further investigating. Especially, it would be instructive to see what role stress plays in the perception of the place contrast of stops.

[23] Cf. V**d**V (6085 items) vs. V**g**V (1259 items), for example.

[24] On the phonotactic layering of the lexicon, see, among others, Itô & Mester (1995) and Rebrus & Trón (2002 : 36–59). The problem, for example, concerns the issue of what counts as 'native', 'non-proper name', etc. in the lexicon of a language. It seems that *token* frequency also plays a role here: even if **gb** is perceptually (and hence phonotactically) a marked cluster, which is also indicated by its low type frequency, the fact that it is frequently used makes it seem unmarked. Cf. for example the Hungarian cluster **ɲv** which is marked in word-final position, but since it occurs in the word *könyv* **køɲv** 'book', speakers will not consider it special or 'odd-sounding'.

(22) English $V_1C_1C_2\acute{V}_2$ – $\acute{V}_1C_1C_2V_2$: $C_1$: **m**, $C_2$: coronal




*anguish, distinguish, language, linguist, penguin* etc.). Phonetic research is needed here to confirm the special status of **dw/gw** clusters.[25]

The last dimension we consider in this section is the occurrence of **m** before coronals (and vowels). The chart (22) confirms what has been indicated about the relationship of stress and English phonotactics above: no gaps occur unless the dimension of the preceding/following vowel is *not* considered:

Another indication that (22) suggests is that **mn** clusters are "better" (at least more frequent) than any other **m**+coronal clusters. Liquids have been established as relatively good contexts for stops, but this is apparently not the case for **m**: their number is fairly low (and they only occur if the vowel before **m** is stressed). Non-homorganic stops are basically impossible after **m**. If a coronal follows labial **m**, it is preferably either **n** or **z**. These factors point towards two well-known phonological facts: nasals prefer to be homorganic with a following stop, and that obstruents following nasals prefer to be voiced. It is these two issues that we turn to next.

## 5.3 Postnasal voicing

(23) summarizes some of the most important phonological facts concerning nasals, place assimilations, and the voicing of postnasal obstruents:

(23) a. In place assimilations, in $C_1C_2$, $\mathbf{C_1}$ tends to assimilate the features of $C_2$.
b. *Nasals* are the most common targets for place assimilation (incl. static place agreement).
c. The target of nasal place assimilation is frequently restricted to *coronals*.
d. Obstruents following nasals prefer to be *voiced*.

There is abundant literature on the phonetic/functional grounding of (23a–c). On the speciality of $C_1$ in $VC_1C_2V$ from a phonetic point of view, cf. Ohala (1990) as well as Kohler (1990), who argue that the place cues of (non-retroflex)[26] consonants in CV positions are more robust than in VC, hence the stability of $C_2$: the place of $C_1$ is not salient before another consonant. The most often cited phonetic reason

[25] The relatively high frequency of **dv** clusters (as opposed to **dw**) is also somewhat surprising. It must be noted though that most of them contain the (obsolete) Latinate prefix *ad-*; cf. *advance, advocate, adverb*, etc.

[26] As it was mentioned earlier, retroflex consonants are better cued in VC transitions. Pre-aspirated consonants are also more salient after a vowel than before it; on this, see Steriade (1997, 1999).

why nasals require a homorganic stop after them is that even though nasals as a group are easily distinguishable from other sounds, yet the identification of the nasals from each other is difficult, as their place is weakly cued in themselves—they need stops so that their place may be more salient (on this type of approach to nasal place sharing, cf. Myers 1997 and Maddieson 1984 : 70f).[27] Accordingly, as Hayes & Steriade (2003 : 29) argue, the scale of the perception difficulty of the place of $C_1$ in $C_1C_2$ is: (strident) fricative < stop < nasal. Among the places, it is velars that are the most and coronals that are the least salient in CC clusters (as the first consonants): velars < labials < coronals[28]—a possible perception-based reason why they are easily confusable and thus why they are the usual targets for place assimilation.

The typology of postnasal obstruents shows that they prefer to be voiced. (24) exhibits historical as well as synchronic examples of the phenomenon:

(24) Postnasal voicing

a. Middle Greek {**p t k**} → {**b d g**} / [+nasal]__:
Ancient Greek *pente* > MGreek *pende* 'five' (Cser 2001 : 58)

b. Hungarian {**p t ʧ k**} → {**b d ɟ g**} / [+nasal]__:
Uralic **kumpa* > *kumba* (> current H. *hab* 'foam'); Finno-Ugric **kunta* > *kunda* (> current H. *had* 'army'); F-U. **lonca* (> current H. *lágy* **laːɟ** 'soft'); F-U. **tuŋke* > **tuŋge** (> current H. *dug* 'stick') (Cser 2001 : 59)[29]

c. Permian {**p t c ʧ k**} → {**b d ɟ ʤ g**} / [+nasal]__:
Uralic **kumpa* > Votyak *gi̯bed* 'peat', Zyrian *gi̯be̯d* 'clump'; U. **lamte* > Votyak/Zyrian **lud* 'field'; U. **kunce* > Vot. *kiz*, Zyr. *kuʒ́* 'urine'; F-U. **wanca* > Vot. *vi̯ǯ*, Zyr. *vuǯ* 'step over'; Finnish-Permian **waŋka* > Votj./Zyr. *vug* 'handle' (Cser *ibid.*)

d. NW New Indo-Aryan {v'less stops & affricates} → {voiced} / [+nasal]__:
Sanskrit *paɲʧa* > Sindhi *paɲʤa* 'five', *aŋka* > *aŋgu* 'sign' (Cser 2001 : 60)

e. Sogdian {**p t t͡s ʧ k**} → {**b d d͡z ʤ g**} / [+nasal]__:
Old Iranian *panca* > Sogd. *panǯ* 'five' (Cser *ibid.*)

f. Wembawamba: **takʌ** 'to hit', **mɪlpʌ** 'to twist' vs. **jandɪn** 'me', **panbʌr** 'shovel' (Hayes & Stivers 2000 : 1)

g. Ecuadorian Quechua: **saʧa** 'jungle' – **saʧa-pi** 'in jungle'; **atam** 'frog' – **atam-bi** 'in frog' (Hayes 1996 : 6)

In Wembawamba (24f), obstruents can only be voiced after nasals (they occur voiceless after liquids, for example). The last example (24g) shows that in Ecuador-

[27] Browman and Goldstein's (1990) paper gives an articulatory account of nasal place assimilation.

[28] Cf. Jun (1995).

[29] Postnasal voicing is the most important source of the voiced stops in Hungarian. The nasals were lost afterwards. See Cser (*ibid.*).

ian Quechua, the locative suffix *-pi* changes to *-bi* if it follows a nasal—postnasal voicing is therefore a truly phonological process in that language. Hayes (1996) as well as Hayes & Stivers (2000) place postnasal voicing into the phonetic perspective by claiming that there are two mechanisms responsible for the process: one is the nasal airflow ("nasal leak") during the production of nasals, the other is the vertical motion of the velum when it closes just at the onset of the obstruent. Both of these mechanisms create an environment in which obstruents are automatically voiced.

In British English RP, as well as many other English dialects, including General American, however, postnasal voicing is not an obligatory phonological process. Examples with postnasal *voiceless* obstruents, such as *antic, bumper, ankle, lance, emphasis,* etc., readily come to one's mind. As a first approximation, we may say that in English postnasal voiced stops are actually more marked than nasal–voiceless stop clusters.[30] However, the situation is more complex than this. Before scrutinizing the case of nasal–stop clusters in English, let us first consider the typology of CC clusters in general; (25) displays some of the implicational universals of such clusters:

(25) Implicational universals for $C_1C_2$ clusters ($C_2$ = stop)
(cf. Rebrus & Trón 2004 : 146f)

a. the presence of voiceless stops before a nasal indicates that of a voiced one (e.g., **nt** > **nd**)
b. the presence of a nonhomorganic nasal–stop sequence indicates that of a homorganic one (e.g., **mt** > **mp**, **nt**)
c. the presence of a liquid + stop indicates that of a nasal + stop (e.g., **rt**, **lt** > **nt**)
d. the presence of a fricative + stop indicates that of a nasal + stop (e.g., **st** > **nt**)

Examples for languages with respect to what non-word-initial CC clusters they allow for are offered in (26). The chart only concentrates on CC clusters whose second element is a stop (the examples all show a coronal stop).[31]

In English, as we have seen, **mt** is rare (hence the bracketing of + for this cluster). Yapese does not permit CC clusters, it is considered to be a "codaless" language (it does have single word-final consonants, though). The implicational

[30] Of course, it is not the fact that there are voiceless stops *at all* after nasals that causes difficulties: postnasal voicing as a phonetic fact is observable in all languages; nevertheless, not all of them enforce it to phonologize it (cf. Hayes 1996 : 6). The existence of voiceless stops after nasals is thus marked (but not impossible) from the viewpoint of postnasal voicing. A more serious problem for the postnasal voicing approach rather concerns languages which apparently only allow for postnasal voiceless stops but lack voiced ones after nasals.

[31] The table is based on Pigott (1999) and Rebrus & Trón (2004 : 147). "Engl." = English, "Ital." = Italian, "Diola F." = Diola Fogny, "Japan." = Japanese. "nh-N" is meant to represent a nasal which is *not* homorganic with the following stop. No distinction has been made here as to the position of the clusters (word-internal vs. word-final). In some cases this overgeneralizes the picture; in Diola Fogny, for instance, liquid-initial clusters do not occur word-finally, only the nasal-initial ones (for details, see Pigott *ibid.*:147).

(26)

| | Engl. | Ital. | Diola F. | Manam | Japan. | Yapese | Ojibwa | Lardil |
|---|---|---|---|---|---|---|---|---|
| nh-N+stop (**mt**) | (+) | | | | | | | |
| fric.+stop (**st**) | + | + | | | | | + | |
| liquid+stop (**lt**) | + | + | + | | | | ! | + |
| N+vless stop (**nt**) | + | + | + | + | | | + | + |
| N+voiced stop (**nd**) | + | + | + | + | + | | + | ! |
| V+stop | + | + | + | + | + | + | + | + |

universals (25) are all exemplified by the languages in (26). Apparently, there are, however, two problematic languages: Ojibwa and Lardil (consider the gaps with an exclamation mark in the table). The difficulty presented by the Ojibwa case is only problematic if it is presumed that clusters of the **st** type are actually more marked than those of the **lt** type: in this case the more marked element would not imply the occurrence of the less marked — an apparent violation of Phonotactic Closedness. Notice, however, that no implication has been established between the existence (or lack) of **st** and **lt** clusters in (25). This means that the gap (the lack of liquid–stop clusters) in the Ojibwa case is not problematic after all: **st** can exist with or without **lt** in a language, and *vice versa* (cf. Diola Fogny vs. Ojibwa). What *is* an important requirement is that the existence of **lt** does imply the occurrence of **nt** (while the reverse does not stand; cf. English, Italian, Diola Fogny vs. Manam, Ojibwa).

The absolute ban on postnasal voiced stops in Lardil is nevertheless a more painful case with respect to postnasal voicing. This problem takes us back to the English case, which is similar to the one presented by Lardil in some ways. It turns out that we cannot treat all nasal–stop clusters the same way: the position where they occur in the word is highly relevant, as the distribution of the clusters is different if the environments are also different. Let us therefore concentrate on CC (including nasal–stop) clusters in English monomorphemic words in two positions: intervocalic (27), and word-final (28).[32]

The perceptual difficulty scale of both the contexts ($y$-axis) and the stops (their place contrast) ($x$-axis) is indicated in (27) and (28), just as it was done in the previous sections. It must be noted here, though, that the difficulty scale is only assumed here on the basis of the actual distribution of the segments in question, a fuller investigation into the perception of stops after consonants based on experimental evidence is needed to ground these scales phonetically.[33]

[32] The empty boxes, as usual, indicate clusters that do not occur. Notice that, unlike in earlier charts, the $x$-axis now has $C_2$, while the $y$-axis $C_1$. **r**+C clusters are, of course, only valid for rhotic dialects (like GA); diphthongs are assumed to be transcribed with vowel symbols, thus, for example, the word *fight* **faɪt** does *not* contain a **jt** cluster. Stops/affricates as the first elements in the CC clusters have been disregarded in the chart; partially at least, they have been tackled in **§5.2**. For more comprehensive lists of English non-initial CC clusters, see Kiss (2001); Szigetvári (2004).

[33] For this reason, for example, the contexts **ʃ/ʒ___**, **j/w___**, **θ/ð___** are presumed to present the same level of difficulty for the perception of the stops, as none of the stops occur in them; hence the difficulty scale is assumed to be {**ʃ/ʒ___**, **j/w___**, **θ/ð___**} > nh-N___> **f/v___**> {**s/z___**, **l___**, **r___**} > N___> V___. No implication between fricatives and liquids can be

(27) English monomorphemic VCCV clusters

| | t | k | p |
|---|---|---|---|
| **ʃ** | | | |
| **j/w** | | | |
| **θ** | | | |
| nh-N | | | |
| **f** | after | | |
| **s** | custom | fiscal | hospital |
| **l** | alter | alcohol | pulpit |
| **r** | forty | turkey | harpoon |
| N | winter | wrinkle | temper |
| V | better | accustom | lepel |

| | d | b | g |
|---|---|---|---|
| **ʒ** | | | |
| **j/w** | | | |
| **ð** | | | |
| nh-N | | | |
| **v** | | | |
| **z** | (asdic$_1$) | (husband$_2$) | |
| **l** | shoulder | album | vulgar |
| **r** | cardigan | turbine | forget |
| N | bandit | gambit | hunger |
| V | ready | ruby | ego |

(28) English monomorphemic VCC# clusters

| | t | k | p |
|---|---|---|---|
| **ʃ** | | | |
| **j/w** | | | |
| **θ** | | | |
| nh-N | (dreamt$_1$) | | |
| **f** | draft | | |
| **s** | mist | ask | wasp |
| **l** | melt | sulk | pulp |
| **r** | part | bark | carp |
| N | vent | think | jump |
| V | cut | make | sip |

| | d | b | g |
|---|---|---|---|
| **ʒ** | | | |
| **j/w** | | | |
| **ð** | | | |
| nh-N | | | |
| **v** | | | |
| **z** | | | |
| **l** | hold | (bulb$_3$) | |
| **r** | guard | disturb | morgue |
| N | find | ! | ! |
| V | bed | stab | fig |

At various points in §**5.2**, we saw that the markedness of a segment might vary if it is placed in different dimensions. The matters regarding postnasal stops are no exceptions either. As the charts in (27) and (28) show, the dimensions "word-internal" vs. "word-final" must also be considered in the distribution of CC clusters as *separate dimensions* that affect markedness. The word-final position—as opposed to the word-internal—appears to be highly restrictive for voiced stops in English: only **d**, but not **b** and **g**, may occur after nasals in that position, for example. Such severe restrictions as these are not observable in the intervocalic (crucially: *prevocalic*) position.

As Rebrus & Trón (2004) observe, an important effect of the independence of markedness scales is that they may sometimes stand in *conflict*. Let us for example consider the voicing contrast of stops in three positions: (i) prevocalically, (ii) after a homorganic nasal, and (iii) word-finally. Using the contrast of **t** – **d** as a hypothetical example, the following markedness scales can be set up in these three environments:

(29) a. Scale 1: **da** > **ta** b. Scale 2: **nt** > **nd** c. Scale 3: **d**# > **t**

established (see the Ojibwa case above), and so, it is *assumed* that there is no difference between them perceptually (as far as the cuing of a following stop is concerned).

All three scales can be grounded phonetically, as we mentioned.[34] An important consequence of the scales in (29) is that Scale 2 (the postnasal context, the position on the *left* of the stop) stands in conflict with both Scale 1 (prevocalic position) and Scale 3 (word-final position), two positions on the *right* of the stop. **d** is *marked* before a vowel (**da**), but it is *unmarked* after a nasal (**nd**). Similarly, **d** is *marked* word-finally (**d**#), but, again, it is *unmarked* postnasally. It can happen that a stop stands in postnasal position *as well as* (i) prevocalically (**nta**/**nda**) or (ii) word-finally (**nt**#/**nd**#). Which markedness scale (the one for the context on the left—the postnasal position, or the one on the right—prevocalic/word-final) "wins" over the other is a language-specific choice.

Let us first consider the case of prevocalic nasal–stop clusters (**nta**/**nda**). If it is the scale for the postnasal context (Scale 2) that wins over the prevocalic scale (Scale 1), then we have a system that will lack ***nta**, but will contain **nda**. If however Scale 1 outweighs Scale 2, the language in question will have **nta**, but no ***nda**. The following systems are thus predicted for the voicing contrast of stops before vowels:

(30) Prevocalic sequences: **ta**, **da**, **nta**, **nda**

a. System 1: voicing is maintained in all sequences: **ta**, **da**, **nta**, **nda** (e.g., Hungarian)
b. System 2: only voiceless stops occur: **ta**, ***da**, **nta**, ***nda** (e.g., Lardil)
c. System 3: voicing neutralization prevocalically, postnasal voicing: **ta**, ***da**, ***nta**, **nda** (e.g., Wembawamba)
d. System 4: voicing contrast prevocalically, postnasal voicing: **ta**, **da**, ***nta**, **nda** (e.g., Japanese)
e. System 5: voicing contrast prevocalically, no postnasal voicing: **ta**, **da**, **nta**, ***nda**

In systems like (30a), even the marked sequences are allowed to occur. In languages like System 2, there is no voicing contrast for stops whatsoever. This is the consequence of the priority of the post-stop vowel (cf. Scale 1 in (29)): the conflict between the influence of the left environment (the nasal) vs. the right environment (the vowel) is won by the latter. Languages that behave like System 3 show that the scales in (29) are independent of each other, it is only when they necessarily come together—in the case of prevocalic nasal–stop clusters—that they stand in conflict. In System 3, there is neutralization on Scale 1 (**ta**, ***da**), while the conflict between Scale 1 and 2 are resolved in favour of Scale 2, where postnasal voicing is more important than prevocalic voicelessness: ***nta**, **nda**. System 4 languages also display the independence as well as the possible conflict of the scales in (29).

[34] For example, Hayes (1996) shows that voiceless stops are less difficult to produce than voiced ones before vowels; especially in English, voiceless aspirated (fortis) stops are also easier to perceive than voiceless unaspirated (lenis) stops prevocalically. In word-final position, important cues for the perception of voicing are missing (see (7)), and thus a voiced stop is marked in that position in this respect. The phonetic basis of postnasal voicing has been tackled above.

A language may allow for marked sequences (i.e., **da** besides **ta**), while it resolves the N__ vs. __V conflict in favour of postnasal voicing. There is another system predicted (System 5): in it there is voicing contrast before a vowel, and the conflict between N__ vs. __V in the case of prevocalic nasal–stop clusters is won in favour of the right-hand side environment. At this point I have not been able to come up with languages exhibiting this option. Crucially, no systems are predicted like **da**, ***ta**, where the less marked sequence is missing, while the more marked exists—this would violate Phonotactic Closedness. Notice that Phonotactic Closedness is not testable in cases of conflicting scales, since the hierarchies are undecided, the choice between them is arbitrary.

The case of the voicing contrast of word-final stops vs. (homorganic) postnasal stops is similar to that of prevocalic stops vs. (homorganic) postnasal stops. If we consider word-final nasal–stop clusters, it is now Scale 3 in (29) that is in conflict with Scale 2. However, the markedness hierarchy of the word-final position seems to be always winning over the postnasal hierarchy: the word-final position is a context where it is difficult to keep up voicing (as well as place) contrast. Accordingly, it is predicted that no system should occur in which there are word-final **nd**# clusters but no ***nt**# sequences (this is what postnasal voicing would suggest). The following cases are predicted thus:

(31) Word-final sequences: **t**#, **d**#, **nt**#, **nd**#

a. System 1: voicing is maintained in all sequences: **t**#, **d**#, **nt**#, **nd**# (e.g., Hungarian)

b. System 2: only voiceless stops occur: **t**#, ***d**#, **nt**#, ***nd**# (e.g., Polish)

c. System 3: voicing contrast for single stops, no postnasal voicing: **t**#, **d**#, **nt**#, ***nd**# (e.g., English)

English (System 3 in (31)) displays the independence of Scale 2 and 3 (29): there can be voicing contrast for stops word-finally, but neutralization into **t** after a nasal.[35] But there are no systems with ***t**#, **d**#, ***nt**#, **nd**#, in accordance with Phonotactic Closedness.

The state of affairs concerning the distribution of word-final nasal–stop clusters in English seems to be even more complex, however. For example, it is not true that *all* voiced stops are missing after nasals word-finally. It is only the noncoronals that are forbidden there. Therefore, another dimension must also be considered—that of place of articulation. As the frequency of the cases in English also indicate, the markedness scales of stops are (32a) (NC#) and (32b) (C#):

(32) a.

| | | |
|---|---|---|
| **nd**#$_{285}$ | ***ŋg**# | ***mb**# |
| **nt**$_{658}$# | **ŋk**#$_{70}$ | **mp**#$_{43}$ |
| cor. | vel. | lab. |

b.

| | | |
|---|---|---|
| **d**#$_{1259}$ | **g**#$_{159}$ | **b**#$_{106}$ |
| **t**#$_{3251}$ | **k**#$_{1416}$ | **p**#$_{406}$ |
| cor. | vel. | lab. |

[35] The situation may well be more intricate than this for English: it is traditionally claimed that English does not contrast obstruents word-finally: they are normally unreleased, voiceless and unaspirated. It is actually the previous vowel (its length/quality) onto which the contrast between fortis ("voiceless") and lenis ("voiced") obstruents is transferred, as it were; thus *beat* **bit** vs. *bead* **biːd̥** ≈ **biːt**; *pint* **paɪnt** vs. *find* **faɪːnd̥** ≈ **faɪːnt**.

The perceptual grounding of these cases definitely needs further research; the expectation is that "heavy" clusters (such as voiced labial **mb**) are perceptually less salient word-finally, than the unmarked coronal clusters.

If we consider yet another dimension, the markedness of word-final nasal–stop clusters alters again. This dimension is the quantity and quality of the vowel before the cluster. Some of the most important facts for English in connection with this dimension are summed up below:

(33) a. V**nt**#: V can be of almost any quality (except **ʊ**); V**nd**#: the V can be any vowel (except **ʊ**; **ɪ** only occurs in *wind*)

b. V**mp**#, V**ŋk**#: most cases occur with the low vowels **æ**, **ʌ**; **ɪ**, **ɒ**, **e** are rare; there are no such clusters with **ə** and **ʊ**

c. long/tense vowels are marked before noncoronal clusters: **V:nt**# (50 items), **V:nd**# (61 items), but: ***V:ŋk**#, ***V:mp**#, ***V:ŋg**#, ***V:mb**#

d. if the V before the word-final nasal–stop cluster is long/tense, it is usually non-high:

| | **nt** | **ntʃ** | **ŋk** | **mp** | | **nd** | **ndʒ** | **ŋg** | **mb** |
|---|---|---|---|---|---|---|---|---|---|
| **i:/u:** | □ | □ | □ | □ | **i:/u:** | fiend$_1$/wound$_1$ | □ | □ | □ |
| **əʊ** | don't$_3$ | □ | □ | □ | **əʊ** | □ | □ | □ | □ |
| **ɔ:** | flaunt$_5$ | launch$_5$ | □ | □ | **ɔ:** | laundry$_4$ | □ | □ | □ |
| **ɔɪ** | point$_6$ | □ | □ | □ | **ɔɪ** | □ | □ | □ | □ |
| **eɪ** | paint$_{13}$ | □ | □ | □ | **eɪ** | □ | change$_9$ | □ | □ |
| **ɑ:** | slant$_6$ | branch$_5$ | □ | □ | **ɑ:** | command$_9$ | □ | □ | □ |
| **aɪ** | pint$_1$ | □ | □ | □ | **aɪ** | find$_9$ | □ | □ | □ |
| **aʊ** | count$_6$ | □ | □ | □ | **aʊ** | sound$_{28}$ | □ | □ | □ |

The exact phonetic grounding of the relationship between the pre-cluster vowel and the distribution of the nasal–stop cluster is still to be clarified. Nevertheless, it seems that the marked clusters cannot occur with either reduced vowels (like **ə**) or long/tense vowels. This is only possible if the place of the cluster is the unmarked coronal. Also, only non-high (mostly low) long vowels can occur with nasal–stop clusters, but, again, only provided that the cluster is coronal. Low vowels are also preferred when the cluster is noncoronal (*camp, lamp, trunk, rank* etc.). Further investigation is needed here, but it seems that clusters which are weakly cued in word-final position need vowels that are the most salient (open)—the low vowels—in order to enhance their own salience.[36]

## 6 Conclusion

The paper has presented that the phonotactics of languages are gradual and that they can be approached in a multidimensional manner. It turned out that the

[36] The findings of Burzio (2002) support these views. He suggests that both a reduced vowel as well as a long vowel result in the loss of perception cues (especially burst cues) for stops; hence after them, neutralization is predicted to the unmarked place, the coronals.

distribution of segments can only be satisfactorily explained if perceptual difficulty scales are set up which delimit what segment combinations can occur where. Most importantly, the outermost (perceptually most marked) elements on these difficulty scales define what other elements may occur in the phonotactic space: only those that are perceptually less marked than the others. Phonotactics in this way is gradual as the scales predict that if a marked element exists in the language, the less marked *must also exist*, the phonotactic space must be filled (closed) by existing forms and surrounded by the marked elements at the outskirts. These scales are thus define the limits of possible languages. As a consequence, the notions "accidental gap" and "exception" are not relevant in this approach: rare/marked clusters (regarded to be exceptional in traditional frameworks) are predicted to be part of the phonotactic space as much as frequent/unmarked clusters, what is more: they are predicted to occur in the "right" areas of the phonotactic space (outskirts vs. around the origin). This is necessarily a different approach from that of other (representational) models, which are either too restrictive (they are *undergenerating* and mark those elements as exceptions which it cannot account for but are nevertheless grammatical) or too "liberal" (they are thus *overgenerating*, and treat the ungrammatical/non-existing forms as accidental gaps). These models thus cannot account for graduality.

Phonotactics is necessarily multidimensional, because the perceptual markedness of a segment is dependent on the contrast in question as well as the context it occurs: on the one hand, a segment which is regarded typologically marked can in fact be unmarked if it does not participate in a contrast; on the other hand, a contrast which is marked in a given position may well be unmarked in another. Therefore, several dimensions must be considered to explain segmental distributions. These ideas were illustrated on a (rather limited) set of data from Hungarian and English non-initial CC clusters in monomorphemic words. The analysis relied only on functional arguments, such as Phonotactic Closedness and perceptual/articulatory grounding. Importantly, no representational constructs (such as the syllable) have been employed in explaining the various distributional phenomena. In fact, it has turned out that the syllable is not necessary in the explanations, linear statements sufficed: what counts is the immediate vicinity of the segments (e.g., whether it is followed by a vowel or not), larger prosodic constituents have not been made use of.

The most important element in the approach of the paper was the functional (phonetic) grounding. However, in many cases, the phonetic grounding was admittedly only assumed, and it is this area, among others, on which research must necessarily focus in the future.

## REFERENCES

Browman, Catherine and Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In: Kingston & Beckman (1990 : 341–376).

Burzio, Luigi. 2002. Phonetics and phonology of English stress and vowel reduction. Paper presented at the Conference on English Phonology, Toulouse, June 26–28, 2002.

Chomky, Noam and Morris Halle. 1968. The sound pattern of English. New York: Harper & Row.

Côté, Marie-Hélène. 2000. Consonant cluster phonotactics: a perceptual approach. Doctoral dissertation, MIT, Cambridge, MA.

Cser, András. 2001. A zörejhangok gyengülési és erősödési folyamatainak tipológiája és modellezése. Doctoral dissertation, Elméleti nyelvészet doktori program, ELTE/MTA, Budapest.

Cyran, Eugeniusz. 2003. Complexity scales and licensing strength in phonology. Lublin: Wydawnictwo Kul.

Diver, William. 1979. Phonology as human behavior. In: Doris Aaronson and Robert Rieber (eds.). Psycholinguistic research: implications and applications. Hillsdale: Lawrence Erlbaum Associates. 161–182.

Flemming, Edward. 1996. Evidence for constraints on contrast: the dispersion theory of contrast. UCLA Working Papers in Phonology 1 : 86–106.

Flemming, Edward. 2001. Contrast and perceptual distinctiveness. Ms. To appear in: Bruce Hayes, Richard Kirchner and Donca Steriade (eds.) *The phonetic bases of markedness.* Cambridge: Cambridge University Press.

Goldsmith, John A. (ed.). 1995. The handbook of phonological theory. Cambridge, MA & Oxford: Blackwell.

Harris, John and Geoff Lindsey. 1995. The elements of phonological representation. In: Jacques Durand and Francis Katamba (eds.). Frontiers of phonology: atoms, structures, derivations. Harlow: Longman. 34–79.

Hayes, Bruce. 1996. Phonetically driven phonology: the role of Optimality Theory and Inductive Grounding. Ms. Also appeared in: Michael Darnell, Edith Moravscik, Michael Noonan, Frederick Newmeyer, and Kathleen Wheatly (eds.) *Functionalism and formalism in linguistics, Volume I: General papers.* Amsterdam: John Benjamins. 243–285.

Hayes, Bruce and Donca Steriade. 2003. The phonetic bases of phonological markedness. Manuscript.

Hayes, Bruce and Tanya Stivers. 2000. Postnasal voicing. Ms. University of California, Los Angeles.

Hooper, Joan B. 1976. An introduction to natural generative phonology. New York: Academic Press.

Hume, Elizabeth and Keith Johnson. 2001. A model of interplay between speech perception and phonology. In: Elizabeth Hume and Keith Johnson (eds.). The role of speech perception in phonology. New York: Academic Press. 3–26, draft version: http://www.ling.ohio-state.edu/~kjohnson/Hume_Johnson.pdf.

Itô, Junko. 1986. Syllable theory in prosodic phonology. Doctoral dissertation, University of Massachusetts, amherst. (Published also in the *Outstanding Dissertations in Linguistics* series, Garland, New York, 1988.

Itô, Junko and Armin R. Mester. 1995. Japanese phonology. In: Goldsmith (1995 : 817–838).

Jakobson, Roman, Gunnar Fant and Morris Halle. 1952. Preliminaries to speech analysis: the distinctive features and their correlates. Cambridge, MA: MIT Press.

Jun, Jongho. 1995. Percpetual and articulatory factors in place assimilation: an Optimality Theoretic approach. Doctoral dissertation, UCLA, Los Angeles.

Kingston, John and Mary E. Beckman (eds.). 1990. Papers in laboratory phonology I: Between the grammar and physics of speech. Cambridge: Cambridge University Press.

Kiss, Zoltán. 2001. Distributional effects in English and Hungarian non-initial CC clusters. M.A. thesis, Eötvös Loránd University, ELTE, Budapest.

Kohler, K. 1990. Segmental reduction in connected speech in German: phonological facts and phonetic explanation. In: W. Hardcastle and A. Marchal (eds.). Speech production and speech modelling. Dordrecht: Kluwer. 69–92.

Lass, Roger. 1984. Phonology. An introduction to basic concepts. Cambridge: Cambridge University Press.

Liljencrants, Johan and Björn Lindblom. 1972. Numerical simulation and vowel quality systems: the role of perceptual contrast. Language 48 : 839–862.

Maddieson, Ian. 1984. Patterns of sounds. Cambridge: Cambridge University Press.

Mohanan, K. P. 1995. The organization of grammar. In: Goldsmith (1995 : 24–69).

Myers, Scott. 1997. Expressing phonetic naturalness in phonology. In: Iggy Rocca (ed.). Derivations and constraints in phonology. Oxford: Clarendon Press. 125–152.

Ohala, John J. 1990. The phonetics and phonology of aspects of assimilation. In: Kingston & Beckman (1990 : 258–275).

Pigott, Glyne L. 1999. At the right edge of words. The Linguistic Review 16 : 143–185.

Postal, P. M. 1968. Aspects of phonological theory. New York: Harper & Row.

Rebrus, Péter and Viktor Trón. 2002. A fonotaktikai általánosításokról. (Kísérlet a magyar mássalhangzó-kapcsolatok nem-reprezentációs leírására). In: Márta Maleczki (ed.). A mai magyar nyelv leírásának újabb módszerei V. Szeged, 2001. szeptember 27–29. Szeged: Szegedi Tudományegyetem, Általános Nyelvészeti Tanszék, Magyar Nyelvészeti Tanszék. 17–63.

Rebrus, Péter and Viktor Trón. 2004. Fonetikai motiváció a fonológiai mintázatokban. In: László Hunyadi (ed.). Kísérleti fonetika labóratóriumi gyakorlatban. Debrecen: Debreceni Egyetem Kossuth Egyetemi Kiadója. 139–164.

Steriade, Donca. 1997. Phonetics in phonology: the case of laryngeal neutralization. Manuscript. UCLA.

Steriade, Donca. 1999. Alternatives to syllable-based accounts of consonantal phonotactics. Ms. Also appeared in: O. Fujimora, B. Joseph and B. Palek (eds.) *Proceedings of the 1998 Linguistics and Phonetics Conference*. Prague: The Karolinum Press. 205–242.

Szigetvári, Péter. 2004. Branching onsets and syncope in English. Ms. Eötvös Loránd University, Budapest. To be released in the Proceedings of the Conference on English Phonology (Université de Toulouse, 26–28 June 2002).

Zipf, George Kingsley. 1949. Human behavior and the Principle of Least Effort. Cambridge, MA: Addison-Wesley Press.