Graduality and closedness in consonantal phonotactics
—a perceptually grounded approach

Zoltán Kiss

The paper suggests that the phonotactics of languages often displays gradual, non-categorical patterns, in the sense that not all possible combinations are grammatical along the various dimensions of segmental contrast. The graduality of phonotactics is a property that will be shown to be better explained by models that employ perception-based functional principles, such as perception cue licensing according to relative context, and phonotactic closedness. The phonotactic space of a language will be claimed to be defined by hierarchical perceptual difficulty scales that linearly order the various syntagmatic relations a segment can enter into. All existing items (marked as well as unmarked — in the perceptual as well as statistical sense) in the language are accounted for in this model, what is more, their place in the phonotactic space is also predicted. In this model, thus, the notions 'exceptional' and 'accidental gap' are meaningless. Where exactly a language draws the line between a contrast that it makes use of and those that it does not, is arbitrary, but if a perception-wise marked cluster is used in a language, then our model will predict that all the other clusters that are perceptually better cued will also occur in the language, in accordance with the principle of phonotactic closedness. Phonotactic graduality and closedness will be demonstrated in the consonantal phonotactics of languages that are often referred to as possessing 'complicated' phonotactics, Hungarian, English, and Slovak.

1. *Introduction: the graduality of 'complicated' phonotactic systems*

The precise explanation of the consonantal phonotactic patterns of languages that display what is sometimes referred to as 'complicated' syntagmatic sequences has proven to be a serious problem for traditional, representational phonological theories. These languages include, among numerous others, Hungarian, Slovak, Polish, and English. The phonotactics of these languages can be characterized by the property of *graduality*. This means that certain consonantal clusters are fairly common (the statistical aspect of which is that their type frequency is high), whereas others are rare (their type frequency is low), and other — otherwise theoretically possible — combinations do not exist at all. In this paper, I would like to approach the phonotactics of

Proceedings of ConSOLE XIII, 2005, 171-195
http://www.sole.leidenuniv.nl

these languages from a perceptually grounded point of view, which I claim to be able to avoid the difficulties that the representational models necessarily have to face. The central notion of the paper is that segmental contrast prefers to occur in environments where its perceptibility is cued the most robustly (cf. Steriade 1997; 1999). Less favourable environments in this respect will result in the contrast being less perceptible, which may even lead to its absolute loss (i.e., neutralization). Segmental contrasts can thus be placed on a perception-grounded hierarchy, one end of which represents the best context(s) for the perception of the given contrast, the other end the worst ones. The relevant consonant clusters will fill in this scale according to how robustly the contrast in question is cued. The model will predict the graduality of the distribution of these clusters: the more common clusters will occur at the better cued position of the scale, while the rare ones at the badly cued positions.

The phonotactics of a language can thus be claimed to be a space defined by these perceptual difficulty hierarchies that linearly order the various syntagmatic relations a segment can enter into. As an example, let us consider one such slice of the phonotactic space of Hungarian consonant clusters, the distribution of palatal stops before labial/dental consonants and vowels in intervocalic position in monomorphemic words.[1]

(1) The distribution of palatal stops before labial/dental Cs, and Vs (intervocalic position) (cf. Siptár & Törkenczy 2000:129)

	C_2=labial	C_2=dental
before Vs:	**c**V (*ku**tya***), **ɟ**V (*bo**gyó***)	**c**V (*ba**tyu***), **ɟ**V (*ra**gya***)
before approximants:	**cʋ** (*ko**tyv**aszt*), **ɟʋ** (*fe**gyv**er*)	**cl** (*tro**tyl**i*$_1$), **ɟl** (*ka**gyl**ó*)
before nasals:	**cm** (*tru**tym**ó*), **ɟm** (*ha**gym**a*)	*cn, *ɲ̟ɟn
before fricatives:	**cf** (*fi**tyf**iritty*$_2$)	*cs, **ɟz** (*je**gyz**ő*$_2$)
before stops:	**cp** (*pi**typ**ang*$_2$), **ɟb** (*bu**gyb**orék*$_1$)	*ct, *ɟd

Glosses: *kutya* ‘dog’, *bogyó* ‘berry’, *batyu* ‘bandle’, *ragya* ‘pockmark’, *kotyvaszt* ‘concoct’, *fegyver* ‘weapon’, *trotyli* ‘tramp’, *kagyló* ‘shell’, *trutymó* ‘suspicious substance’, *hagyma* ‘onion’, *fityfiritty* ‘imp’, *jegyző* ‘town clerk’, *pitypang* ‘dandelion’, *bugyborék* ‘bubble’.

The table tells us that not all combinations are possible. The contrast of a palatal stop is *categorically neutralized* (that is, no words occur) before dental nasals and dental stops. Also, the voiceless palatal stop does not occur before a voiceless dental fricative. Moreover, as indicated by the type frequency numbers, some clusters are *marked* frequency-wise, as they occur in but a handful of words (this we may call *partial neutralization*).

Traditional phonological theories, making use of such representational devices as distinctive features, gross natural classes, prosodic constituents (like the syllable), regardless of whether they are derivational, principles/parameters or constraint-based, face difficulties in giving a precise account of generalizations, like those tabulated in Table (1). For example, a syllable-based

---

[1] An independent phonotactic constraint does not allow two obstruents with different voicing to stand next to each other in Hungarian. That is, only voiceless – voiceless, and voiced – voiced sequences are allowed. In the first row, I also included examples of the palatal stops being followed by a vowel. The subscript numbers in some cases indicate that the cluster in question only occurs in one or two words, that is, its type frequency is low. I’m using IPA symbols (they are in bold face).

model must declare that the clusters in (1) are coda–onset clusters (they cannot be complex or branching onsets as they cannot stand word-initially, their sonority profile does not make them a suitable branching onset; they cannot be complex codas either, because they do not occur word-finally, or before other consonants). If it is only Prosodic Licensing that is the driving force behind the distribution of segments, than we cannot explain why some clusters are grammatical, why others are not well-formed, and why others are marginal. More concretely, if we allow for palatals in the coda, then our model will *overgenerate*, as it does not account for the lack of palatal–nasal, palatal–stop clusters (their ungrammaticality in Hungarian will have to be regarded as accidental, the non-existent clusters as *accidental gaps*).[2] If we do not allow for palatals in the coda, then the situation is reversed: our model will *undergenerate*: the existing palatal–consonant clusters will have to be treated as *exceptions*. The syllable-based models will have to resort to additional devices to account for these regularities, most of which are rather arbitrary. They include 'Syllable Contact Laws' (Vennemann 1988, Clements 1990), 'inter-constituent government' (Kaye et al. 1990, Rice 1992, Harris 1994), just to name but a few.[3] Note further that the so-called *context-independent*, or absolute *universal markedness* considerations cannot play a role in the explanation of the distributional asymmetry in (1) either. According to Maddieson (1984:32), in the UPSID database, more languages have consonants in the coronal area than in the labial one, in this 'universal' sense then labials are more marked. The problem is that in the case of (1), the clusters whose second member is a dental are actually the ungrammatical, missing ones. What is thus universally marked proves to be unmarked in Hungarian.

To sum up the discussion so far, we have seen that languages with consonant clusters of the 'complicated' kind, like Hungarian, cannot precisely be accounted for by traditional (such as syllable-based) phonological models. In the case of languages with 'simple' phonotactics,[4] those theories are more successful because a given contrast (in a given dimension of the phonotactic space) either always occurs, or never does so. Such a phonotactic space can be represented by a rectangle shape, as in (2a). Since the syllable-based models aim at a simple model, they are successful at capturing a simple (i.e., non-gradual) phonotactic space. In the case of a more complicated phonotactic space, displaying a gradual, or 'terraced' shape (2b), the simplicity-oriented model can either partially cover the gradual space (hence leave out parts, which will be treated as exceptions) or will also include areas that do not belong to the original space (those will be treated as accidental gaps); both cases are shown by the interrupted lines in (2c).[5]

---

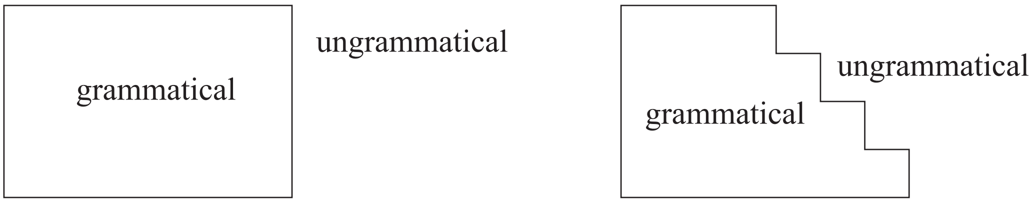
[2] The term 'accidental' here does *not* mean the accidentalness represented by the (well-known example of the) non-existent word *blick* in the English lexicon, which is a possible word phonotactically (as it contains licit sequences, occurring in existing words), it just happens to be a nonsense word in the language.

[3] For example, Törkenczy (1994:384) as well as Siptár & Törkenczy (2000) introduce the Antipalatal Condition, claiming that '[**c**, **ɟ**, **ɲ**] make an interconstituent cluster ill-formed irrespective of whether they occur in the first or the second position' (Siptár & Törkenczy *op.cit.*:137); clearly, they have to treat existing words like *pitypang* 'dandelion', *pletyka* 'rumour', etc., as exceptional, not to mention the fact that the condition against palatals must be introduced *in addition to* the general (trans)syllable-building algorithm, which obviously weakens its explanatory force.

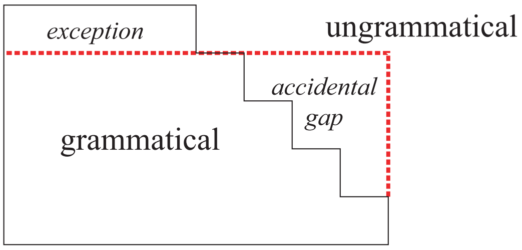
[4] They include languages where a consonant may only occur before a vowel (CV languages), or what are referred to as the 'Prince languages', where only clusters of the homorganic nasal–stop kind or geminates occur.

[5] For arguments and further illustrations of the representation of the phonotactic space as a sum of two-dimensional coordinate systems, see Rebrus & Trón (2002).

(2) a. ‘simple’ phonotactics b. ‘complicated’ phonotactics

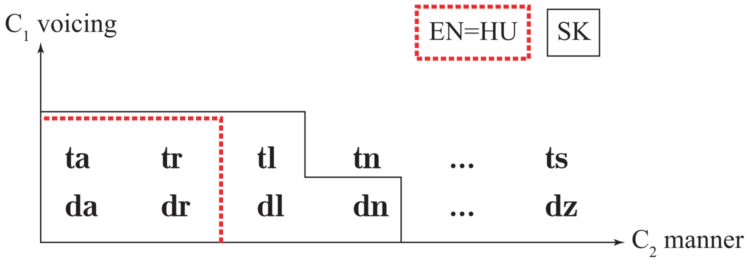


c. the simplification of ‘complicated’ phonotactics



As an example how a simplicity-oriented model may fail in some cases, let us consider the voicing contrast of coronal stops in word-initial position in three languages, English, Hungarian, and Slovak. We may represent this chunk of the consonantal phonotactic space in a two-dimensional coordinate system, the x-axis of which exhibits the manner of the second consonant, the y axis the voicing of the first consonant (cf. (3)).[6]

(3) Word-initial alveolar stops+coronal C clusters



We can see that English and Hungarian impose the same constraint on these clusters, which can be defined as an ‘antihomorganic constraint’:[7] the segments in a word-initial branching onset may not share the same place. This will allow **tr** and **dr** (supposing that the stops are alveolar and the **r** is post-alveolar), but exclude all the other alveolar–alveolar clusters. We can see that this general constraint covers a rectangle-shaped area of the phonotactic space and is successful at accounting for the grammatical vs. ungrammatical sequences, because this given dimension of the English/Hungarian phonotactic space can also be represented as having a rectangle shape (in other words: all the relevant—ungrammatical—clusters are excluded by the antihomorganic constraint). However, the same dimension in Slovak cannot be accounted for by this constraint, as even though in this language *some* of the word-initial coronal–coronal

---

[6] On the ordering of the segments along the two axes, see the discussion in the following section.

[7] Cf., among others, Harris (1990:277ff, 1994:171), and Brockhaus (1990:282).

clusters are ungrammatical, some of them are not, there *are* words with initial **tl**, **dl** and **dn**. Furthermore, the asymmetry of **dn** vs. ***tn** still awaits explanation. The terraced shape of this chunk of the Slovak phonotactic space thus cannot be covered by the antihomorganic constraint, the use of which would necessarily introduce the ‘accidental gap /exceptionality’ fallacy again.

What is thus clearly needed is a model that precisely predicts the graduality of the phonotactic space of languages with ‘complicated’ phonotactics. In such a model, both the unmarked clusters as well as the marked (rare) ones are predicted to fill the phonotactic space, and the notions ‘accidental gap’ and ‘exceptionality’ will not have to be evoked. It is this model the discussion of which I will turn to in the remaining parts of the paper.

2. *Functionalist principles in phonology: contrast, segmental markedness and perceptual robustness*

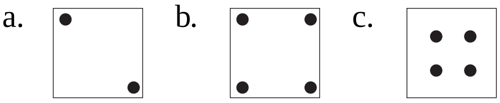
As it has long been established by functionalist accounts, phonological systems of languages are claimed to be shaped by the interaction of the following (partially conflicting) factors.[8]

(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 (2004) 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)



---

[8] Cf. Zipf (1949), Liljencrants & Lindblom (1972), Diver (1979), Flemming (1996; 2004), Rebrus & Trón (2002).

It seems clear then that in a functional theory of contrast, some weighting of potentially conflicting principles is inevitable and the weighing may well be language-specific.[9]

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*.

2.1. 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.[10] 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 counterpart 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 (2004) puts it, '[the] markedness of sounds is indeed dependent on the contrasts that they enter into.'[11]

---

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

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

[11] 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.

2.2. *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.

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

a. (i) V_1__V_2: *a**p**a, a**b**a*; (ii) V_1__son: *a**p**ra, a**b**ra*
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) #__: ***p**a, **b**a, **p**ra, **b**ra*; (ii) obstr__son: *as**p**a, as**b**a, as**p**ra, as**b**ra*
cues: voicing of closure; length of closure; length/strength of release; VOT value; F0 and F1 of V_2

c. V__#: *a**p**, a**b***; cues: voicing of closure; length of closure; length of V; length/strength of release

d. V__obstr: *a**p**sa, a**b**sa*; cues: voicing of closure; length of closure; length of V_1; F1 values of V_1

e. obstr__obstr: *as**p**ta, as**b**ta*; cues: voicing of closure; length of closure

f. obstr__#: *as**p**, as**b***; cues: voicing of closure; length of closure

g. #__obstr: ***p**sa, **b**sa*; 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 has 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 prevocalic 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).

3. *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.[12]

(8) Perceptual difficulty scale for the voicing contrast of obstruents:
{O__O , O__#, #__O} > V__O > V__# > {#__, O__R} > V__R
‘more difficult’ ←———————————— ‘less difficult’

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).

---

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

(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.

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).[13]

(12)	a. #__O (*psa–bsa*)	∗	b. #__O (*psa–bsa*)	∗
	O__# (*spa–sba*)	∗	O__# (*spa–sba*)	∗
	R__O (*apta–abta*)	✓	R__O (*apta–abta*)	✓
	R__# (*ap–ab*)	✓	R__# (*ap–ab*)	$\boxed{*}$
	__R (*pa–ba*)	✓	__R (*pa–ba*)	✓
	R__R (*apa–aba*)	✓	R__R (*apa–aba*)	✓
		T–D		T–D

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)), the intersonorant context (R__R) being the least marked environment for the voicing contrast. The tick mark

---

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

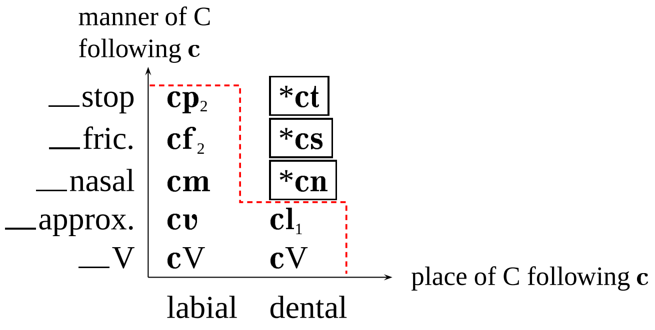
indicates that the contrast in question is attested, while the asterisk shows that the contrast is missing (thus only the unmarked occurs, the voiceless stop). (12b) violates Phonotactic Closedness because there is **p** – **b** contrast between a sonorant and an obstruent (R__O), but no contrast after a sonorant word-finally (R__#)—a better cued environment than the other. Closedness predicts that *there cannot be gaps between existing sequences* in hierarchies like (12).

Closedness is a consequence of the functional principles introduced in (4): the ordered markedness hierarchies define a phonotactic space. This phonotactic space is filled with lexical items 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'). This means that a system, in order to maximize the number of contrasts, only uses a particular context provided that other, perceptually more favourable contexts are also allowed, since in the reverse case, the last two functionalist principles in (4) would be violated.

3.1. Phonotactic Closedness in Hungarian

The phonotactic space as defined by Closedness can thus be represented as a multidimensional space, in which the markedness of sequences increases along each dimension. In a two-dimensional space, this can be illustrated by a coordinate system where the markedness of the items increases as we move from the origin. Returning to our initial example in (1), the relevant phonotactic space of the Hungarian lexicon is shown in (13), where we see the place of the voiceless palatal stop in the phonotactic space with respect to the manner and place dimensions of the following consonant.

(13)

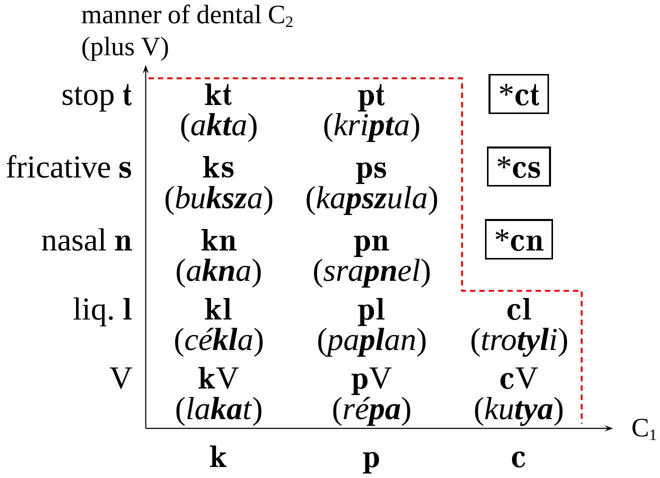


The vertical axis is the ordered perceptual hierarchy of the manner of the consonants following the voiceless stop. The horizontal axis is the ordered perceptual scale of the place of the consonants after the voiceless palatal stop. According to the perceptual properties of these positions, the ones close to the origin are the ones providing the best cues for the voiceless palatal stop. As predicted by Phonotactic Closedness and Licensing By Cue, forms like *kutya* 'dog', *batyu* 'bag', that is the ones where **c** stands before a vowel, will be more common, more frequent, hence unmarked; while those with the palatal stop followed by say an approximant will be rare, or nonexistent. Closedness declares that the phonotactic space is closed from below; there are no gaps between the existing forms along the dimensions. The phonotactic space is thus defined by the most marked items, as ‘below’ them, all the other existing forms will fill in the space until the least marked item, without interruptions, or gaps.

Let us see other dimensions in the Hungarian phonotactic space. (15) presents voiceless non-dental plus vowel and dental clusters in monomorphemic stems. The vertical axis is the ordered hierarchy of the the manner of the second consonant; the horizontal axis is the ordered scale of the place of the first consonant. The perceptual difficulty hierarchy of these contexts for the relevant place contrasts is (14a). According to Rebrus & Trón (2002), the perceptual difficulty scale for the voiceless non-coronal stops before coronal consonants is (14b).

(14) a. __stop > __fric. > __nasal > __liquid > __vowel
b. **c** > **p** > **k**

(15) VC_1V, VC_1C_2V sequences: C_1 = voiceless non-dental stop, C_2 = dental
(cf. Rebrus & Trón 2002:23)



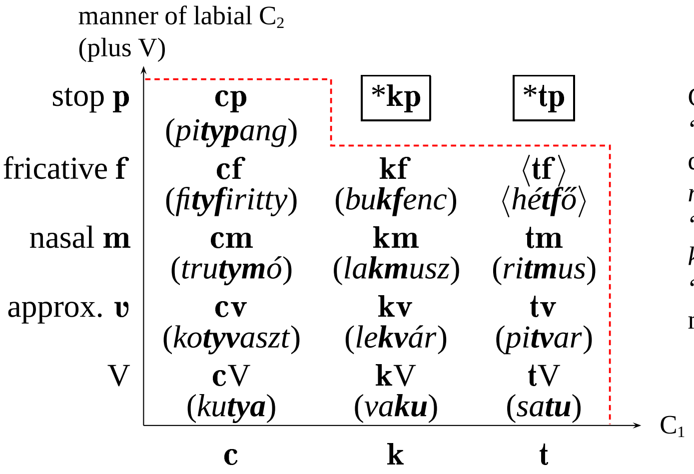
manner of dental C_2 (plus V)	**k**	**p**	**c**
stop **t**	**kt** (*a**kt**a*)	**pt** (*kri**pt**a*)	***ct**
fricative **s**	**ks** (*bu**ks**za*)	**ps** (*ka**ps**zula*)	***cs**
nasal **n**	**kn** (*a**kn**a*)	**pn** (*sra**pn**el*)	***cn**
liq. **l**	**kl** (*cé**kl**a*)	**pl** (*pa**pl**an*)	**cl** (*tro**tyl**i*)
V	**k**V (*la**ka**t*)	**p**V (*ré**pa***)	**c**V (*ku**tya***)

Glosses: *akta* ‘document’, *kripta* ‘crypt’, *buksza* ‘purse’, *kapszula* ‘capsule’, *akna* ‘shaft’, *srapnel* ‘srapnel’, *cékla* ‘beet’, *paplan* ‘duvet’, *trotyli* ‘tramp’, *lakat* ‘lock’, *répa* ‘carrot’, *kutya* ‘dog’.

As we can see in (15), this slice of the phonotactic space is also closed towards the unmarked items; also, again, the marked elements occur in the ‘outskirts’ of the space. As we go along the vertical axis, the contexts are increasingly more and more difficult for the perception of the first stops in question. The horizontal axis shows the first consonants in the order of their difficulty of perception. Velar and labial stops are said by Rebrus & Trón (2002) to be less difficult to perceive in these environments than the palatal place.

Closedness predicts that we should get the same shape of the phonotactic space in all dimensions (and, actually, universally in all languages). Let us consider now voiceless non-labial stop plus labial clusters.

(16) VC_1V, VC_1C_2V sequences: C_1 = voiceless non-labial stop, C_2 = labial



stop **p**	**cp** (*pi**typ**ang*)	***kp**	***tp**
fricative **f**	**cf** (*fi**tyf**iritty*)	**kf** (*bu**kf**enc*)	⟨**tf**⟩ ⟨*hé**tf**ő*⟩
nasal **m**	**cm** (*tru**tym**ó*)	**km** (*la**km**usz*)	**tm** (*ri**tm**us*)
approx. **ʋ**	**cv** (*ko**tyv**aszt*)	**kv** (*le**kv**ár*)	**tv** (*pi**tv**ar*)
V	**c**V (*ku**tya***)	**k**V (*va**ku***)	**t**V (*sa**tu***)
	c	**k**	**t**

Glosses: *pitypang* ‘dandelion’, *fityfiritty* ‘imp’, *bukfenc* ‘somersault’, *hétfő* ‘Monday’, *trutymó* ‘suspicious substance’, *lakmusz* ‘litmus’, *ritmus* ‘rhythm’, *kotyvaszt* ‘concoct’, *lekvár* ‘jam’, *pitvar* ‘porch’, *kutya* ‘dog’, *vaku* ‘flash’, *satu* ‘vice’; *hétfő* ‘Monday’ is regarded by some as a polymorphemic word (*hét* ‘week’ + *fő* ‘head’).

Again, the space is closed towards the unmarked items; the infrequent items, as well as those in which the contrast is neutralized, occur at the edges. The vertical axis is the same as in (15); however, the markedness of the stops in (15) and in (16) is not the same, it is actually reversed. The relative perception of palatals is more difficult before dentals than that of velars (see (15)); on the other hand, labials provide better cues for the palatal place of stops than for velars, as is shown in (16). In other words, if the first consonant is the voiceless palatal stop, a labial sound is a better choice for the next position than a dental. This markedness reversal is the direct upshot of the fact that markedness is a relational/contextual notion.

3.2. Phonotactic Closedness in 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 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[14] 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 (17) 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 has primary stress (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. 14).

---

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

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

manner of coronal C_2
(plus V)

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	**t**

C_1

manner of coronal C_2
(plus V)

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*$_{1365}$
	k	**p**

C_1

As (17) 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 (17), 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 least suggest that for this contrast (**k** – **p**) the pre-nasal position is more marked than any of the others.

Let us turn our attention to the **g** – **b** contrast in the same environments as in (17) (except that now the obstruents following the two segments are *voiced*).

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

manner of coronal C_2
(plus V)

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**

C_1

manner of coronal C_2
(plus V)

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*$_{1490}$
	g	**b**

C_1

The first thing that is apparent in the first chart of (18) 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 somewhat 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**.[15] 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.[16] 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 *op.cit.*:12).

The other choice that is suggested by (18) 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 (19a)).

(19) a. English $V_1C_1C_2V_2$ C_1C_2: voiced, C_1: non-cor. stop, C_2: cor.

manner of coronal C_2 (plus V)

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**

C_1

b. English $V_1C_1C_2V_2$ C_1C_2: voiced, C_1: non-cor. stop, C_2: cor.

manner of coronal C_2 (plus V)

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**

C_1

If we reverse the relative difficulty markedness of **b** and **g** (as suggested above), then we get the chart in (19b). If the aerodynamics argument is valid, then the ordering has to be changed accordingly, as it is done in (19b). 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's (2002) 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.

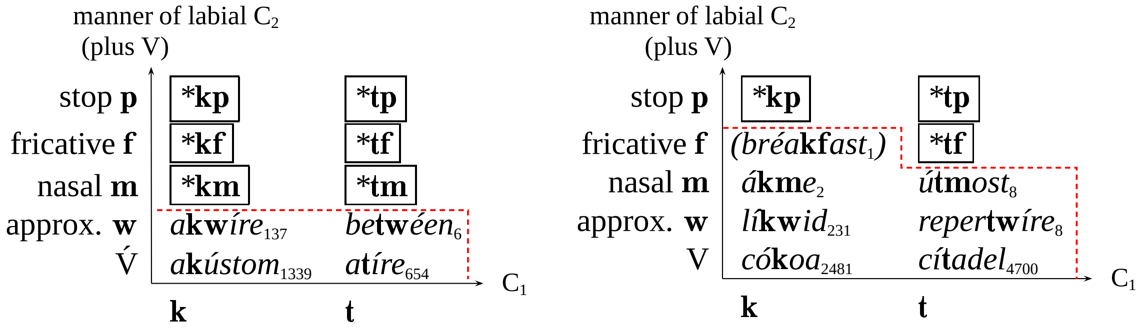
The following dimension of the phonotactic space of English (20) shows voiceless non-labial consonants before labials (and vowels); notice that the voiceless velar stop is less marked than the coronal before labials (cf. (16)).

---

[15] **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**.

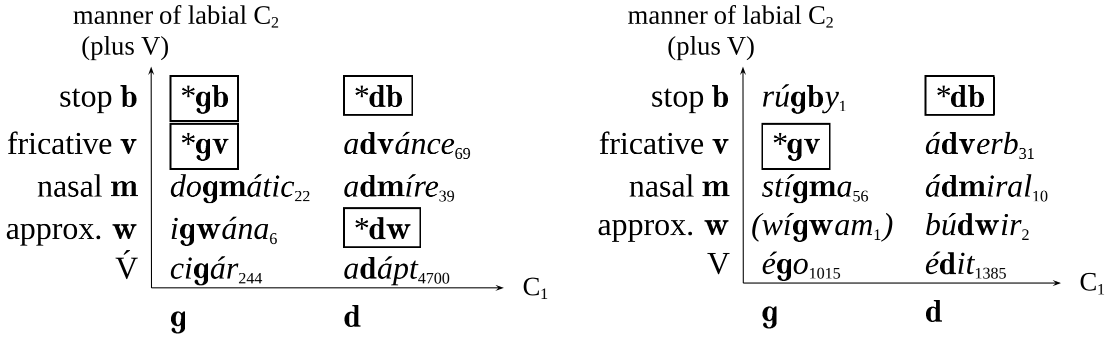
[16] This is also a reason why *long* voiced stops are typologically marked: their production is in this sense more difficult to sustain than that of voiceless geminate stops.

(20) English $V_1C_1C_2\acute{V}_2 - \acute{V}_1C_1C_2V_2$: C_1C_2: voiceless, C_1: non-labial stop, C_2: labial



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.[17] 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.

(21) English $V_1C_1C_2\acute{V}_2 - \acute{V}_1C_1C_2V_2$: C_1C_2: voiced, C_1: non-labial stop, C_2: labial



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 (21), 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**,[18] if we do this, as well as collapse the two relevant charts, we obtain (22).

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

---

[17] 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 **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.

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

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

manner of labial C_2 (plus V)		
stop **b**	*db	**gb**$_{1}$
fricative **v**	**dv**$_{100}$	*gv
nasal **m**	**dm**$_{49}$	**gm**$_{78}$
approx. **w**	**dw**$_{2}$	**gw**$_{7}$
V	**d**$_{6085}$	**g**$_{1259}$
	d	**g**

C_1

done is, again, a matter of future research.[19] The relative markedness of **dw** and **gw** is also conspicuous in (22). Possibly all of these words could be considered non-core vocabulary (cf. *boudoir* ˈ**buːdwɑː**, *iguana* **ɪˈgwɑːnə**, *wigwam* ˈ**wɪgwæm**). What is also curious is that the number of **gw** clusters raises provided they follow the velar nasal **ŋ** (e.g., *anguish, distinguish, language, linguist, penguin* etc.). Phonetic research is needed here to confirm the special status of **dw**/**gw** clusters.[20]

The last dimension we consider in this section is the occurrence of **m** before coronals (and vowels). The chart (23) confirms what has been indicated about the relationship of stress and English phonotactics above: no gaps occur unless the dimension of the stressing of the preceding/following vowel is *not* considered.

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

stop **t d**	*asy***mt***ótic*$_{2}$ *hu***md***ínger*$_{2}$
fric. **θ ð s z**	*mθ *mð *ms *mz
nasal **n**	**amn***ésia*$_{20}$
liquid **l r**	*ml *mr
V́	*a***m***óunt*$_{1442}$
	m

stop **t d**	*mt *md
fric. **θ ð s z**	*mθ *mð *ms *cl**ú***mz***y*$_{23}$
nasal **n**	*á***mn***esty*$_{66}$
liquid **l r**	*ó***ml***et*$_{11}$ *có***mr***ade*$_{5}$
V	*clí***m***ax*$_{2167}$
	m

Another indication that (23) 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

---

[19] 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'.

[20] 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.

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.

3.3. Postnasal voicing and Phonotactic Closedness

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

(24) a. In place assimilations, in C_1C_2, C_1 tends to assimilate the features of C_2.
  b. *Nasals* are the most common targets for place assimilation (including 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 (24a–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)[21] 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 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).[22] 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[23] — 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. For example, postnasal voicing was the most important source of the rise of voiced stops in Hungarian.

(25) 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’) (cf. Cser 2001:59)[24]

In British English RP, as well as many other English dialects, 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–

---

[21] 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).

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

[23] Cf. Jun (1995).

[24] For further—synchronic as well as diachronic examples of postnasal voicing, cf. Kiss (2004).

voiceless stop clusters.[25] 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; (26) displays some of the implicational universals of such clusters.

(26) 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 (27). The chart only concentrates on CC clusters whose second element is a stop (the examples all show a coronal stop).[26]

(27)

	English	Italian	Diola Fogny	Manam	Japanese	Yapese	Ojibwa	Lardil
nh-N+stop (**mt**)	(+)							
fric.+stop (**st**)	+	+					+	
liquid+stop (**lt**)	+	+	+				!	+
N+vless stop (**nt**)	+	+	+	+			+	+
N+voiced stop (**nd**)	+	+	+	+	+		+	!
V+stop	+	+	+	+	+	+	+	+

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 universals (26) are all exemplified by the languages in (27). 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,

---

[25] 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.

[26] The table is based on Pigott (1999) and Rebrus & Trón (2004:147). '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 *op.cit.*:147).

however, that no implication has been established between the existence (or lack) of **st** and **lt** clusters in (26). 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 (28), and word-final (29).[27]

(28) English monomorphemic VC$_1$C$_2$V clusters (C$_2$=stop)

j/w/h						
θ/ð						
v						
ʃ/ʒ						
nh.-N						
ʧ/ʤ	(*vegetable*)					
z				(*asdic*)	(*husband*)	(*Glasgow*)
g				(*Magdalene*)	(*rugby*)	
b	obtain	(*subcutaneous*)		abdominal		
d		(*vodka*)	(*jodhpurs*)			(*Edgar*)
p	chapter					
k	actor			(*anecdote*)		
t		(*Atkins*)				
f	after					(*Afghan*)
s	custom	fiscal	hospital			
l	alter	alcohol	pulpit	shoulder	album	vulgar
r	forty	turkey	harpoon	cardigan	turbine	forget
N	winter	wrinkle	temper	bandit	gambit	hunger
V	better	accustom	lepel	ready	ruby	ego
	t	**k**	**p**	**d**	**b**	**g**

---

[27] The words that are in brackets indicate clusters that occur in just few words (sometimes only in that word). Notice that, unlike in earlier charts, the x-axis now houses 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. For more comprehensive lists of English non-initial CC clusters, see Kiss (2001).

(29) English monomorphemic VC$_1$C$_2$# clusters (C$_2$=stop)

j/w/h						
θ/ð						
v						
ʃ/ʒ						
ʧ/ʤ						
z						
g						
b						
d						
nh.-N	(*dreamt*)					
p	apt					
k	act					
t						
f	soft					
s	mist	ask	wasp			
l	melt	sulk	pulp	hold	(*bulb*)	
r	part	bark	carp	guard	disturb	mourge
N	vent	think	jump	find	!	!
V	cut	make	sip	bed	stab	fig
	t	**k**	**p**	**d**	**b**	**g**

As we can see, the word-final position—as opposed to the word-internal—appears to be highly restrictive for voiced stops in English: only the dental one, and not the labial or the velar may occur after nasals in that position, for example. Such severe restrictions as these are not observable in the intervocalic (crucially: *prevocalic*) position.

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.

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

All three scales can be grounded phonetically, as we mentioned.[28] An important consequence of the scales in (30) 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.

---

[28] 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.

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.

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

- a. *System 1*: Scale 1 wins over Scale 2 (i.e., voicing is not preferred before a vowel): **ta**, ***da**, **nta**, ***nda** (e.g., Lardil)
- b. Scale 2 wins over Scale 1 (i.e., voicelessness is not preferred after a nasal):
  - (i) *System 2*: (**ta**, **da**), ***nta**, **nda** (e.g., Japanese)
  - (ii) *System 3*: (**ta**, ***da**), ***nta**, **nda** (e.g., Wembawamba)
- c. *System 4*: the markedness statements are not enforced (i.e., voicing is maintained in all sequences): **ta**, **da**, **nta**, **nda** (e.g., Hungarian)

In languages like System 1, there is no voicing contrast for stops whatsoever. This is the consequence of the priority of the post-stop vowel (cf. Scale 1 in (30)): 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 2 and 3 show that the scales in (30) 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. A language may allow for marked sequences along Scale 1 (i.e., **da** besides **ta**), while it resolves the N__ vs. __V conflict in favour of postnasal voicing (System 2). In System 3, there is neutralization on Scale 1 (**ta**, ***da**), while the conflict between Scale 1 and 2 is resolved in favour of Scale 2, where postnasal voicing is more important than prevocalic voicelessness: ***nta**, **nda**. Lastly, languages of type System 4 do not enforce the markedness scales, they allow for the marked sequences, too. Crucially, no systems are predicted like **da**, ***ta** or **nta**, ***nda**, 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 (cf. the case of System 2 languages, like Lardil).

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 (30) 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.

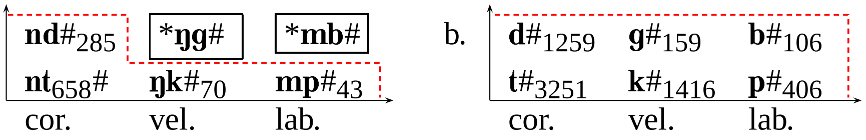
(32) Word-final sequences: **t#**, **d#**, **nt#**, **nd#**

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 (32)) displays the independence of Scale 2 and 3 (30): there can be voicing contrast for stops word-finally, but neutralization into **t** after a nasal.[29] 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 (33a) (NC#) and (33b) (C#).

(33) a. b.



The perceptual (and/or articulatory) grounding of these cases definitely needs further research; the expectation is that ‘heavy’ clusters (such as voiced labial **mb**) are perceptually less robust 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.

(34) 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: **Vːnt#** (50 items), **Vːnd#** (61 items), but: ***Vːŋk#**, ***Vːmp#**, ***Vːŋg#**, ***Vːmb#**
d. if the V before the word-final nasal–stop cluster is long/tense, it is usually non-high:

---

[29] 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**.

iː/uː				
əʊ	don't$_3$			
ɔː	flaunt$_5$	launch$_5$		
ɔɪ	point$_6$			
eɪ	paint$_{13}$			
ɑː	slant$_6$	branch$_5$		
aɪ	pint$_1$			
aʊ	count$_6$			
	nt	**ntʃ**	**ŋk**	**mp**

iː/uː	fiend$_1$/wound$_1$			
əʊ				
ɔː	laundry$_4$			
ɔɪ				
eɪ		change$_9$		
ɑː	command$_9$			
aɪ	find$_9$			
aʊ	sound$_{28}$			
	nd	**ndʒ**	**ŋg**	**mb**

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 — the low vowels — in order to enhance their own salience.[30]

4. Conclusions

The paper suggested that the phonotactics of languages is often gradual, non-categorical, in the sense that not all possible combinations are grammatical along the various dimensions of segmental contrast. The graduality of phonotactics is a property that has been shown to be better explained by a model that employs perception-based functional principles, such as cue licensing according to relative context, and phonotactic closedness. The phonotactic space is defined by hierarchical perceptual difficulty scales that linearly order the various syntagmatic relations a segment can enter into. All existing items (marked as well as unmarked) in the language are accounted for in this model, what is more, their place in the phonotactic space is also predicted: contrasts in environments with robust perceptual cues are predicted to occur around the 'origin' of the scales making up the space (their number will also be predicted to be relatively high), whereas the marked clusters will occur at the edges. In this model, thus, the notions 'exceptional' and 'accidental gap' are meaningless. Where exactly a language draws the line between a contrast that it makes use of and those that it does not, is arbitrary, but if a perception-wise marked cluster is used in a language, then our model will predict that all the other clusters that are perceptually better cued will also occur in the language, in accordance with the principle of phonotactic closedness. In some cases the perception hierarchies may stand in conflict, it seems that the actual 'output' of the conflict is arbitrary; Phonotactic Closedness cannot play a role in these cases *if the conflicting scales are independent of each other*; Closedness only deals with

---

[30] 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.

spaces that are defined by *related* markedness scales. Actually, if in the investigation of two dimensions of a contrast it turns out that closedness is violated, it can usually be taken as an indication that the two dimensions are incompatible, that they cannot be related, or compared.

Further research is needed in finding and further specifying the phonetic bases of the various segment clusters, as phonetic grounding is the cornerstone of this approach, the role of other factors needs to be also considered, such as token frequency, analogy, paradigm uniformity, phonotactic layers, etc., as well as their interaction with Phonotactic Closedness.

Acknowledgements

My thanks are due to Miklós Törkenczy, Péter Rebrus, Zsuzsanna Bárkányi, and Anna Hamp-Kürti for the inspiring discussions on the topic and for reading earlier versions of the paper. I would like to thank the comments and suggestions of the reviewer of the paper, Sara Mackenzie. I would also like to thank Sylvia Blaho for all the help at ConSOLE XIII. As always, the remaining mistakes are all mine.

Zoltán Kiss
Eötvös Loránd University (ELTE), Budapest
cash@nytud.hu
http://budling.nytud.hu/∼cash

References

Brockhaus, W. (1990). Colourful leagues: a Government Phonology approach to final obstruent devoicing in German. *UCL Working Papers in Lingusitics* 2, pp. 270–297.

Browman, C. & L. Goldstein (1990). Tiers in articulatory phonology, with some implications for casual speech. Kingston & Beckman (1990), pp. 341–376.

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

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. Kingston & Beckman (1990), pp. 283–333.

Cser, A. (2001). *A zörejhangok gyengülési és erősödési folyamatainak tipológiája és modellezése* [*The typology and modelling of the lenition and fortition processes of obstruents*]. Doctoral dissertation, Elméleti nyelvészet doktori program, ELTE/MTA, Budapest.

Diver, W. (1979). Phonology as human behavior. Aaronson, D. & R. Rieber (eds.), *Psycholinguistic research: implications and applications*, Lawrence Erlbaum Associates, Hillsdale, pp. 161–182.

Flemming, E. (1996). Evidence for constraints on contrast: the dispersion theory of contrast. *UCLA Working Papers in Phonology* 1, pp. 86–106.

Flemming, E. (2004). Contrast and perceptual distinctiveness. Hayes, B., R. Kirchner & D. Steriade (eds.), *Phonetically based phonology,* Cambridge University Press, Cambridge, pp. 232–276.

Harris, J. (1990). Segmental complexity and phonological government. *Phonology* 7, pp. 255–300.

Harris, J. (1994). *English Sound Structure*. Blackwell, Cambridge, MA & Oxford.

Hayes, B. (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, B. & D. Steriade (2003). The phonetic bases of phonological markedness. Manuscript.

Itô, J. & A. R. Mester (1995). Japanese phonology. Goldsmith, J. A. (ed.), *The Handbook of Phonological Theory*, Blackwell, Cambridge, MA & Oxford, pp. 817–838.

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

Kaye, J., J. Lowenstamm & J.-R. Vergnaud (1990). Constituent structure and government in phonology. *Phonology* 7, pp. 193–231.

Kingston, J. & M. E. Beckman (eds.) (1990). *Papers in Laboratory Phonology I: Between the grammar and the physics of speech*. Cambridge University Press, Cambridge.

Kiss, Z. (2001). *Distributional effects in English and Hungarian non-initial CC clusters*. M.A. thesis, Eötvös Loránd University, ELTE, Budapest, URL `budling.nytud.hu/~cash/papers/papers.html`.

Kiss, Z. (2004). Markedness, graduality and closedness in phonotactics—a phonetically grounded approach. *The Even Yearbook* 6, pp. 85–110, URL `www.btk.elte.hu/delg/publications/even/index.html`.

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

Liljencrants, J. & B. Lindblom (1972). Numerical simulation and vowel quality systems: the role of perceptual contrast. *Language* 48, pp. 839–862.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge University Press, Cambridge.

Myers, S. (1997). Expressing phonetic naturalness in phonology. Roca, I. (ed.), *Derivations and constraints in phonology*, Clarendon Press, Oxford, pp. 125–152.

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

Pigott, G. L. (1999). At the right edge of words. *The Linguistic Review* 16, pp. 143–185.

Rebrus, P. & V. Trón (2002). A fonotaktikai általánosításokról. (Kísérlet a magyar mássalhangzó-kapcsolatok nemreprezentációs leírására) [On phonotactic generalisations. (An attempt at the non-representational description of Hungarian consonant clusters)]. Maleczki, M. (ed.), *A mai magyar nyelv leírásának újabb módszerei V. Szeged, 2001. szeptember 27–29* [*Current methods in the description of Hungarian 5. Szeged, 29–29 September 2001*], Szegedi Tudományegyetem, Általános Nyelvészeti Tanszék, Magyar Nyelvészeti Tanszék, Szeged, pp. 17–63.

Rebrus, P. & V. Trón (2004). Fonetikai motiváció a fonológiai mintázatokban [The phonetic motivation of phonological patterns]. Hunyadi, L. (ed.), *Kísérleti fonetika labóratóriumi gyakorlatban* [*Experimental phonetics in laboratory practice*], Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, pp. 139–164.

Rice, K. D. (1992). On deriving sonority: a structural account of sonority relationships. *Phonology* 9, pp. 61–99.

Siptár, P. & M. Törkenczy (2000). *The Phonology of Hungarian*. Oxford University Press, Oxford.

Steriade, D. (1997). Phonetics in phonology: the case of laryngeal neutralization. Manuscript. UCLA.

Steriade, D. (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.

Törkenczy, M. (1994). A szótag [The syllable]. Kiefer, F. (ed.), *Strukturális magyar nyelvtan 2: Fonológia* [*A Structural Grammar of Hungarian 2: Phonology*], Akadémiai Kiadó, Budapest, pp. 273–392.

Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change*. Mouton de Gruyter, The Hague & Berlin.

Zipf, G. K. (1949). *Human behavior and the Principle of Least Effort*. Addison-Wesley Press, Cambridge, MA.