Acta Linguistica Hungarica, Vol. 49 (3–4), pp. 427–456 (2002)

H AS IN HUNGARIAN

PÉTER SIPTÁR – SZILÁRD SZENTGYÖRGYI

Abstract

The paper discusses the possible analyses of the behaviour of [h] and [x] in Hungarian. It argues that in a derivational, rule-based framework two types of analyses are possible: one that assumes two separate underlying segments, /x/ and /h/, and thus misses the generalisation that the two segments are in complementary distribution, a typical characteristic of allophones. The second kind of approach argues that [h] and [x] come from the same underlying segment; this type of analysis can be further divided into two subtypes. According to one of these, the underlying segment is /h/. To be able to derive the attested output forms, three separate strengthening rules must be posited, an obvious disadvantage. The other possible approach, on the other hand, argues that the underlying segment is always /x/ weakened into a [h] in onsets and deleted in a group of lexically marked words by a minor rule. Besides, we also consider the behaviour of *H*-type segments in voice assimilation: they trigger but do not undergo that process. Siptár and Törkenczy (2000) suggest that if a filter disallowing surface voiced dorsal fricatives, i.e., ***ɣ**, is proposed, then the desired result is obtained.

While such a filter is an ad hoc device in rule-based theories, it is an organic part of a solution in Optimality Theory (OT), which argues that both /h/ and /x/ may occur in the input and the constraint hierarchy must be such that they should always select well-formed output candidates as optimal regardless of the input. As a result of this and Lexicon Optimization (LO), non-alternating forms will have /h/ or /x/ in their underlying representation depending on the output forms while alternating forms may have an underlying /x/ or /h/ as a result of the alternation sensitive LO (Inkelas 1994). Finally, we will show that the treatment of the behaviour of /x/ or /h/ in voice assimilation is simple in OT if we assume the constraint proposed by Siptár and Törkenczy (2000), prohibiting voiced dorsal fricatives, ***ɣ**, which, interacting with the ones suggested by Petrova et al. (2001), will be able to select the actual surface form as optimal in all cases.

1. Introduction

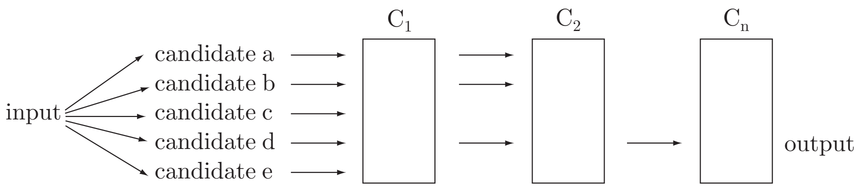
The paper is concerned with the description of the behaviour of *H*-type segments in two types of framework: it starts with a proposal within a rule-based derivational theory and ends with a non-derivational account given in Optimality Theory (OT).

Optimality Theory as first described in Prince–Smolensky (1993) is an input-output device without serial derivation or intermediate levels. The OT

1216–8076/02/$ 5.00

model of grammar consists of three major components: the Lexicon, which contains the underlying representations of all the words and morphemes of the language; the Generator function (Gen), which maps each input form onto an infinite number of output candidates; and the Evaluation function (Eval) mapping this infinite set of possible candidates onto one candidate, the optimal one, which is the output form corresponding to the input. Eval is made up of a set of ranked constraints and it is these constraints that constitute Universal Grammar (UG) in the sense that all languages or language varieties have exactly the same constraints. It is the relative priority or ranking of the constraints that distinguishes one language from another. The constraints can be violated by the output candidates, but these violations are not necessarily fatal as we shall see. It is important to note that evaluation is parallel, no serialism is involved.

(1) OT AS AN INPUT-OUTPUT DEVICE



In (1), Gen generates an infinite candidate set from the input, five of which are shown. C_1, C_2 and C_n are the constraints of Eval and they map the infinite set onto a one-member set, the optimal candidate, which is the output form.

Since we are making use of the Correspondence Theory of Faithfulness (McCarthy–Prince 1995), we also have to mention a mapping between input and output forms, which sets up a correspondence relation between segments/units of the input and those of the output by a simple indexation. Input and output segments correspond to each other if they have the same index.

Constraints can be of two major kinds: markedness constraints requiring that output forms are only made up of unmarked units (e.g., “all front vowels must be unrounded” = no front rounded vowels), and faithfulness constraints, which penalize changes to the input form (e.g., “all input segments must have an output correspondent” = no deletion). The evaluation of output candidates by the constraint hierarchy can be illustrated by tableaux, which show the output candidates in the first column followed by the constraints, left-to-right, starting with the dominant, highest ranked ones. Asterisks indicate violations of the constraint by the candidate and exclamation marks show which violation is fatal, while the cells containing the violations that are not relevant because they are lower than all the fatal violations are shaded for easier understanding. The rightward pointing hand shows the optimal candidate.

(2)

	C_1	C_2
☞ candidate a		∗
candidate b	∗!	

If all the candidates satisfy or violate a high ranked constraint, then the decision is passed on to the next constraint in the hierarchy as illustrated by the schematic tableaux in (3) and (4).

(3)

	C_1	C_2
☞ candidate a	∗	
candidate b	∗	∗!

(4)

	C_1	C_2
☞ candidate a		
candidate b		∗!

Since according to one of the important properties of OT, the Richness of the Base, any kind of form is possible in the input, the constraint hierarchy has to be such that independent of the input form, the optimal output candidate should always conform to the requirements of the language. Because of this, it may happen that evaluating two different underlying forms results in phonetically identical optimal outputs. In such cases, it is Lexicon Optimization (LO) that decides which underlying form should be preferred by the language: it is always the one in the case of which the output candidate has fewer and less serious violations. This is shown in tableaux (5)–(6).

(5)

Input 1	C_1	C_2
☞ candidate a		
candidate b	∗!	

(6)

Input 2	C_1	C_2
☞ candidate c	∗	
candidate d	∗	∗!

where *candidate a* is phonetically identical to *candidate c*

As we can see, from the two different input forms two winning candidates are selected, candidates (a) and (c), which are phonetically identical. Lexicon Optimization decides that the input form in (5) should be preferred as the underlying representation of the word because if we compare the two winning candidates, it can be seen that while candidate (c) does violate the higher ranked C_1, (a) does not violate either constraint and is thus "closer" to the input.[1]

The organisation of the paper is as follows: section **2** introduces the distribution of *H*-type segments and their possible representations. Section **3** discusses the three derivational analyses, one assuming that both /x/ and /h/ are underlying, another one proposing only /h/ as an underlying segment, and finally the one assuming only an underlying /x/. Section **4** draws the reader's attention to the problems of such derivational analyses while section **5** introduces the facts concerning the behaviour of *H*-type segments in voice assimilation. Section **6** shows how an analysis in OT is superior to the derivational ones discussed before and we give an account of the voice assimilation facts in the framework of OT in section **7** followed by the conclusion.

2. The distribution of *H*-type segments

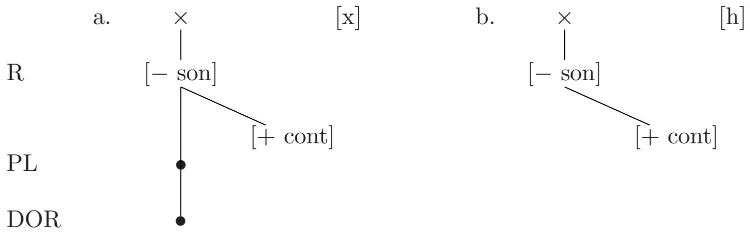
In present-day Hungarian speech, there are four different "*H*-type" segments: a voiceless glottal fricative [h] (as in *hó* 'snow'), a voiced glottal fricative [ɦ] (as in *ruha* 'dress'), a voiceless velar fricative [x] (as in *doh* 'musty smell'), as well as a slightly fronted variant of the latter (as in *pech* 'bad luck') that is often erroneously identified with palatal [ç]. The difference between [h] and [x] **is** phonologically relevant: the distribution and phonological analysis of those two segments is the subject matter of the present paper. Voiced [ɦ] is merely a phonetic (coarticulatory) variant of glottal [h] occurring in a post-sonorant (including intervocalic) position; the fronted velar fricative as in *pech*, on the other hand, is related to the [x] of *doh* in the same manner as e.g., the fronted [k] of *fék* 'brake' to the non-fronted [k] of *fok* 'degree'. That is, the small phonetic difference between them is phonologically irrelevant. In what follows, the difference between [h] and [ɦ], as well as that between velar [x] and its fronted variant, will be disregarded.

In this paper, then, we will discuss [h] as in *hó* and [x] as in *doh*. (Capital *H* will be used to refer to the two segment types together until we decide whether they are variants of the same underlying segment or else two distinct members of the consonant inventory of Hungarian.) Their phonological rep-

[1] For further general introduction to OT, cf. Archangeli – Langendoen (1997); Kager 1999; Roca – Johnson (1999, ch. 19); McCarthy (2002).

resentations are as follows[2] (R = root node, PL = place node, DOR = dorsal node; the lack of place specification in the case of [h] is meant to suggest that it is articulated outside the oral cavity, in the glottis):

(7) REPRESENTATIONS OF THE TWO TYPES OF *H*



In coda position, *H* is either deleted (e.g., *méh* [meː] 'bee', *cseh* [ʧɛ] 'Czech'), or else it occurs in the form shown in (7a) (e.g., *jacht* 'yacht', *technika* 'engineering', *ihlet* 'inspiration', *Ahmed*; *doh* 'musty smell', *potroh* 'abdomen', *sah* 'Shah', *Allah*, *padisah* 'Padishah', *sarlach* 'scarlet fever', *almanach* 'yearbook', *moloch* 'Moloch', *eunuch* 'id.', etc.) Geminate *H* is always like (7a), i.e., a velar fricative, irrespective of whether it occurs in a branching coda (*pech* 'bad luck', *cech* 'bill', *Bach*, *krach* 'crash', *fach* 'pigeon-hole') or is divided between a coda and a subsequent onset (*peches* 'unlucky', *Bachot* 'Bach' (acc.)). Elsewhere, an *H* occurring in an onset is always (7b), a glottal fricative (*hó* 'snow', *ruha* 'dress', *konyha* 'kitchen'). How should we account for these facts?

3. A derivational analysis of the distribution facts

Let us start with the alternation of the type [ʧɛ] ∼ [ʧɛhɛk] 'Czech' (sg.) ∼ (pl.). As in any ∅ ∼ X alternation (that is, in all cases where something alternates with nothing, the lack of itself), two basic types of accounts suggest themselves: deletion (syncope) and insertion (epenthesis). In the latter case we could say that *cseh* and *doh* behave in two different ways because the lexical representation of *doh* includes an *H*, whereas that of *cseh* does not. In *cseh*-type words, then, there would be a rule of *H*-insertion that would apply before vowel initial suffixes. Given that the syllabic position at hand is an onset, this rule would obviously insert a [h], that is, a segment of the type (7b), as in *csehül* 'in Czech', *csehek* 'Czechs'. However, this solution has two

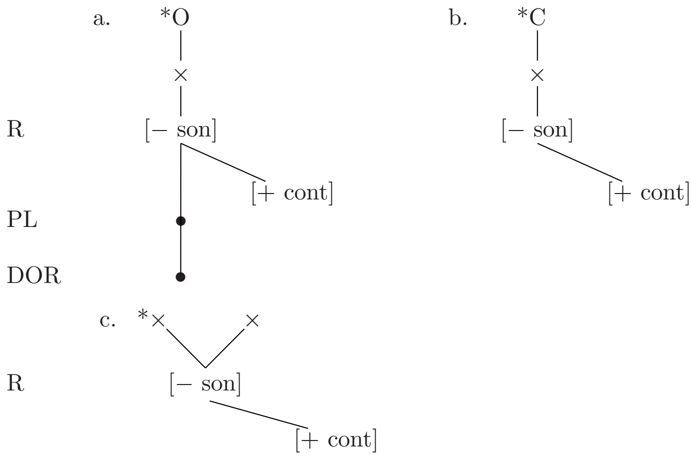
[2] See Siptár – Törkenczy (2000, 7–9), with respect to the feature geometry assumed in this paper.

serious drawbacks. First, the overwhelming majority of vowel final words do not exhibit such insertion: **kávéhok* (cf. *kávé-k* 'coffee' (pl.)), **kapuhok* (cf. *kapu-k* 'gate' (pl.)). This problem could be solved by taking *H*-insertion to be a **minor rule**: one that only applies to words specifically marked in the lexicon to that effect (i.e., the entry of *cseh* would include the instruction 'apply *H*-insertion', whereas that of *kávé* would not). But there is an even more serious difficulty that this solution would have to face. Suffixes consisting of a single consonant are attached to regular vowel final stems with no intervening 'linking vowel' (*kávé-t* (acc.), *kávé-k* (pl.)). In other words, the condition on *H*-insertion that the insertion site must be followed by a vowel would not be met in these cases. What is more, in the case of "*cse*", low vowel lengthening would also come into the picture (cf. Siptár–Törkenczy 2000, 170–3). That is, the expected forms would be **csét*, **csék* (as in *kefét* 'brush' (acc.), *kefék* 'brush' (pl.)). Thus, *cseh*-type words, were they to end in a vowel in their lexical representation, would **exceptionally** be exempted from low vowel lengthening, would **exceptionally** require a linking vowel before suffixes consisting of a single consonant, and would **exceptionally** insert a [h] before that linking vowel. These three sorts of exceptionality are the easiest to record in the lexicon in the form of assuming that the lexical item ends in an *H*. That is, the insertion account is untenable.

In the deletion case, on the other hand, we would have to be able to tell why some words require deletion (*cseh*), whereas others do not (*doh*). According to the traditional view, deletion is the regular event for *H*-final words, and all items in which it does not apply are exceptions (or, worse still, "not proper Hungarian words"). But, first of all, a lot more words behave like *doh* than like *cseh*, and whenever a new word is introduced (either by borrowing or, e.g., as an acronym like *MÉH* 'name of a company for collecting waste material', *APEH* 'name of the tax office', and *BAH* 'name of an intersection in Budapest'), it invariably joins the *doh* group, not the *cseh* group. The pattern exemplified by *doh* is the productive one. Secondly, there is a significant amount of vacillation in the *cseh* group: *juh* 'sheep', *pléh* 'tin', *céh* 'guild', *düh* 'anger', *rüh* 'scabies', *oláh* 'Wallachian', *éh* 'hunger', *keh* 'wheeziness', *tereh* 'burden' are all traditionally of the *cseh* type but all of them exhibit extensive variability and, for most of them, *doh*-type behaviour, that is, the lack of *H*-deletion seems to gain the upper hand. Thus, we are forced to draw the conclusion that *H*-deletion is a minor rule: words that undergo it are exceptional, not words that do not. The exact way of formulating this rule of *H*-deletion will be decided on when we have accounted for the distribution of the two kinds of *H*.

This can be done in three different ways in principle. First, we could assume (with Ritter 2000, 28–9) that we have to do with two distinct underlying segments here. Their distribution would then be accounted for by way of filters of the following form (O = onset, C = coda, × = timing slot):

(8) DISTRIBUTION OF THE TWO TYPES OF *H*



(No [x] *in an onset (except if it is also dominated by a coda node, i.e., it is long); no* [h] *in a coda; long H cannot be glottal, whether it occurs in a branching coda or in a coda + onset position.)*

The bonus in this solution would be that we would get deletion for free in the *cseh* type: all we would have to assume is that the words belonging here (exceptionally) include a glottal [h] in the lexicon that can only surface in case it gets into onset position by suffixation (*csehek* 'Czech' (pl.)). Otherwise (e.g., *cseh*, *csehben* 'in (a) Czech', *Csehország* 'Czech Republic', *cseh ellenzék* 'Czech opposition')—given that it cannot be parsed as a coda in view of (8b)—it could not be syllabified at all and as a stray segment it would have no audible effect on phonetic implementation (would fail to be pronounced).

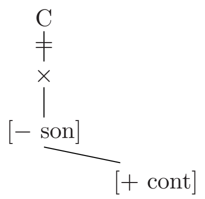
However, this solution would also have serious drawbacks in a derivational framework. First, discounting the handful set of *cseh*-type words, the two kinds of *H* are in complementary distribution: in contexts where one of them occurs, the other one never does, and *vice versa*. This means that it is impossible to find a pair of words such that the only difference between them is that one has (7a) whereas the other has (7b) in the same position. That is, the distribution of the two types of *H* is predictable and as such it is not to be recorded in the lexicon but rather to be formulated as a phonological rule. Furthermore, there is also alternation between the two types of *H* (e.g., *doh*

[x] ∼ *dohos* [h] 'musty'). Hence, a rule that turns one into the other is required as part of the grammar anyway, a fact that makes the solution involving two distinct lexical (underlying) representations totally superfluous (again, if we think in terms of a derivational account).

The second and third solutions have one thing in common: they assume a single *H* in the lexicon. In other words, they claim that the two different *H*'s appearing in pronunciation are context-dependent surface representations of the same underlying segment. The only remaining question is which to derive from which.

Let us assume, first, that all *H*'s are represented in the lexicon as (7b), i.e., the placeless (glottal) version: /h/. Then we need a rule that inserts a place specification into all *H*'s that are in coda position (as well as a minor rule of /h/-deletion that removes the /h/ of a *cseh*-type word if it finds itself in coda position; the latter rule will have to precede (bleed) the former). These two rules would look like this (see (9) and (10), respectively):

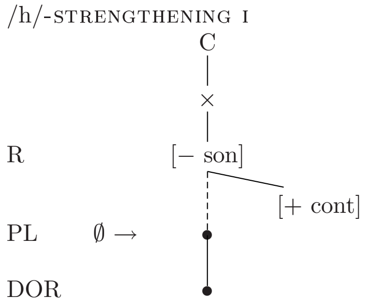
(9) /h/-DELETION



(A placeless continuous obstruent (= a [h]*)—in words that are specified in the lexicon so as to undergo this rule—is delinked together with its timing unit, if it is in coda position.)*

The deletion rule has to remove the ×, too, because /h/-deletion never involves compensatory lengthening (e.g., *csehnek* 'for a Czech' is pronounced [ʧɛnɛk], not [ʧɛːnɛk]; the latter can be a pronunciation of *cselnek* 'for a trick' with *l*-deletion, but it is not a possible rendering of *csehnek*).

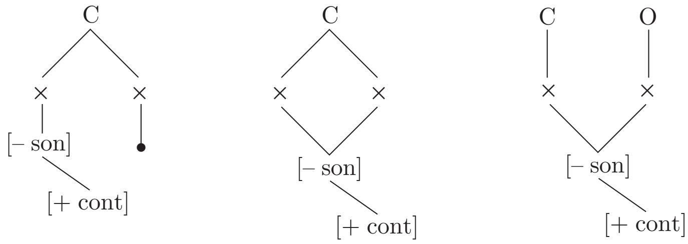
(10) /h/-STRENGTHENING I



(Insert a place node dominating DOR into a /h/ in coda position.)

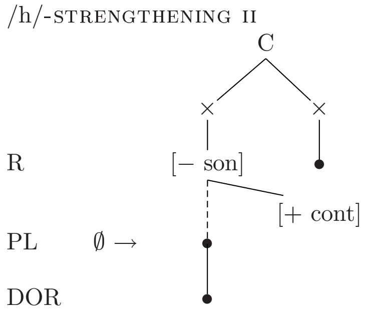
But this is not the whole story. A rule like (10) will insert a place node dominating DOR into e.g., *doh* 'musty smell', *potroh* 'abdomen', *almanach* 'id.', as well as *ihlet* 'inspiration', *technika* 'technique', but say nothing about *jacht* 'yacht' that has a branching coda, or words like *pech* 'bad luck', *krach* 'crash', in which both branches of the coda are filled by /h/, or about *peches* 'unlucky', *krachok* 'crash' (pl.), whose long /hː/ is divided between a coda and a subsequent onset. These three configurations are presented in (11):

(11) a. *jacht* b. *pech* c. *peches*



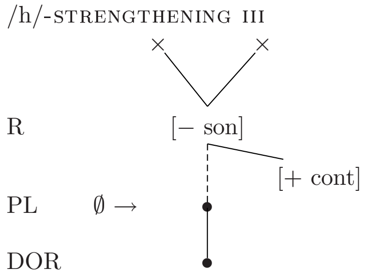
It can be seen clearly that rule (10) does not fit any of the configurations in (11). Hence we need an additional rule of the form in (12) to tackle the situation in (11a), and one of the form in (13) to apply in cases like (11b–c).

(12) /h/-STRENGTHENING II



(Insert a place node dominating DOR into a /h/ in the first position of a branching coda.)

(13) /h/-STRENGTHENING III



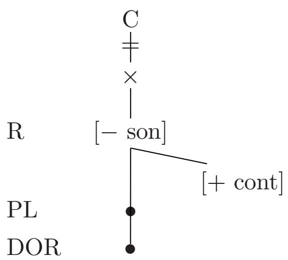
(Insert a place node dominating DOR into a long /hː/.)

The advantage of this solution is that phonological segments that have several surface realizations are usually represented in the lexicon by their word initial versions (or more generally, by their versions appearing in onsets) and it is those versions that serve as input to the derivation of their syllable final (coda) realizations. For instance, the lexical shape of /j/ is the sonorant [j] as in *jó* 'good'—and this is what the voiced obstruent [ʝ] as in *dobj* 'throw' (imp.) and the voiceless obstruent [ç] as in *lopj* 'steal' (imp.) are derived from (cf. Siptár–Törkenczy 2000, 205–6). However, this is not always the case: for instance, the onset and coda versions of /v/ are both derived from a third, neutral version (cf. Siptár–Törkenczy 2000, 198–205).

The most important drawbacks of the solution in (9)–(13) are as follows. First, it is rather complicated: in addition to the deletion rule, three different strengthening rules are required by it. Second, strengthening in coda position is cross-linguistically marked, or "unnatural": that position is usually a lenition site. And third, this solution is also arbitrary: why is it just a dorsal place specification that is inserted (i.e., why is /h/ strengthened into a velar—rather than, say, dental—fricative)? And where does that DOR come from?[3]

The third solution avoids all three of the above pitfalls. It says that all *H*'s are lexically represented as (7a), i.e., /x/. That underlying /x/ then loses its place specification in an onset, cf. rule (15),[4] whereas in coda position nothing happens to it (except in the exceptional *cseh* set where it is deleted in full,[5] with its timing slot and all):

(14) /x/-DELETION



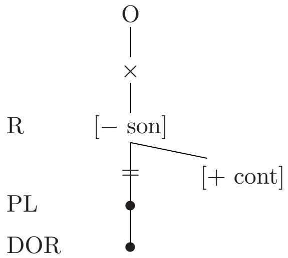
(A dorsal continuous obstruent (= a [x]*)—in words that are specified in the lexicon so as to undergo this rule—is delinked together with its timing unit, if it is in coda position.)*

[3] If all relevant examples involved a back vowel (like *doh*), we could claim that the source of the spreading of DOR is that back vowel. However, in cases like *ihlet* 'inspiration', *pech* 'bad luck' etc. (where the adjacent vowel has COR) the DOR node has to come literally from nowhere.

[4] This is a typical lenition process occurring, in at least some of the cases, in a typical lenition site: intervocalically (cf. Harris 1990; 1997). However, it also occurs in **initial onsets** where lenition typically does not take place. This is just as problematic as the strengthening in the coda in the alternative analysis.

[5] Alternatively, we could also claim that *cseh*-type morphemes have two underlying allomorphs, a consonant-final and a vowel-final one, whose selection is morphological. This solution is suggested in Siptár – Törkenczy (2000, 276).

(15) /x/-WEAKENING



(Delink the place node dominating DOR from a /x/ occurring—exclusively—in onset position.[6]*)*

In sum, all Hungarian consonants articulated behind the palatal area, i.e., not only /k/ and /g/ but also *H* (that is, /x/) are velar (contain DOR as a place specification). This place specification is delinked (and lost) in an onset, whereby glottal [h] is produced: *ha* 'if', *nátha* 'flu'; *ruha* 'dress', *konyha* 'kitchen'; *csehek* 'Czech' (pl.). In all other cases, the underlying representation surfaces (*doh* 'musty smell', *technika* 'technology', *peches* 'unlucky'), except in a few words (*cseh* 'Czech', *méh* 'bee'), where the whole /x/ is deleted in order to avoid ending up in coda position.

4. Problems with the derivational analysis

We may distinguish two types of problems concerning the derivational analyses of the behaviour of *H*-type segments. The first kind of problem is general theoretical while the second kind is particular to the aforementioned derivational solutions.

As far as general theoretical problems are concerned, we can refer to the issues that have led to the emergence of non-derivational theories. Among these we find the question whether a derivational type of grammar purely consists of a series of rules or rather rules and constraints. Since the Obligatory Contour Principle (OCP), as well as other morpheme well-formedness conditions, was suggested as a constraint not to be violated, we can claim that such a grammar cannot only be made up of rules. If, however, there are both rules and constraints, how are their possible interactions limited?

[6] The non-application of (15) to a geminate /xː/ (i.e., to the configuration (11c)) is an instance of geminate inalterability (cf. Siptár – Törkenczy (2000, 276–7) and references cited there). As the rule explicitly refers to the timing tier, it is to be interpreted exhaustively, i.e., it does not apply to an input in which the segmental content is multiply linked to two timing slots and is in coda and onset position at the same time.

Such a theory would definitely be too powerful, it would result in predicting a number of possible but unattested grammars. Optimality Theory (OT) offers a different alternative: it suggests that grammar only consists of constraints and that these constraints are ranked with respect to each other; thus the question of the interaction between rules and constraints does not even arise.

The other important theoretical type of problem is that of intermediate representations: since there are several rules in a grammar, ordering them in the appropriate manner may give us whatever we want, i.e., the possibility to order rules extrinsically is too powerful a device, too, in itself. Since OT is an input-output device, that is one without intermediate representations, this problem may be avoided as well.

Also, conspiracies are very difficult to explain in a rule-based theory, where the structural description, i.e., what triggers the rule, and the structural change, i.e., what happens or how the undesirable situation is amended, are contained within the same rule. On the other hand, OT captures this kind of functional unity of the rules in a simple but elegant way: the trigger constraint only states the undesirable configuration but it is the other constraints and their relative ranking that selects what the best, or rather optimal, solution is.[7]

As far as the particular problems with the above derivational solutions are concerned, some of them have already been mentioned: if we stipulate both underlying segments, /x/ and /h/, then the generalisation that the surface [x] and [h] are allophones is missed. On the other hand, if we assume that in every instance of a surface *H*-type segment there is an underlying /h/, then we need three different strengthening rules in three slightly different environments: in non-branching codas, in branching codas whose first segment is [x], and in codas containing a geminate [xː] or the first part of a geminate [xː]. The third analysis proposing an underlying /x/ also runs into problems. Namely, although such a solution is simpler than the previous two, it is counter-intuitive as the number of non-alternating stems containing a surface [h], e.g., *hegy* [hɛɟ] 'hill', is significantly higher than the number of alternating stems, i.e., those sometimes containing [x] and sometimes [h], e.g., *doh* ∼ *dohos* [dox] ∼ [dohoʃ] 'musty smell' ∼ 'with a musty smell', and non-alternating stems containing a surface [x], e.g., *jacht* [jɒxt] 'yacht', taken together.

A further complication is the exceptional behaviour of *H*-type segments in voice assimilation. This is what we turn to in the next section.

[7] For a detailed comparison of rule-based and constraint-based grammars see Kager (1999); cf. also Roca (1997), Hermans – van Oostendorp (1999), and McMahon (2000).

5. The behaviour of [x] with respect to voice assimilation

In Hungarian, adjacent obstruents must agree in terms of voicing (cf. Ritter 2000 and references cited there).

Of word initial consonant clusters, those that do not contain a sonorant are always voiceless throughout as in (16a);[8] even irregular initial clusters conform to this pattern, see (16b):

(16) (a) sport [ʃp] 'sports', stég [ʃt] 'landing-stage', skála [ʃk] 'scale', szpáhi [sp] 'Turkish cavalryman', sztár [st] 'leading man/lady', sztyeppe [sc] 'prairie in Russia', szkíta [sk] 'Scythian'

(b) psziché [ps] 'psyche', xilofon [ks] 'xylophone', szfinx [sf] 'sphinx'

Other morpheme-internal (intervocalic or morpheme-final) obstruent clusters are either all-voiceless as in (17a,b) or else all-voiced as in (17c,d):

(17) (a) pitypang [cp] 'dandelion', puszpáng [sp] 'boxwood', ráspoly [ʃp] 'file'; szeptember [pt] 'September', bukta [kt] 'sweet roll', kaftán [ft] 'Turkish coat', asztal [st] 'table', este [ʃt] 'evening'; kesztyű [sc] 'glove', bástya [ʃc] 'bastion'; sapka [pk] 'cap', patkó [tk] 'horseshoe', butykos [ck] 'pitcher', dafke [fk] 'obstinacy', deszka [sk] 'plank', táska [ʃk] 'bag', kocka [tsk] 'cube', bocskor [ʧk] 'moccasin'; klopfol [pf] 'beat (steak)', bukfenc [kf] 'somersault', aszfalt [sf] 'asphalt', násfa [ʃf] 'lavalier'; kapszula [ps] 'capsule', buksza [ks] 'purse'; tepsi [pʃ] 'frying-pan', taksál [kʃ] 'estimate'; nátha [th] 'cold (n)'; kapca [pts] 'foot clout', vakcina [kts] 'vaccine'; kapcsol [pʧ] 'link (v)'

(b) kopt [pt] 'Coptic', akt [kt] 'nude', szaft [ft] 'gravy', liszt [st] 'flour', test [ʃt] 'body', jacht [xt] 'yacht'; maszk [sk] 'mask', barack [tsk] 'apricot'; copf [pf] 'plaited hair'; gipsz [ps] 'gypsum', koksz [ks] 'coke'; taps [pʃ] 'applause', voks [kʃ] 'vote (n)'; also in placenames like Apc, Detk, Batyk, Recsk, Szakcs, Paks, etc.

(c) rögbi [gb] 'rugby football', azbeszt [zb] 'asbestos'; labda [bd] 'ball', Magda [gd] ⟨a name⟩, bovden [vd] 'Bowden cable', gazdag [zd] 'rich', rozsda [ʒd] 'rust'; mezsgye [ʒɟ] 'ridge'; izgul [zg] 'be excited', pezsgő [ʒg] 'champagne'; udvar [dv] 'yard', fegyver [ɟv] 'weapon', özvegy [zv] 'widow'; kobzos [bz] 'minstrel', madzag [dz] 'string', lagzi [gz] 'wedding'; habzsol [bʒ] 'devour'

(d) smaragd [gd] 'emerald', kezd [zd] 'begin', pünkösd [ʒd] 'Whitsun', kedv [dv] 'temper', edz [dz] 'train (v)'

Loanwords that originally contained an obstruent cluster of heterogeneous voicing (or happen to have a spelling suggesting one) automatically get adjusted to this pattern:

[8] In this paper, we abstract away from the behaviour of /v/; see Szentgyörgyi (2000); Ritter (2000); Siptár – Törkenczy (2000, 298–305).

(18) (a) abszolút [ps] 'absolute', obstruens [pʃ] 'obstruent', abcúg [pts] 'down with him!', abház [ph] 'Abkhaz', Buddha [th], joghurt [kh] 'yogurt'

(b) futball [db] 'football', Macbeth [gb], matchbox [ʤb] 'toy car', Updike [bd], anekdota [gd] 'anecdote', afgán [vg] 'Afghan'

In suffixed forms, stem-final voiceless obstruents become voiced if the suffix begins with a voiced obstruent (19a) and vice versa: stem-final voiced obstruents become voiceless if the suffix begins with a voiceless obstruent (19b):

(19) (a) kalap-ban [bː] 'in (a) hat', kút-ban [db] 'in (a) well', fütty-ben [ɟb] 'in (a) whistle', zsák-ban [gb] 'in (a) sack', széf-ben [vb] 'in (a) safe', rész-ben [zb] 'in part', lakás-ban [ʒb] 'in (a) flat', ketrec-ben [dzb] 'in (a) cage', Bécs-ben [ʤb] 'in Vienna'

(b) rab-tól [pt] 'from (a) prisoner', kád-tól [tː] 'from (a) bath-tub', ágy-tól [ct] 'from (a) bed', meleg-től [kt] 'from the heat', szív-től [ft] 'from (a) heart', víz-től [st] 'from water', garázs-tól [ʃt] 'from (a) garage', bridzs-től [ʧt] 'from bridge (the card game)'

This assimilation process is regressive and (right-to-left) iterative:

(20) liszt-ből [stb] → [sdb] → [zdb] 'from flour'
pünkösd-től [ʒdt] → [ʒtt] → [ʃtt] (→ [ʃt]) 'from Whitsun'

It also applies across a compound boundary (*rab*][*szolga* [ps] 'slave', lit. 'captive-servant'), across a word boundary (*nagy kalap* [ck] 'large hat') and indeed across any higher boundary as long as no pause intervenes; furthermore, as the examples in (18) show, it applies in non-derived environments as well, hence it is postlexical (but obligatory and non-rate-dependent).

Sonorants do not participate in the process: they do not voice a preceding obstruent (21a), nor do they get devoiced by a following voiceless obstruent (21b):

(21) (a) kalap-nak 'to (a) hat', kút-nak 'to (a) well', fütty-nek 'to (a) whistle' zsák-nak 'to (a) sack', széf-nek 'to (a) safe', rész-nek 'to (a) part', más-nak 'to sg else', léc-nek 'to (a) lath', csúcs-nak 'to (a) peak'

(b) szem-től 'from (an) eye', bűn-től 'from (a) sin', torony-tól 'from (a) tower', fal-tól 'from (a) wall', őr-től 'from a guard', száj-tól 'from (a) mouth'

There are two segments that behave asymmetrically with respect to this process. One is /v/ that undergoes devoicing (*szívtől* [ft] 'from (a) heart') but does not trigger voicing (*hatvan* *[dv] 'sixty'); cf. Szentgyörgyi (2000); Ritter (2000); Siptár–Törkenczy (2000, 189–205). The other is *H* that triggers devoicing (*adhat* [th] 'he may give') but does not undergo voicing before an obstruent. The usual solution for *H* is to assume that this segment is /h/ at the underlying level, specified as [− cons] (this is quite appropriate phonetically as long as [+ cons] is defined as 'constriction **in the oral cavity** at least equal to that found in fricatives') and restrict the input of voice assimilation to [+ cons, − son] segments. However, as we saw in section **2**, the glottal realization of this segment does not occur preconsonantally; what does occur is its velar realization [x]. It is this [x] that resists voice assimilation (e.g., *pechből* [xb], *[ɣb] 'out of bad luck')—but then it cannot be claimed to be specified as [− cons]. Several possibilities suggest themselves at this point, none of them very satisfactory. One would be to order the /h/-strengthening rule /h/ → [x] after voice assimilation, such that this rule, formulated as (10) above, counterfeeds voicing.[9] Another possibility would be not to restrict voice assimilation to [+ cons] segments and let /h/ undergo it (in principle, at least).[10] The solution given in Siptár–Törkenczy (2000, 79) and adopted here stipulates a filter to the effect that *[ɣ] is disallowed in Hungarian surface representations (or representations at any level, for that matter). This will do the job: we can simplify the rule of voice assimilation (by omitting [+ cons] which, without rule ordering, and especially if the underlying segment is /x/ rather than /h/, would be useless anyway), yet keep our grammar from generating *[ɣ]. Such a filter is a totally *ad hoc* device in a derivational account, but is a completely legitimate tool in an OT analysis, to which we now turn.

6. An OT analysis of the behaviour of *H*-type segments

As we have seen above, several possible analyses exist for the treatment of *H*-type segments in a derivational framework and some parts of such solutions seem to be *ad hoc* in such a theory. However, non-derivational approaches to grammar are different in this very respect: constraints penalizing marked segments or segment types form an organic part of Optimality Theory (OT).

[9] Both rules being postlexical, this ordering would have to be based on stipulation.

[10] Zsigri (1994) suggests to (do that and yet) exempt [x] from undergoing the rule by introducing the notion of 'phonetic quotations'. He points out that voiceless obstruents that are clearly non-Hungarian do not get voiced: *Bath-ba* [θb], *[ðb] 'to Bath', as if they were 'encapsulated' or surrounded by 'quotation marks'. He then claims that all Hungarian [x]-final lexical items are exactly like this example in that they refuse to be affected by Hungarian phonological rules (in particular, voice assimilation). This suggestion would be perfectly all right if [x]-final items were indeed few and clearly non-native. However, as we saw, this is far from being the case. We are therefore left with the solution proposed in the text.

Thus, if we assume an OT framework, a completely different solution might become available. This is exactly what we try to find out now.

One of the most important characteristics of OT is the Richness of the Base (RB), which means that there are no constraints on input forms. Or, to put it in a different way, any kind of input forms may occur in the Lexicon because of RB. Consequently, the constraint hierarchy should be such that it should select well-formed output forms whatever the input is. That is, even though an input form may contain segments or structures never occurring on the surface, the constraints should never allow these to surface; instead, they should prefer forms that are possible output forms in the particular language. This is a most important characteristic of OT as we will see in the following.

Let us first try to translate the generalisations of the last derivational solution into OT, i.e., into constraints instead of derivational rewrite rules. In this solution we assumed that *H*-type segments are underlyingly specified as voiceless dorsal fricatives, as in (7a), and are lenited into a placeless continuant in onset positions, as in (7b). Such a restriction may be directly translated into a positional markedness constraint prohibiting voiceless dorsal fricatives in onsets, i.e., the configuration described in (8a). The constraint is given in (22):

(22) ***Onset-x** Voiceless dorsal fricatives are prohibited in onsets (unless licensed).[11]

This constraint only prohibits a [x] in an onset but does not imply the repair strategy chosen by the language. This is done by some other constraints, the most relevant of which are given below.

(23) **MAX** Every input segment must have a correspondent in the output.

(24) **DEP** Every output segment must have a correspondent in the input.

(25) **IDplace** Corresponding input and output segments have identical specifications for place features.

(26) ***ɣ** No voiced dorsal fricatives in output forms.

The faithfulness constraint in (23), **MAX**, prohibits deletion of segments while the other such constraint in (24), **DEP**, does just the opposite, i.e., it penalizes epenthetic segments. Constraint (25) is a featural faithfulness constraint that

[11] Note that this constraint has to be such that it should allow for a geminate [x], i.e., an onset [x] must be allowed if it is licensed by the preceding coda [x]. Thus, the bracketed part has to be added so that we could interpret the constraint this way.

prohibits any kind of change—or deletion or addition—in the specification of place of articulation. Finally, the markedness constraint in (26) penalizes the occurrence of the relatively marked voiced dorsal fricative. Since such segments never occur in Hungarian surface forms, this constraint will probably be high ranking.

The following tableaux demonstrate the operation of the above constraints. As will be seen, in some cases the constraint hierarchy is able to select the optimal output form but there are some cases when it cannot do so.

(27)

UR: /dox/	*ɣ	MAX	DEP	*Onset-x	IDplace
a. ☞ dox					
b. do		∗!			
c. do.xɔ			∗!	∗	
d. doɣ	∗!				
e. doh					∗!

In tableau (27) an input with a coda /x/ is evaluated. Four of the candidates violate at least one of the constraints except for candidate (a) which, thus, wins as optimal. Candidate (d) violates the markedness constraint against voiced dorsal fricatives;[12] candidate (b) violates **MAX** since the input coda segment does not have a correspondent in the output;[13] candidate (c) violates not only the constraint prohibiting epenthesis, but also the constraint against voiceless velar fricatives in an onset. Finally, candidate (e) violates **IDplace** since the place of articulation of the input dorsal fricative is not present in the output form. Hence candidate (a) is correctly selected as optimal.

(28)

UR: /xeːt/	*ɣ	MAX	DEP	*Onset-x	IDplace
a. xeːt				∗!	
b. eːt		∗!			
c. ☞ seːt					∗
d. ɣeːt	∗!				
e. ☞ heːt					∗
f. ɛxteːt			∗!∗		

*ɣ, MAX, DEP, *Onset-x ≫ IDplace

[12] This candidate also violates **IDvoice** because the corresponding input and output segments, /x/ and /ɣ/, do not have the same specification for voice. Such a faithfulness constraint will play a significant role in the treatment of voice assimilation, or rather the explanation of the lack of voice assimilation. See section **7** for details.

[13] Note that this candidate also violates a syllable well-formedness or markedness constraint that does not allow short mid rounded vowels word finally.

Tableau (28) shows an input form (for *hét* 'seven') with a dorsal fricative in the onset, a configuration that is not permitted. Similarly to tableau (27), candidate (d) is ruled out because of the relatively marked segment it contains. Candidates (b) and (f) are excluded because of their respective violations of the constraints penalizing deletion and epenthesis, **MAX** and **DEP**, respectively.[14] The completely faithful candidate in (a) has a fatal violation of the positional markedness constraint, ***Onset-x**, and is ruled out. The remaining two candidates in (c) and (e) tie on all the constraints because they only violate the faithfulness constraint requiring identity of place features in input and output forms. Since the place of articulation is changed in them, although differently, they both have one violation. To have these two candidates as the best, **IDplace** must be dominated by all the other constraints, indicated by the solid line between **IDplace** and the dominating constraints. As candidate (e) is the actual surface form, there must be some other constraint that prefers it to the form in (c). Such a constraint, or rather a family of constraints, may be the one penalizing change in the identity of place of articulation features one by one, e.g., **IDdorsal**, **IDcoronal**, **IDlabial**, etc. Supposing that place features are privative, candidate (28e) violates **IDdorsal** only because of the deletion of the underlying dorsal feature. It does not violate the other such constraints since no features appear in the output form that were not present in the input. Candidate (28c), on the other hand, would violate not only **IDdorsal**, but **IDcoronal**, too, as the input-output specifications of the word initial fricative for the feature coronal are not identical. Thus, any candidate with a specified place of articulation in the output would lose to candidate (e), which is a placeless, i.e., glottal, segment.

(29)

UR: /ʧɛx/	*ɣ	MAX	DEP	*Onset-x	IDplace
a. ☹ ʧɛx					
b. ʧɛh					∗!
c. ʧɛ		∗!			

(30)

UR: /ʧɛx+Vk/[15]	*ɣ	MAX	DEP	*Onset-x	IDplace
a. ʧɛxɛk				∗!	
b. ☞ ʧɛhɛk					∗
c. ʧɛɛk		∗!			

[14] A candidate like [ɛ.xeːt] would violate **DEP** only once, but would also violate ***Onset-x**; it would similarly turn out to be suboptimal.

[15] The surface quality of the suffix vowel may be partially predicted on the basis of the last stem vowel, i.e., by the rules/constraints governing vowel harmony (cf. Ringen – Vago

Tableaux (29) and (30) show one of the words that surface without the underlying /x/ in coda position, i.e., the coda /x/ is deleted, but with a surface [h] in onset position when followed by a vowel initial suffix. As can be seen in (30) containing the suffixed form, candidate (c) is ruled out by **MAX** because the input stem final fricative does not have a surface correspondent. Candidate (a) has a fatal violation of ***Onset-x**, while candidate (b), the actual surface form, only violates the lower ranked **IDplace** because of the missing output dorsal specification.

Unfortunately, in tableau (29) showing the bare stem, it is not the actual surface form that is selected as optimal (this is indicated by '☹' in the tableau). Candidate (b) is ruled out because of the input dorsal specification missing in the output and candidate (c), the actual surface variant, is excluded by **MAX** because of the deletion of the /x/. It seems then that our constraint hierarchy is unable to cope with the above form and the others in its class. Note, however, that as we have already mentioned at the beginning of section **3**, *cseh* is one of the exceptional stems that behave differently from the normal /x/-final stems, e.g., *doh*, in that their /x/ is deleted in a coda. Thus, whatever kind of solution we propose, derivational or not, these stems will have to be marked as exceptional in the Lexicon.

One of the advantages of OT is that there is a way to deal with exceptional forms: they are not different from normal forms in that the constraints are not valid for them; rather the constraints are evaluated the same way but in a different hierarchy, i.e., the constraints may be reranked for exceptional lexical items. This means then that exceptions are marked for the reranking of certain constraints. This is what we turn to now.

(31) ***x** Voiceless dorsal fricatives are prohibited.

(32) ***Coda-h** Voiceless glottal fricatives are prohibited in codas.

The constraint in (31) prohibits [x] in the output. Since such segments normally occur in surface forms, it has to be ranked low, at least below **IDplace** (cf. (27)), in the case of normal stems, i.e., the ones that do not have /x/-deletion in coda position unlike the one in (29). This way, ***x** will never have an effect on the evaluation of the candidates in normal stems. The constraint in (32) is needed for independent reasons since if an input form contains a coda /h/, which may occur because of RB, then such a positional constraint

(1998); Siptár – Törkenczy (2000) and Szentgyörgyi (2000), for instance). We are not going to discuss the input quality of the vowel and whether it is part of the input or it is epenthetic.

is needed to rule such output forms out. Note that the positional markedness constraints (22) and (32) are exactly the ones mentioned in the discussion of the distribution of *H*-type segments in (8a) and (8b). *H*-deleting stems on the other hand are marked for the reranking of the constraints: if ***x** dominates **MAX** then the actual surface forms win as shown below.

(33) *H*-deleting words are marked for the reranking:
***x ≫ MAX**

(34)

UR: /ʧɛx/	*ɣ	*Onset-x	*Coda-h	*x	MAX	IDplace
a. ʧɛx				∗!		
b. ʧɛh			∗!			∗
c. ☞ ʧɛ					∗	

*x, *Coda-h ≫ MAX

Tableau (34) shows the effect of adding the positional markedness constraint (32) and the reranking shown in (33) required by an *H*-deleting stem. The fully faithful candidate in (34a) is ruled out by the reranked ***x** constraint. Candidate (b) containing a glottal fricative violates the newly added positional markedness constraint, ***Coda-h**. The actual output form only violates **MAX** because of the deleted stem final segment.

(35)

UR: /ʧɛx+töːl/	*ɣ	*Onset-x	*Coda-h	*x	MAX	IDplace
a. ʧɛx.töːl				∗!		
b. ʧɛh.töːl			∗!			∗
c. ☞ ʧɛ.töːl					∗	

Tableau (35) shows the same stem followed by a consonant initial suffix. Candidates (a) and (b) are ruled out by ***Coda-h** and ***x** respectively and thus allow candidate (c), the actual surface form, to win.

It remains to show that if a normal stem ending in a /x/ is followed by a vowel initial suffix, the underlying /x/ surfaces as [h] but it stays [x] before consonant initial suffixes.

(36)

UR: /dox+Vʃ/	*ɣ	*Onset-x	*Coda-h	MAX	IDplace	*x
a. do.xoʃ		∗!				∗
b. ☞ do.hoʃ					∗	
c. do.oʃ				∗!		

Candidates (a) and (c) are ruled out by ***Onset-x** and **MAX** respectively, while the actual surface form only violates the lower ranking **IDplace** dominated by the constraints violated by the other candidates. This way, candidate (b) is correctly selected as optimal.

(37)

UR: /dox+toːl/	*ɣ	*Onset-x	*Coda-h	MAX	IDplace	*x
a. ☞ dox.toːl						*
b. doh.toːl			*!		*	
c. do.toːl				*!		

Candidates (37b) and (37c) both violate some of the highest ranked constraints, ***Coda-h** and **MAX** respectively. This way, candidate (a), the actual surface form, is correctly allowed to be selected as optimal.

Thus we can conclude that so far we have been able to prove that supposing an underlying /x/ all kinds of forms may be accounted for if we assume that *H*-deleting stems are marked for the reranking of ***x** $\gg$ **MAX**.

Let us now turn to the problem of the Richness of the Base. As a result of RB, it is not only /x/ that may occur underlyingly, but also /h/. We have to show that even in such cases the constraint hierarchy is able to predict the correct surface forms. Then we will show how the input forms are selected by Lexicon Optimization (LO).[16]

(38)

UR: /doh/	*Onset-x	*Coda-h	MAX	IDplace	*x
a. ☞ dox				*	*
b. doh		*!			
c. do			*!		

Tableau (38) demonstrates what happens if a stem whose surface form ends in [x] contains an underlying /h/ stem finally. The fully faithful candidate in (38b) is ruled out by ***Coda-h** while candidate (38c) is excluded because of a **MAX** violation, i.e., the underlying stem final /h/ not having a surface correspondent. This way, candidate (38a) is correctly selected as optimal.[17]

(39)

UR: /doh+Vʃ/	*Onset-x	*Coda-h	MAX	IDplace	*x
a. do.xoʃ	*!			*	*
b. ☞ do.hoʃ					
c. do.oʃ			*!		

[16] Candidates violating ***ɣ** are not shown. ***ɣ** and **DEP** are left out of the tableaux unless relevant.

[17] We should bear in mind that other candidates containing a labial or a coronal fricative (**dof, *dos*, etc.) are also evaluated. However, since **IDplace** is just a shorthand for a family of constraints, **IDlabial**, **IDcoronal** and **IDdorsal**, these candidates would incur more violations than the optimal one. For details see the discussion after tableau (28).

Tableau (39) shows the same stem followed by a vowel initial suffix, where the surface quality of the vowel is determined by constraints governing vowel harmony as above. Candidate (a) containing a voiceless dorsal fricative incurs a fatal violation of ***Onset-x**, while candidate (c) violates **MAX** fatally. Again, the actual surface form in (b), which is also the fully faithful candidate in this case, is selected as optimal.

(40)

UR: /doh+toːl/	*Onset-x	*Coda-h	MAX	IDplace	*x
a. ☞ dox.toːl				∗	∗
b. doh.toːl		∗!			
c. do.toːl			∗!		

Similarly to the above, a normal stem ending in an H-type segment is predicted to behave the same way before a consonant initial suffix if we assume an underlying /h/ stem finally as shown in (40). Since the violations of the highest ranked constraints do not depend on what the underlying stem final segment is, they are exactly the same as in tableau (37) containing the same stem with underlying /x/. Thus, the same output form, the actual surface form, is selected as optimal once again.

(41)

UR: /heːt/	*Onset-x	*Coda-h	MAX	IDplace	*x
a. xeːt	∗!			∗	∗
b. eːt			∗!		
c. ☞ heːt					

Tableau (41) contains a word starting with an H-type segment, which always surfaces as a placeless (glottal) continuant. This prediction is borne out by the constraint hierarchy in the tableau since candidate (a) containing a dorsal fricative and the unfaithful candidate (b) containing no correspondent of the input /h/ violate ***Onset-x** and **MAX** respectively, allowing candidate (c) to win.

Let us now turn to exceptional H-deleting stems like *cseh* 'Czech' and see how they behave in various environments.

(42)

UR: /ʧɛh/	*Onset-x	*Coda-h	*x	MAX	IDplace
a. ʧɛx			∗!		∗
b. ʧɛh		∗!			
c. ☞ ʧɛ				∗	

In tableau (42), the bare H-deleting stem is shown without any suffix. In such cases, the stem final H should be dropped and this is exactly what we see in the optimal candidate, (42c). The other two candidates are ruled out by higher ranked constraints, namely ***Coda-h** and ***x**. We must bear in mind that, as we have noted above, *H*-deleting stems are lexically marked for the reranking of ***x** $\gg$ **MAX**. As a result of this reranking, the violation of **MAX** by the optimal candidate is less serious than either the violation of ***Coda-h** by (42b) or that of ***x** by candidate (42a).

(43)

UR: /ʧɛh+Vk/	*Onset-x	*Coda-h	*x	MAX	IDplace
a. ʧɛ.xɛk	∗!		∗		∗
b. ☞ ʧɛ.hɛk					
c. ʧɛ.ɛk				∗!	

The same stem behaves differently if followed by a vowel initial suffix as shown in (43). In such environments the stem final *H* is syllabified into the onset of the next syllable and thus surfaces as [h] as in candidate (43b), the optimal output form. Candidate (43a) violates ***Onset-x** and ***x** because of the dorsal continuant syllabified in the onset while candidate (43c) violates **MAX** because of the input /h/ being unparsed.

(44)

UR: /ʧɛh+töːl/	*Onset-x	*Coda-h	*x	MAX	IDplace
a. ʧɛx.töːl			∗!		∗
b. ʧɛh.töːl		∗!			
c. ☞ ʧɛ.töːl				∗	

Finally, tableau (44) demonstrates what happens to the same type of stem if a consonant initial suffix is added. As can be seen, the hierarchy selects the actual output form as optimal again, ruling out candidates (44a) and (44b) because of the violations of ***x** and ***Coda-h**, respectively. In comparison, the optimal candidate only violates the relatively low ranked **MAX** constraint because of the /h/ not having a surface correspondent.

Thus, we can conclude that whatever the input, whether /h/ or /x/, the constraint hierarchy always selects the actual surface forms as optimal both for normal stems and *H*-deleting stems, provided that the latter are marked for the reranking of ***x** $\gg$ **MAX**.

The next problem to be discussed is that of selecting the best input form, i.e., the one closest to the optimal output form. This is performed by Lexicon Optimization. Since the words ending in an *H*-type segment may have a surface [x], [h]—or nothing in the case of *H*-deleting stems—, we have to use the context sensitive version of LO as discussed in Inkelas (1994), which compares the violations for all kinds of possible environment types and ranks the different input forms accordingly.

Let us first re-examine normal stems ending in an *H*-type segment concerning the violations depending on the input forms.

(45)

UR's	Sample environments	*Onset-x	*Coda-h	MAX	IDplace	*x
a. /dox/	dox					*
	do.hoʃ				*	
	dox.toːl					*
b. /doh/	dox				*	*
	do.hoʃ					
	dox.toːl				*	*

The tableau in (45) is a summary of the violations incurred by the winning candidates generated from underlying /dox/ or /doh/ in various environments, i.e., it repeats the relevant candidates (27a), (36b), (37a), (38a), (39b) and (40a). As can be seen in (45a), the optimal candidates incur one violation of **IDplace** in the three logically possible environment types: word final, prevocalic and preconsonantal. It is so because it is only in the prevocalic environment, i.e., when the surface segment is [h], that the place of articulation of the underlying /x/ has to be changed under the pressure of the higher ranked constraints. On the other hand, in (45b), the phonetically identical optimal output candidates incur two violations of **IDplace** as the underlying /h/ has to be changed into [x] both in word final and preconsonantal position because of the high ranked contextual markedness constraint ***Coda-h**. As a result of this we can conclude that it is in the case of (45a), i.e., underlying /dox/, that the optimal output candidates incur the fewest violations of the constraint hierarchy proposed. Thus, the alternation sensitive version of LO selects /dox/ as the optimal input form on the basis of the optimal output forms. Of course, the same is true of all normal stems ending in an *H*-type segment.

Let us turn to the exceptional stems and see what the prediction is if the reranking of ***x** $\gg$ **MAX** is performed.

(46)

UR's	Sample environments	*Onset-x	*Coda-h	*x	MAX	IDplace
a. /ʧɛx/	ʧɛ				*	
	ʧɛ.hɛk					*
	ʧɛ.töːl				*	
b. /ʧɛh/	ʧɛ				*	
	ʧɛ.hɛk					
	ʧɛ.töːl				*	

As tableau (46) suggests the situation is different from that of normal stems: in *H*-deleting stems the number of **MAX** violations is the same for output candidates in the case of the two underlying forms but while there is no violation of **IDplace** if the UR ends in a /h/, there is one violation if it ends in /x/. As a result, the UR ending in /h/, /ʧɛh/ is selected as optimal over /ʧɛx/.

The last type of stem is non-alternating: these stems contain an *H*-type segment either in initial or medial position where no alternation resulting from suffixation is possible. In these cases, LO always selects the input identical to the non-alternating output form. We demonstrate this with two examples below, *hét* [heːt] 'seven' and *jacht* [jɒxt] 'yacht'.

(47)

a. /heːt/	*Onset-x	*Coda-h	MAX	IDplace	*x
heːt					
b. /xeːt/					
heːt				*	

Since in (47b) the output candidate violates **IDplace** because of the discrepancy between the place features of the input /x/ and its output correspondent [h] while the optimal candidate in (47a) does not incur any violations, LO selects (47a) as the lexical representation of the stem.

(48)

a. /jɒht/	*Onset-x	*Coda-h	MAX	IDplace	*x
jɒxt				*	*
b. /jɒxt/					
jɒxt					*

In contrast to (47), where a non-alternating stem has an underlying /h/, (48) displays a non-alternating stem which has underlying /x/ according to LO. Both the output forms in (48a) and (48b) violate ***x** but only (48a) incurs a violation of **IDplace** since underlying /h/ corresponds to surface [x] there. Thus, the UR of (48b) is selected as the lexical representation for this stem.

In conclusion to this section it can be said that the OT approach to the problem provides us with a solution dispreferred by the derivational approach, namely that both underlying /x/ and /h/ can be found in Hungarian. Note that this is not an arbitrary stipulation but rather a consequence of the application of Lexicon Optimization to phonetically identical optimal candidates arising from different underlying forms.

7. Voice assimilation and *H*-type segments in OT

In this section, we are going to show how the above OT solution also provides valuable insight to the behaviour of *H*-type segments with respect to voice assimilation.

Recall from the previous section that a constraint penalizing voiced dorsal fricatives, ***ɣ**, is necessary to rule out candidates containing this segment. As we will see below such a constraint is crucial for the analysis of voice assimilation in terms of OT.

We will base our account on Petrova et al's (2001) analysis of voice assimilation and other laryngeal phenomena. According to their approach, Hungarian voice assimilation can be described by the constraints in (49)–(53) if they are ranked as in (54) except for /v/ and *H*-type segments, which behave asymmetrically: the former undergoes voice assimilation but does not trigger it while the latter (set) does trigger it but does not undergo it.

(49)	**IDvoice**	Corresponding input and output segments have identical specifications for voice.
(50)	**IDpreson voice**	Corresponding input and output segments before a sonorant have identical specifications for voice.
(51)	**IDwf voice**	Corresponding segments in word final position have identical specifications for voice.
(52)	**Share**	Obstruents in a cluster share their voice specifications.
(53)	***voice**	No voiced obstruents.

(54) **Share**, **IDpreson voice** $\gg$ **IDwf voice** $\gg$ **IDvoice** $\gg$ ***voice**

The highest ranked **Share** constraint ensures that there is always voice assimilation in obstruent clusters while the positional faithfulness constraints govern the direction of the assimilation: it is always the segment before a sonorant or a word final segment, i.e., the rightmost obstruent in the cluster, that must be faithful to its underlying voice specification and thus triggers regressive assimilation. **IDvoice** and especially ***voice** will not play a significant role since they are ranked too low in the hierarchy to interfere with the selection of the optimal candidate.

First, let us see how words containing *H*-type segments in a cluster behave according to Petrova et al's constraint hierarchy and then decide where our constraints should be placed in the hierarchy, still observing the dominance relations determined so far.

(55)

UR: /dox+bɒn/[18]	Share	IDpreson voice	IDwf voice	IDvoice	*voice
a. dox.bɒn	*!				**
b. doh.bɒn	*!				**
c. ☹ doɣ.bɒn				*	* * *
d. dox.pɒn		*!		*	*

Tableau (55) contains a normal stem ending in an *H*-type segment followed by a voiced obstruent initial suffix, i.e., a potential cluster for voice assimilation. The input form is the one chosen by the alternation sensitive Lexicon Optimization as selected in (45) above. Tableau (55) shows that the constraint hierarchy that gives the right predictions for obstruent clusters in general does not pick the actual surface form, (55a), as optimal. Since this candidate violates one of the highest ranking constraints, there must be a constraint dominating all the ones in (55) to rule out the incorrect forms.

We have already seen that there is independent evidence for the two markedness constraints ***ɣ** and ***Coda-h**. If we assume that these constraints, which were shown to be high ranking among those responsible for the distribution of *H*-type segments, are ranked above **Share**, then the correct result is achieved as indicated in (56).[19]

(56)

UR: /dox+bɒn/	*ɣ	*Coda-h	Share	IDpreson voice	IDvoice	*voice
a. ☞ dox.bɒn			*			**
b. doh.bɒn		*!	*			**
c. doɣ.bɒn	*!				*	* * *
d. dox.pɒn				*	*!	*

*ɣ, *Coda-h ≫ Share

In this modified tableau we can see the result of adding the two markedness constraints to Petrova et al's (2001) hierarchy in an undominated position. Candidates (56b) and (56c) violate these constraints and are thus ruled out.

[18] The harmonising suffix vowel is most probably [− back] underlyingly since it always occurs with that value on the surface when used as a root, e.g., *bennem* [bɛnːɛm] 'in me'. For the sake of simplicity we use the back vowel variant as underlying after back stems and the front vowel variant after front stems in the tableaux. See Ringen – Vago (1998) for details.

[19] We may also assume that all the constraints used to predict the distribution of *H*-type segments dominate all the constraints governing voice assimilation although it is not necessary. The only ranking between the two blocks of constraints we have to assume so far is ***ɣ**, ***Coda-h** ≫ **Share**.
Note furthermore that **IDwf voice** is not shown in the tableau since the relevant cluster is not word final; thus the constraint cannot play a role.

Candidates (a) and (d) violate **Share** and **IDpreson voice** respectively but since these two constraints are not crucially ranked with respect to each other, these violations count as equal. Decision between the two candidates is passed on to the next constraint, **IDvoice**, which is fatally violated by candidate (d) but satisfied by (a), the fully faithful candidate, the actual surface form.

Let us now turn to *H*-deleting stems and see how they behave in this respect. Since the stem final *H* is deleted before consonants and a pause, it will not assimilate to a following obstruent, e.g., *cseh* ∼ *csehben* [ʧɛ] ∼ [ʧɛbɛn].

(57)

UR: /ʧɛh+bɛn/	*ɣ	*Coda-h	*x	MAX	Share	IDpreson voice	IDvoice	*voice
a. ʧɛx.bɛn			∗!		∗			∗
b. ʧɛh.bɛn		∗!			∗			∗
c. ʧɛɣ.bɛn	∗!						∗	∗∗
d. ʧɛx.pɛn			∗!			∗	∗	
e. ☞ ʧɛ.bɛn				∗				∗

*ɣ, *Coda-h ≫ MAX

In (57) we have shown some of the rest of the proposed constraints relevant for *H*-type segments. We have already seen that *H*-deleting stems are marked for the reranking of ***x ≫ MAX**. Assuming this reranking and that these constraints together with ***ɣ** and ***Coda-h** dominate the ones responsible for voice assimilation,[20] we get the desired result. Candidates (57a–d) violate the three highest ranked constraints, the ones responsible for the distribution of *H*-type segments. Since candidate (57e) only violates **MAX** (and the lowest ranked ***voice**), it is correctly selected as optimal.

Finally we have to show that the hierarchy established above does not interfere with an *H*-type segment triggering regressive voice assimilation. This is what can be seen in (58).

(58)

UR: /haːz+hoz/	*ɣ	*Onset-x	MAX	Share	IDpreson voice	IDvoice	*voice
a. haːz.hoz				∗!			∗∗
b. ☞ haːs.hoz						∗	∗
c. haːz.xoz		∗!		∗			∗∗
d. haːz.ɣoz	∗!				∗	∗	∗ ∗ ∗
e. haː.zoz			∗!				∗∗

[20] **MAX** does not have to dominate **Share** according to (56) but in general we can claim that **MAX ≫ Share** since in no other case are segments deleted to avoid violating Share, not even in the case of the other segment exceptional for voice assimilation, /v/.

In the input in (58), a stem final voiced obstruent is followed by an underlying /h/ as determined by LO. Since /h/ is rightmost in the cluster, it will trigger regressive assimilation. This prediction is actually borne out in the tableau. Candidates (c) and (d) violate ***Onset-x** and ***ɣ**, two markedness constraints, respectively, and these violations are fatal. Candidate (a) is fully faithful to the input and thus contains an obstruent cluster with heterogeneous voice, a clear violation of **Share**. Candidate (e), on the other hand, violates **MAX** because of the unparsed /h/ of the suffix, also a fatal violation. This way, candidate (b) is correctly allowed to win.

8. Conclusion

To sum up, it can be said that although the three derivational analyses proposed in the first part of the paper have the descriptive power to explain the distribution of *H*-type segments, they all suffer from similar shortcomings addressed in section **4**. On the other hand, the OT analysis presented in sections **6** and **7** can account for all the phenomena involving *H*-type segments, the [x] ∼ [h] alternations, *H*-deleting stems, and the asymmetric behaviour of *H*-type segments in voice assimilation as well, without having to make unreasonable stipulations. This is achieved by a hierarchy of constraints made up of both general, e.g., ***x** and ***ɣ**, and positional markedness constraints, e.g., ***Coda-h** and ***Onset-x**, and some faithfulness constraints, e.g., **IDplace**, **MAX** and **DEP**. The exceptional class of *H*-deleting stems can simply be treated by a reranking of two constraints, ***x** $\gg$ **MAX**, one of the most straightforward ways of dealing with exceptionality in OT. Also, the role of *H*-type segments in voice assimilation is described without stipulating any new constraints: the hierarchy proposed by Petrova et al. (2001) for the treatment of voice assimilation in a number of languages including Hungarian amended by the constraints suggested by us to account for the distribution of *H*-type segments is sufficient to deal with this problem, too.

References

Archangeli, Diana – Terence D. Langendoen 1997. Optimality Theory: an introduction. Blackwell, Cambridge MA & Oxford.

Harris, John 1990. Segmental complexity and phonological government. In: Phonology 7: 255–300.

Harris, John 1997. Licensing Inheritance: an integrated theory of neutralisation. In: Phonology 14 : 315–70.

Hermans, Ben – Marc van Oostendorp 1999. The derivational residue in phonological Optimality Theory. John Benjamins, Amsterdam.

Inkelas, Sharon 1994. The consequences of optimization for underspecification. Rutgers Optimality Archive. [ROA-40-1294, HTTP://RUCCS.RUTGERS.EDU/ROA.HTML]

Kager, René 1999. Optimality Theory. Cambridge University Press, Cambridge.

McCarthy, John 2002. A thematic introduction to Optimality Theory. Cambridge University Press, Cambridge.

McCarthy, John – Alan Prince 1995. Faithfulness and reduplicative identity. In: Jill Beckman – Laura Dickey – Susan Urbanczyk (eds) UMOP 18: Papers in Optimality Theory. GLSA, Amherst MA.

McMahon, April 2000. Change, chance and optimality. Oxford University Press, Oxford.

Petrova, Olga – Rosemary Plapp – Catherine Ringen – Szilárd Szentgyörgyi 2001. Constraints on voice: and OT typology. Paper presented at the Utrecht Typology Workshop.

Prince, Alan – Paul Smolensky 1993. Optimality Theory: constraint interaction in generative grammar. Technical Report of the Rutgers Center for Cognitive Science.

Ringen, Catherine O. – Robert M. Vago 1998. Hungarian vowel harmony in Optimality Theory. In: Phonology 15 : 393–416.

Ritter, Nancy A. 2000. Hungarian voicing assimilation revisited in Head-Driven Phonology. In: Gábor Alberti – István Kenesei (eds) Approaches to Hungarian 7. Papers from the Pécs Conference, 23–49. JATEPress, Szeged.

Roca, Iggy (ed.) 1997. Derivations and constraints in phonology. Clarendon Press, Oxford.

Roca, Iggy – Wyn Johnson 1999. A course in phonology. Blackwell, Cambridge MA & Oxford.

Siptár, Péter – Miklós Törkenczy 2000. The phonology of Hungarian. Oxford University Press, Oxford.

Szentgyörgyi, Szilárd 2000. A magyar /v/ és a zöngésségi hasonulás az optimalitáselméletben [Hungarian /v/ and voicing assimilation in Optimality Theory]. In: Tibor Szécsényi (ed.) LingDok I. Nyelvész-doktoranduszok dolgozatai [LingDoc I. Papers by Ph.D. students of linguistics], 42–56. JATEPress, Szeged.

Zsigri, Gyula 1994. Magyar mássalhangzószabályok [Rules for Hungarian consonants]. Ph.D. dissertation, JATE, Szeged.

Address of the authors:

Péter Siptár
Research Institute for Linguistics
Hungarian Academy of Sciences
Benczúr utca 33.
H–1068 Budapest
Hungary
siptar@nytud.hu

Szilárd Szentgyörgyi
Dept. of English and American Studies
University of Veszprém
Egyetem utca 10.
H–8200 Veszprém
Hungary
szentsz@almos.vein.hu