Péter Siptár

Degemination in Hungarian

The traditional insight concerning Hungarian degemination is that geminates do *not* occur in this language (i) word initially, or (ii) flanked by another consonant on either side.[1] In other words, the occurrence of geminates is only possible (i) intervocalically (e.g., *állat* 'animal', *áll-ok* 'I stand', *áll Attila* 'Attila stands') and (ii) utterance finally (i.e., before a pause) if preceded by a vowel (e.g., *áll* 'stand'). The latter type is degeminated, however, if a consonant follows, irrespective of whether that consonant comes from synthetic (or Level I) suffixation (e.g., *áll-t* 'he/she stood', *áll-tam* 'I stood'), analytic (or Level II) suffixation (e.g., *áll-hat* 'may stand'), compounding (*áll-kapocs* 'jawbone') or even from a different word (*áll Tamás* 'Tom stands'). However, this traditional view is oversimplified and has to be revised, to be at least observationally adequate, in various ways.[2] This revision (as well as an analysis of the issue of degemination) is the topic of the present paper.

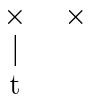
In a detailed study of degemination in Hungarian, Nádasdy (1989) distinguishes UNDERLYING vs. DERIVED geminates and LEFT-FLANKED vs. RIGHT-FLANKED geminates. Within the class of derived geminates he further distinguishes what is usually referred to as TRUE (assimilation-based) vs. FAKE (juxtaposition-based) geminates (where the former type involves linked structure as in *lábbal* 'with leg' and the latter type involves a sequence of identical short consonants across an analytic boundary as in *lábban* 'in leg'). The following discussion is partly based on Nádasdy's data and classification but the actual analysis differs from his in some important respects.

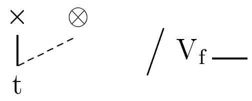
[1] On the small overall functional load of geminate consonants in Hungarian, cf. Obendorfer (1975).

[2] A point of minor significance concerns the actual examples in this paragraph rather than the issue of degemination. In a number of lexical items there is free variation between short and geminate consonants; one of the most characteristic combinations where this holds is /aː/ followed by /lː/ as in *áll* 'chin', *áll* 'stand', *állat* 'animal', *istálló* 'stable', *szakáll* 'beard', *száll* 'fly' (verb), *váll* 'shoulder', *vállal* 'undertake', *vállalat* 'company', etc.

Right-flanked underlying geminates behave roughly in the way described above, except that across word boundary degemination is optional and varies in terms of speech style and boundary strength (cf. Dressler and Siptár 1989): the "stronger" the boundary and/or the more formal the register, degemination is the less likely to apply.

Left-flanked underlying geminates do not normally occur since no morpheme *begins* with a geminate consonant. There are two possible candidates for morphemes *consisting of* a geminate consonant: comparative *-bb* and past tense *-tt.* The former hardly ever occurs in a degemination context: it is a 'Type A' suffix[3] (e.g., *nagy-obb* 'bigger', *csúnyá-bb* 'uglier') that, however, exceptionally "drops" its unstable vowel in a handful of lexicalised forms: *külön-b* 'superior', *idős-b* 'elder', and *nemes-b* 'nobler'; also in some forms containing the verbalising suffix *-ít*: *helyes-b-ít* 'rectify', *öreg-b-ít* 'enhance', *súlyos-b-ít* 'aggravate'. With respect to the past tense suffix, Siptár and Törkenczy (S&T, section 8.1.4.4) suggest that it exhibits degemination effects without actually undergoing degemination. In particular, they suggest that this suffix is a /t/ whose root node is underlyingly associated to a *single* timing slot followed by an *empty* timing slot, i.e., a timing slot devoid of melodic content (see (1a)). They further assume that a rule of *t*-spread applies to this configuration *if a full vowel precedes it* (see (1b), where $\mathrm{V_f}$ = full (nonempty) vowel). Thus, in a form like *fal-t* 'he/she devoured', a geminate never occurs in the first place, hence there is nothing to degeminate.

(1) a. 

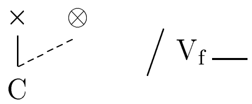
b. 

Another type of suffix showing degemination effects without actually undergoing degemination is the set of ALTERNATING *v*-SUFFIXES whose initial consonant copies the stem-final consonant, if any, and otherwise surfaces

[3] Siptár & Törkenczy (2000) (henceforward S&T) distinguish two types of unstable-vowel-initial suffixes (in traditional terms, suffixes requiring a linking vowel in some contexts). Type A includes suffixes (like plural *-k*) whose initial unstable vowel is only unrealised when they are added to a vowel-final stem; a phonetically realised vowel is always present after a consonant-final stem, regardless of the identity of the stem-final consonant. Type B, on the other hand, whose typical instance is accusative *-t,* exhibits *phonotactically motivated* vowel–zero alternation, i.e., the unstable vowel is phonetically unrealised iff the suffixal consonant can syllabify as (part of) a well-formed coda. This means that no linking vowel appears after vowels and after stem-final consonants with which the suffixal consonant can form a branching coda. See section 8.1.2.2 of S&T for details.

as [v] (instrumental *-val/vel*: *só-val* 'with salt', *méz-zel* 'with honey', and translative *-vá/vé*: *só-vá* '(turn) into salt', *méz-zé* '(turn) into honey').[4] This case (if it did involve degemination) would be that of a left-flanked derived true geminate: *domb-bal* [mb] 'with hill', *vers-sel* [rš] 'with poem', *lánc-cá* [nt^s] '(turn) into chain', *férj-jé* [rj] '(turn) into husband'. The analysis S&T offer for this case involves the generalisation of *t*-spread into a rule of C-SPREAD (see (2)) that applies in e.g., *méz-zel* 'with honey', etc. but not in *domb-bal* 'with hill', etc., giving the desired degemination effect.

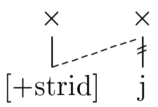
(2) **C-spread**



Let us now consider if this treatment can be extended to other instances of left-flanked derived true geminates as well.

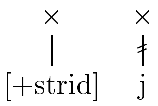
Take for instance the rule of full assimilation (STRIDENT *j*-ASSIMILATION, as proposed in section 7.2.1 of S&T) whereby a sequence of strident consonant + /j/ emerges as a long strident consonant, e.g., *hozzon* /hoz-j-on/ → [hozːon] 'let him/her bring':

(3)



Suppose we simplify this rule so that it becomes just a delinking rule and let the rule of C-spread (2) apply to its output in a case like *hozzon.* In addition to the improvement that this move represents with respect to the form of the rule, it has the side effect that in a case like *rajzzon* /rajz-j-on/ → [rɔjzon] 'let it swarm' we get the degemination effect for free:

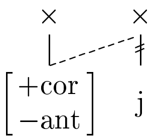
(4) **Strident *j*-assimilation**



[4] There are NON-ALTERNATING *v*-SUFFIXES as well (such as *-van/ven* '-ty': *hatvan* 'sixty', deverbal noun-forming *-vány/vény*: *lát-vány* 'sight', deverbal adverb-forming *-va/ve*: *lop-va* 'stealthily') which are likewise [v]-initial after vowel-final stems, but whose initial /v/ is unchanged even after consonant-final stems.

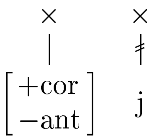
Another rule that can be reformulated in the same manner is PALATAL *j*-ASSIMILATION as in *báty-ja* [tʸː] 'his brother', *hagy-ja* [dʸː] 'he leaves it/let him leave it', *hány-ja* [nʸː] 'he throws it/let him throw it':

(5)



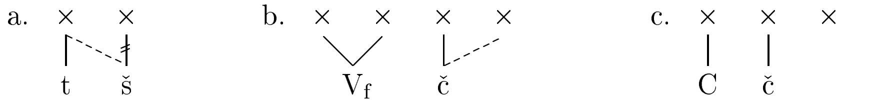
This rule also applies to the output of palatalisation in cases like *lát-ja* [tʸː] 'he sees it', *ad-ja* [dʸː] 'he gives it/let him give it'. Again, if we omit the spreading part, we get the degemination effect in *tart-ja* [rtʸ] 'he holds it', *hord-ja* [rdʸ] 'he leaves it/let him leave it':

(6) **Palatal *j*-assimilation**



Similarly, it would be a good idea to let the difference between *fűts* [čː] 'heat!' and *önts* [č] 'pour!' fall out automatically. In principle, this could be done as follows:[5]

(7)



a. (i.e., [tš]→[č]) b. (e.g., *fűts* [füːčː]) c. (e.g., *önts* [önč])

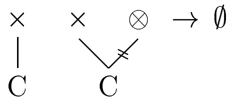
If all other cases falling under the rubric left-flanked derived true geminate turned out to show automatic degemination effects without an actual rule of degemination, it would be worth following the above line of analysis (even

[5] S&T propose that the imperative of *t*-final verbs is not derived via a rule of *t*-palatalisation (followed by (4)), as had been assumed at least since Vago (1980), but rather via a rule of *j*-obstruentisation (into [š]). Hence, rather than /tj/ → čj → [čː], the derivation—in cases like *taníts* 'teach!'—is now assumed to be /tj/ → tš → [čː]. See section 7.2.1 in S&T for details.

though it would entail some loss of generality in another part of the grammar, in particular, it would prevent us from treating the coalescence of *tš* into [čː] as an instance of the general rule of Palatalisation, see section 7.2.3 of S&T).

However, there is at least one further type of case that can by no means be analysed without a degemination rule. The output of voice assimilation[6] may or may not come out as a geminate (it does if the two segments previously only differ in terms of voicing: *adta* [ɔtːɔ] 'he gave it'; it does not otherwise: *dobta* [doptɔ] 'he threw it'). The "merger of all class nodes dominating identical material" that is involved in *adta* is an automatic OCP-effect, not a rule—hence we cannot manipulate it in a way similar to what we did above (i.e., we cannot factor out the spreading part from a spreading-cum-delinking complex and constrain it in terms of left context). Thus, we need a degemination rule for cases like *küldte* [kültɛ] 'he sent it': ldt → ltː → lt. That rule can be informally stated as in (8):

(8) **Degemination I**



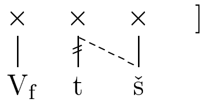
Once we need this rule anyway, we can derive the *taníts/önts*-type cases as was suggested in footnote 5 above (i.e., tj → tš → čː) and apply (8) to the output where appropriate.

Turning to right-flanked derived true geminates as in *üsd* /üt-j-d/ [üžd] 'hit it!', one way to let the degemination effect fall out automatically would be, again, to take the full assimilation rule that produces the intermediate [šː] as in [üšː-d] (*t*-assimilation, (9),[7] see section 7.2.1 in S&T) and simplify it by omitting the spreading part as in (10):

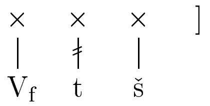
[6] This problem is by no means particular to voice assimilation. It arises in all cases where a spreading operation involving a single feature or a single class node leads to complete identity (hence to a derived true geminate) by accident, as it were.

[7] The morphological bracket] is included in this rule in order to restrict its application to *tš* sequences that derive from /tj/ in imperatives (to the exclusion of other *tš* sequences as e.g., in *hátsó* 'posterior' or *vétség* 'offence').

(9) ***t*-assimilation**

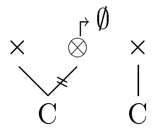


(10) ***t*-deletion**



However, for this idea to work, a separate C-spread rule would be required that — unlike (2) — (i) spreads the C *leftwards,* (ii) has a *negative* condition ("unless another consonant follows") rather than a positive environment, and (iii) is *not* independently motivated. But the whole attempt is superfluous anyway since the mirror image of (8) will be independently needed to handle right-flanked underlying geminates as in *hall-gat* [hɔlgɔt] 'listen' etc. Therefore, we leave (9) as it is and formulate the following rule for all right-flanked true geminates, whether underlying or derived:

(11) **Degemination II**



The next question is whether we should collapse (11) with (8) into a single mirror-image rule schema ("Delete one of the two ×'s of a true geminate if it is *preceded or followed* by another consonant") or not. The answer is in the negative since (11) applies postlexically and — as was mentioned above — it shows optionality effects across a word boundary, whereas (8) is always strictly obligatory (as long as true geminates are concerned).

Let us now consider fake geminates (i.e., sequences of identical consonants arising across analytic morpheme boundaries). Note first of all that — with the possible exception of geminate affricates as in *kulcscsomó* 'bunch of keys' and unlike geminate vowels as in *kiirt* 'exterminate' — fake geminate consonants surface phonetically as if they were true geminates. This means that at some point they will undergo merger (one which is either OCP-driven or rule-based, depending on one's general assumptions).

That merger can take place either too early or too late: if it takes place before (postlexical) degemination is considered for application, the difference between the behaviour of true and fake geminates may become inexpressible; if, on the other hand, merger is later than degemination, it may be difficult to refer to adjacent identical consonants that do not form a linked structure (coindexing is one possibility but not a very pleasant one).

Consider the following data (partly based on Nádasdy 1989):

(12) a. **Left-flanked fake geminates**

'Obligatory' degemination if the flanking consonant is an obstruent

koszt-tól	[kos**t**ol]	'from food'	(analytic suffix)
direkttermő	[dirɛk**t**ɛrmöː]	'a type of vine'	(compound)
lakj jól	[lɔk**j**oːl]	'eat enough!'	(phrase)

'Optional' degemination if the flanking consonant is a nasal

comb-ból	[tˢom**b**(ː)ol]	'from thigh'	(analytic suffix)
csonttányér	[čon**t**(ː)aːnʸeːr]	'bone plate'	(compound)
tank körül	[tɔŋ**k**(ː)örül]	'around tank'	(phrase)

'No degemination' if the flanking consonant is a liquid

sztrájk-kor	[straj**k**ːor]	'during a strike'	(analytic suffix)
talppont	[tɔl**p**ːont]	'foot-end'	(compound)
szerb bor	[sɛr**b**ːor]	'Serbian wine'	(phrase)

b. **Right-flanked fake geminates**

'Obligatory' degemination if the flanking consonant is an obstruent

kisstílű	[ki**š**tiːlüː]	'petty'	(compound)
olasz sztár	[olɔ**s**taːr]	'Italian (film) star'	(phrase)

'Optional' degemination if the flanking consonant is a nasal

őssmink	[öː**š**(ː)miŋk]	'proto-make-up'	(compound)
kész sznob	[keː**s**(ː)nob]	'a perfect snob'	(phrase)

'No degemination' if the flanking consonant is a liquid

széppróza	[seː**p**ːroːzɔ]	'prose fiction'	(compound)
ügyes srác	[üdʸɛ**š**ːraːtˢ]	'smart boy'	(phrase)

The expressions 'obligatory', 'optional' and 'no degemination' appear in quotation marks in (12) since we want to claim that there is a continuous gradient of optionality here in which 'most likely', 'less likely' and 'least likely' would be more appropriate labels. The type of degemination we are considering is simply an optional process whose likelihood covaries with the type of the flanking consonant as indicated.

The question is whether the phenomenon displayed in (12) is a postlexical phonological process or rather part of phonetic interpretation. An argument that supports the latter option is that the merger of fake geminates into true ones is most probably a phonetic issue and — unless we want to formulate a deletion rule referring to (coindexed) identical segments — the earliest point where this simplification process can be stated in terms of linked structures is after that merger has taken place. Therefore, we will assume the following two statements as part of the phonetic implementation module of the grammar of Hungarian:

(13) **Long Consonant Formation**

Merge a sequence of two identical short consonants into a single long consonant (applies in all speech styles/tempos with respect to consonants other than affricates; applies to affricates in fast/casual speech only)

(14) **Degemination III**

Optionally realize a long consonant as short if it is flanked by another consonant (applies with decreasing likelihood when the flanking consonant is (i) an obstruent, (ii) a nasal, or (iii) a liquid)

In this paper, we have proposed three different degemination rules, applying at word level, postlexically, and in the phonetic implementation module, respectively. (8) is the word level rule that applies obligatorily to all left-flanked true geminates that emerge from the lexical phonology as such, irrespective of the identity of the flanking consonant and of the morphological make-up (underlying vs. derived) of the geminate itself. Instances of this process are cases like *önts* [ön**č**] 'pour!'[8] and *hordtam* [hor**t**ɔm] 'I carried'

[8] Note that cases like *öntse* [önčɛ] 'let him pour it' and *öntsd* [önǰd] ~ [önžd] ~ [ȫːžd] 'pour it!' also belong here, i.e., it does not matter whether the geminate is followed by nothing, a vowel, or a consonant; what is important is the left-hand consonant, *n* in this case, that is the necessary and sufficient condition for degemination to apply.

where degemination is directly fed by palatalisation in the first example and by voice assimilation in the second.[9]

The postlexical rule is (11) that applies obligatorily within words and optionally in phrasal domains (with decreasing likelihood across increasingly "stronger" syntactic boundaries and in increasingly formal speech styles). However, the rule is insensitive to the identity of the flanking consonants and to whether the geminate is underlying or derived. Instances of this process include *hallgat* [hɔ**l**gɔt] 'listen', *üsd* /üt-j-d/ [ü**ž**d] 'hit it!', *adj neki* [ɔ**d**$^{\mathbf{y}}$nɛki] 'give him!', *evett banánt* [ɛvɛ**d**bɔnaːnt] 'he ate some bananas'.

Finally, the phonetic rule is (14) that applies optionally and targets — primarily — long consonants that are (phonologically) fake geminates. The gradience of optionality is as stated in (14); examples appear in (12) above.

REFERENCES

Dressler, Wolfgang U. and Péter Siptár. 1989. Towards a natural phonology of Hungarian. Acta Linguistica Hungarica 39 : 29–51.

Nádasdy, Ádám. 1989. The exact domain of consonant degemination in Hungarian. Hungarian Papers in Phonetics 21 : 104–107.

Obendorfer, Rudolf. 1975. The ambiguous status of Hungarian long consonants. Lingua 36 : 325–336.

Siptár, Péter and Miklós Törkenczy. 2000. The Phonology of Hungarian. Oxford: Clarendon Press/Oxford University Press.

Vago, Robert M. 1980. The Sound Pattern of Hungarian. Washington, D.C.: Georgetown University Press.

[9] Recall that a number of cases that are traditionally analysed as degemination were reinterpreted here as *lack of gemination.* The major cases include (i) past-tense verb forms like *kap-**t**-a* 'he got it' and *fal-**t*** 'he devoured', (ii) noun forms involving alternating *v*-suffixes like *domb-bal* [dom**b**ɔl] 'with hill', (iii) imperatives of sibilant-final verbs like *rajzzon* /rajz-j-on/ → [rɔj**z**on] 'let it swarm', and (iv) verb forms, both indicative and imperative ones, involving Palatal *j*-assimilation as in *tartja* [tɔr**t**$^{\mathbf{y}}$ɔ] 'he holds it'.