Surface geminates (dageš forte) in Israeli Hebrew

Shmuel Bolozky, University of Massachusetts - Amherst

Available at: https://works.bepress.com/shmuel_bolozky/33/
Surface geminates (*dageš forte*) in Israeli Hebrew
Shmuel Bolozky, University of Massachusetts/Amherst


1. **Gemination in pre-modern Hebrew**

   In traditional Hebrew linguistics, *dageš xazak*, or *dageš forte*, is a term used for gemination, i.e., for a long consonant, represented either by a special mark, or by doubling the consonant. In syllable division, a geminate consonant is assumed to close one syllable and serve as the onset to the next one, as in *dib-ber* ‘he spoke.’ Spirantization transforms a stop into the corresponding fricative, e.g., /b/ into [v], when that stop follows a vowel. Gemination blocks spirantization, e.g., while /dabar/ ‘a thing; an act of speech’ is realized as *[davar]*, *dibber* is unaffected. One of the most salient differences between Standard Israeli Hebrew (as distinguished from some Arabicized traditions) and earlier phases of the language is the loss of gemination in the former. What is left of it are only its historical residues. A word like Biblical Hebrew *dibber* is realized as *diber* in Israeli Hebrew. Spirantization is still blocked, but now it can no longer be explained why *davar* has undergone spirantization, while *diber* has not, since the surface environment is the same. Furthermore, the loss of gemination can no longer account for the blocking of vowel deletion where expected. In an open syllable, the vowel *a* is deleted (or was historically reduced to a schwa) two syllables before the main stress, as in:

   (1)  
   *davár* ‘a thing’ ~ /davar+ím/ > d(⟨e⟩)varím ‘things’ 
   *paqíd* ‘clerk’ ~ /paqid+ím/ > p(⟨e⟩)qidím ‘clerks’

   Historically, when a geminate consonant was present, the syllable concerned was closed, and deletion was blocked, so as prevent the formation of a three-consonant sequence (even with a schwa in between the first two, the sequence is still hard to articulate):

   (2)  
   *pattiš* ‘hammer’ ~ *pattiším* ‘hammers’ (*p(⟨e⟩)ttiším)*
   *gannáv* ‘thief’ ~ *gannavím* ‘thieves’ (*g(⟨e⟩)nnavím*)

   Since in Israeli Hebrew such words are realized without gemination, e.g., *patiším*, there is no *a priori* reason why /pakid+ím/ would undergo reduction, whereas /patiš+ím/ would not. One might argue that historical geminates should still be posited underlingly in Israeli Hebrew, to account for differences in behavior between superficially-similar strings with respect to rules like spirantization and *a*-deletion; once the rules have applied (or in this case have failed to apply), we get rid of all geminates. But by the same token, one could simply mark the relevant segments for not undergoing spirantization, or for not allowing the preceding vowel to undergo reduction. There is no independent evidence that an actual geminate must be assumed there, since

Bolozky, Surface geminates, page 1
geminates of this type are never realized on the surface.

2. **Avoiding intra-morphemic gemination in Israeli Hebrew**

Pattern comparison suggests that in environments where potential geminates could arise within the morpheme, they are either broken with the minimal vowel e:

(3) zalelan ‘glutton’ xatetan ‘meddler’
    noxexut ‘presence’ holelut ‘foolly, hilarity’ cf. rokxut ‘pharmacology’

or the elision of a vowel separating between identical consonants is blocked (although reduction to e still takes place), as in:

(4) xagag ‘he celebrated’ ~ xagega ‘she celebrated’
    cf. katav ‘he wrote’ ~ katva ‘she wrote’
    kucec ‘it was cut’ ~ kucecu ‘they were cut’
    cf. supac ‘it was overhauled’ ~ supcu ‘they were overhauled’
    hitpalel ‘he prayed’ ~ hitpalela > ‘she prayed’
    cf. hitlabes > ‘he got dressed’ ~ hitlabsha > ‘she got dressed’

McCarthy regards the blocking of e-deletion in xagega, etc. as “antigemination”:\(^5\) syncope rules are prohibited from creating clusters of identical consonants. This is an immediate corollary of his Obligatory Contour Principle (OCP), which prohibits adjacent identical elements at the melodic level (either consonantal or vocalic, in an autosegmental analysis).

3. **Geminates across morpheme boundary and in connected speech**

Phonetically, geminate consonants do occur in Israeli Hebrew in some environments. Potentially, they **may** be formed across morpheme boundary, although their occurrence is only optional (stress is word-final, unless marked otherwise):

(5) (a) šavat > ‘he was on strike’ šaváti > ‘I was on strike’ ~ šaváti > ‘I was on strike’
    avad > ‘he worked’ /avádi/ > ‘I worked’ > aváti > avádeti > aváti
(b) dan > ‘he discussed’ dánnu > ‘we discussed’ ~ dánu
    yašen > ‘he slept’ yašánu > ‘we slept’ ~ yašánu
(c) tamim > ‘naïve’ hittamem > ‘he feigned naïveté’ ~ hitamem
    dirder ‘he caused to deteriorate’ /hiddarder/ > hiddarder > he deteriorated ~ hiddarder

Autosegmental theory allows such geminates: the OCP does not apply to the across-morpheme-boundary cases, since different morphemes are represented on different tiers.

In practice, however, it is mostly group (a) geminates that tend to occur in connected speech. In a study conducted by Aniv,\(^6\) involving 12 speakers and some radio recordings, there
were no instances at all of geminates in group (c), and one marginal geminate occurred in group (b). Where potential geminates could arise in group (a), about half of the instances were split with $e$, and the other half were about evenly divided between cases in which the geminate was maintained and ones in which complete assimilation (or geminate simplification) occurred. Aniv’s data, as well as other data collected by the author, suggest that:

(i) suffix-related $\ldots t+et$ or $\ldots d+et$ realizations can be found at all registers;
(ii) suffix-related $\ldots t+t$ gemination tends to occur when attention is high and in deliberate speech;
(iii) complete assimilation of the same sequences ($\ldots t+t > t$), i.e. geminate simplification, is characteristic of casual/fast speech.

While suffix-related $t+t$ gemination is usually found in deliberate speech, the $e$ which splits intra-morphemic geminates may elide when casualness and tempo are high, particularly when sonorant consonant and fricatives are involved (see also Aniv 1997), e.g.:

(6) $\text{od ló rа̱tī zа̱lelán kazè } > \text{‘I have never seen such a glutton’ } > \ldots \text{zallán...}$
   yet not I saw glutton like this

   $\text{em kílelú et àmorá } > \text{‘They cursed the teacher’ } > \text{èm kíllú tamorá}$
   they cursed ac. the teacher

   $\text{yéš lo nòxexút norá maršimá } > \text{‘He has an awfully impressive presence’ } >$
   there is to him presence awful impressive f.s.
   …noxxut…

   $\text{em šàxexú otò babáit } > \text{‘They forgot him at home’ } > \text{èm šaxxú to babáit}$
   they forgot him at home

but:  $\text{i šàtetá bakbúk šálém } > \text{‘She drank a whole bottle’ } > \text{??i šattá bakbúk šálém}$
   she drank bottle whole

In other words, casual/fast speech may further reduce any reducible sequence, even if the result turns out to be a surface geminate within what used to be the morpheme boundary before it was erased. Thus, although geminates occur only marginally in Israeli Hebrew, they are not totally excluded from any register. They are only constrained from being realized within the morpheme in non-casual, deliberate speech. Bolozky shows how the notion of ‘strict cyclicity’\textsuperscript{8} may explain this constraint.\textsuperscript{9} The ‘strict cyclicity’ principle blocks rules from applying to environments derived in the same cycle, i.e., to environments arising either by morpheme combination or by application of a previous process. Casual/fast speech $e$-deletion is consequently blocked until the level of connected speech: it may apply neither at the base morpheme level, nor at the inflected word level, where $e$ results from the appending of the suffix and the ensuing reduction of $a$ to $e$. Even if the resulting geminates are within what used to be

Bolozky, Surface geminates, page 3
the morpheme, it boundaries have already been erased at earlier stages of the derivation by the
time the processes of connected speech apply.

4. **Why some geminate configurations can only undergo geminate simplification, and are never subject to geminate splitting**

Except for the possible occurrence of geminates of the *zallan*-type in very fast/casual speech, the restriction on intra-morphemic geminates suffices to explain their absence within morphemes. The question is why across morphemes, some types of geminates may be broken, while others may not. As shown above, the šavát+ti type often undergoes e-insertion, which breaks the surface geminate or a homorganic *d+t* sequence -- just like in English *prodded*, *wanted*, except that in English the process is obligatory. But the same never happens to forms like */hit+tamem/* or */hit+darder/*, nor to */dán+nu/* or */yašán+nu/*:

(7) šavátti > ‘I was on strike’ ~ šaváteti /avádti/> ‘I worked’ > avátti ~ avádeti  *yašánenù > ‘we slept’ *dánenù > ‘we discussed’
   *hitètamém > ‘he feigned naïveté’ *hitèdardér > ‘he rolled down/deteriorated’

As already noted, the */hit+tamem/* and */dán+nu/* types avoid gemination by cluster simplification rather than by cluster split; the question is why the option degeminating by insertion of an epenthetic *e* is totally excluded for these types. Bolozky\(^\text{10}\) proposes that inapplication of *e*-insertion in the */hit+tamem/* type means that the process is restricted to sequences involving an *inflectional* affix (/hit+/ being a derivational affix), and that it is blocked in *dánnu* (/+nu/ being inflectional) since *dánenù* would have been interpreted as stemming from a geminate root *d.n.n* instead of the correct *d.w.n* (or *š.n.n* instead of *y.š.n* in the case of *yašánenù*). Reevaluation of the data suggests that the reasons are rooted in more solid *phonological* grounds. It essentially has to do with the role of sonority -- or its inverse, consonantal strength -- in determining syllable structure.

4.1 **Sonority-based explanation for the blocking of geminate splitting in some configurations**

Broadly speaking, sonority is a characterization of the relative amount of audible sound that comes out of the vocal tract in speech production. Consonantal strength is essentially the inverse; it characterizes the energy that is expended on forming speech sounds rather than the volume of what comes out. The greater the constriction, the more the energy that is spent on production, reflecting consonantal strength, and the smaller the amount of energy that is emitted. And vice versa: the lesser the constriction, the greater the volume coming out, reflecting sonority. If we go by sonority, for instance, the most sonorant segments are low, open vowels, like *a*; the least sonorant are voiceless plosives, like *t*. The universal sonority scale, starting with the least sonorant (or strongest), is roughly the following:

Bolozky, Surface geminates, page 4
(8) voiceless plosives (like p t k)
    voiced plosives (like b d g)
    voiceless fricatives (like f θ s š x)
    voiced fricatives (like v ð z ž γ)
    nasals (like m n ŋ)
    lateral liquids (l-sounds)
    central liquids (r-sounds)
    semi-vowels (like y w)
    high vowels (like I u)
    mid vowels (like e o)
    low vowels (like a a)

Constraints on syllable structure are often based on the sonority scale. Thus, for instance,
the syllable nucleus must be a sonority peak, i.e., a vowel. But in the absence of a vowel, a
syllabic consonant may also fulfill this function, e.g., the n in English bu[t] > ‘button,’ or the r
in Czech [br-no] ‘Brno.’ Also, the consonants in the syllable onset must be arranged in order of
increasing sonority towards the nucleus, which is why [breyk] > ‘break’ is acceptable but not
[*rbeyk]. The reverse order (i.e., gradually decreasing sonority) applies to the progression from
the nucleus to the end of the coda, i.e., pu[ŋk] > ‘punk,’ but not *pu[kŋ]. If sequences violate the
sonority scale arise, e.g., in the process of borrowing, the sequence must either be broken, or
divided between two syllables, or simplified. Hebrew often adopts the first two strategies: /hércel/
> ‘Herzl’ > [hércel], /stratégyal ‘strategy’ > [as-tratógya]. Another principle (Vennemann’s
1988 “Contact Law”): a syllable contact is preferred when the offset of one syllable is more
sonorant than the onset of the next one, and the greater the difference in sonority, the more
preferred is the contact.

One of the criteria for preferred syllable head is that the sonority of the consonantal onset
be as low as possible (Venneman’s “Head Law (b)”). Again, the principle underlying it is
maximization of sonority distinctions: since the nucleus is a vowel (or a sonorant consonant), and
any vowel is more sonorant than a consonant, then the lesser the sonority of the consonantal
onset, the more optimal the consonant-vowel sequence is. In our case, n is not a preferred onset,
since it is a sonorant consonant, i.e., a consonant of high sonority. Furthermore, the ne that
would have been generated by epenthesis in /dán+nuł > *dánenù would also be non-optimal in
view of the similarly-motivated “Sequence Law”: a sequence of elements (e.g., segments,
syllables) is preferred when the adjacent components are less alike (in sonority and otherwise).
In the case of ne, the vowel and the sonorant consonant are not far from each other on the
sonority scale. te in aváteti, on the other hand, is preferred because of the significant
sonority/strength difference between t and e. The same explanation may be used to account for
the preference for sonorants when intra-morphemic geminates resurface, e.g., in zalelan > zallan
etc. above, where l and e are close in sonority. Although ll appears to constitute an instance of
“ultimate similarity,” we are actually dealing with a single long consonant, which is far easier to
articulate in fast speech than a lel sequence. Starting with a sonorant consonant, continuing with
a vowel that is of not much greater sonority, and then shifting back to l, is not easy to articulate.
Not so in the case of tet, where the sonority difference between the stops and the intervening vowel is significant. Although it could be argued that by the same reasoning, *hitètamém should have been as good an alternative of getting rid of the geminate as geminate simplification is (/hitamem/ > hitamem), the difference is in the number of unstressed syllables involved. In the case of avátetì, there are two unstressed vowels following the main stress; mechanical secondary stress may apply to the final i, which is natural. A sequence of three unstressed syllables in *hitètamém is highly irregular, and for mechanical secondary stress to fall on the inserted e is unusual.

4.2. Orthography-based explanation for the blocking of geminate splitting

Another approach would invoke the orthography: whereas dánnu and hittamem are always represented with a single consonant in Hebrew writing, šavát(e)ti can be found with either one or two consonants, and avád(e)ti certainly maintains the two (orthographic) consonants regardless of how it is realized phonetically. Since we are talking of the formal/slow register, which is clearly affected by literacy, straightforward explanations from orthography are not to be pushed aside as irrelevant.

In fact, the possibility that orthography plays a role here may be more significant than it initially appears to be. Although the loss of gemination is clearly a function of the non-Semitic linguistic infrastructure of the Ashkenazi speaker community, it is also feasible that the absence of gemination marking in normal Hebrew texts was a factor in making this loss easy to accept. Had there been an unbroken chain of speech, it would not have occurred. And as for Arabicized traditions that maintained gemination, in spite of the absence of live Hebrew speech, it is obviously the Arabic infrastructure that encouraged its preservation.

5. Conclusion

Gemination is not realized phonetically in Israeli Hebrew, and there is no convincing evidence that it should be employed as such at the underlying abstract level, only to be disposed of later across the board. It is clearly excluded from within morphemes, and where potential geminates could arise intra-morphemically, they are split by the minimal vowel e. Geminates may optionally arise cross morphemes boundaries. When they result from adjoining +t... suffixes to stems ending in t or d, the geminate may be maintained in formal/deliberate speech. Also, in connected speech, at high speech tempo and increased casualness, e-deletion may result in the type of geminates that are normally prevented from occurring in normal speech by this very same e. In most environments, however, geminates undergo geminate simplification. Since geminates can be gotten rid of either by cluster simplification or e-epenthesis, the choice of degemination device can be accounted by referring to sonority considerations, or to the effect of orthographic representation.
Gemination blocks spirantization either because a long consonant is more resistant to lenition processes, or mechanically, it might be argued that in a form like *dibber*, spirantization yields intermediate *divber*, as expected, but then regressive assimilation restores *dibber*.

A starred form is an ungrammatical form, or a non-existing one.

Note that in Standard Israeli Hebrew, /q/ merges with /k/.

Actually, in the case of spirantization, the rule has become heavily morphologized, and reduced to certain limited environments. It is possible that instead of marking exceptions to it, it is the items undergoing it or the environments allowing it that should be marked.


Aniv, H., “*hageminácyá hafonéítit baivrit hayisreelit* (Phonetic gemination in Israeli Hebrew,)” 1997, Ben-Gurion University ms.

Bolozky, Surface geminates, page 7
stress in Israeli Hebrew tends to alternate fairly regularly, i.e., is normally assigned to every other syllable away from the main stress. If the secondary stress of one word clashes with the primary stress of another, the former undergoes destressing.


