Constraints imposing minimal word sizes (in terms of feet, syllables, or sometimes morae) have been found to hold in many languages (Dixon 1983, McCarthy and Prince 1986, Garrett 1999, and many others). Usually these have been viewed as purely phonological constraints on the size of content words in a given language. Ito and Hankamer (1989), however, proposed a word size constraint (henceforth WSC) for Turkish which is not purely phonological: while monosyllabic roots are plentiful, it prohibits polymorphemic monosyllabic words. Ito and Hankamer attributed the constraint's sensitivity to morpheme count to a derived environment effect; Inkelas and Orgun (1995) carried that idea a step further, suggesting that the WSC holds only at certain levels in a level-ordered phonology.

The question addressed in this paper is how the WSC interacts with the rest of the grammar, in the framework of Distributed Morphology (Halle and Marantz 1993, et seq.). Ito and Hankamer suggest that it is responsible for the selection of allomorphs of the Aorist suffix, which would mean, in DM terms, that it is active at the point of Vocabulary Insertion. In this paper I will argue that the WSC is not in effect at the point of VI, and cannot be involved in the selection of Aorist allomorphs in the way envisioned by Ito and Hankamer. In fact, it appears that the WSC is never involved in allomorph selection at all.

1. The Word Size Constraint

The existence of a Word Size Constraint for Turkish was first demonstrated by Ito and Hankamer (1989). What Ito and Hankamer noticed was that many Turkish speakers, when confronted with monosyllabic polymorphemic words (which should be licit given the usual rules of combination of morphemes, with the usual morphophonemics) reject them:¹

(1) *ye-n ye-n-di cf. at-il
   eat-PASS eat-PASS-PST throw-PASS
   "be eaten!" "was eaten" "be thrown!"

(2) *be-m be-m-de be-ler
   'b'-POSS1S 'b'-POSS1S-DAT 'b'-PL
   "my 'b'" "in my 'b'" "b's"

(3) *do-m do-m-da do-lar
   'do'-POSS1S 'do'-POSS1S-DAT 'do'-PL
   "my 'do'" "in my 'do'" "'do's"
To make the corresponding utterances natural, we must imagine such contexts as stage directions, penmanship classes, and music practice, but the contrasts appear to be robust for a large number of speakers of Turkish. It will be important in what follows that, for the speakers we are concerned with here, the forms in the middle column (ye-n-di, be-m-de, etc.) are grammatical, even though the base onto which the third affix is attached would be ungrammatical in isolation. This indicates that the WSC is not a constraint on the combination of morphemes per se, but truly a constraint on the size of words. Ito and Hankamer state the WSC as follows:

\[(4) \text{WSC}^3: \; ^*[\sigma] \quad \text{Word}\]

The WSC, as stated by Ito and Hankamer, would rule out monosyllabic roots, which would be incorrect. Ito and Hankamer suggest that the WSC is only active in derived environments, so that it comes into play only when affixation takes place. I will not dispute this, though it seems equivalent in this case to saying that a monosyllabic word cannot contain more than one morpheme:

\[(5) \quad [\sigma] \quad \Rightarrow \quad [M] \quad \text{Word} \quad \text{Word}\]

A monosyllabic word must be a monomorphemic word.

2. Ito and Hankamer proposal for Aorist allomorphy

Ito and Hankamer suggested that the WSC could provide an explanation for part of the pattern of allomorphy of the Turkish Aorist suffix. The pattern in question is that the Aorist appears to have three allomorphs: \(-r\) after vowels, \(-Lr\) after all but 13 monosyllabic C-final roots, and \(-Hr\) elsewhere.

\[(6) \quad \text{anla} + r \rightarrow \text{anla}-r \quad \text{koş} + Lr \rightarrow \text{koş}-ar \quad \text{çalış} + Hr \rightarrow \text{çalış}-ir\]

\(\text{understand + AOR} \quad \text{run + AOR} \quad \text{work + AOR}\)

The mystery that Ito and Hankamer wanted to solve was what makes (most) monosyllabic C-final roots take the \(-Lr\) allomorph, and why is that allomorph not available to other stems? Their answer, briefly, was that there really are just two allomorphs of the Aorist suffix, \(-r\) and \(-Lr\). When \(-r\) is affixed to a C-final stem, an independently attested process of vowel epenthesis introduces a high vowel, which then undergoes vowel harmony to determine its backness and rounding. The question then is what determines the selection of the \(-Lr\) allomorph.
Ito and Hankamer suggested the following: -r is the unmarked allomorph, and will appear wherever it can. But if -r is attached to a monosyllabic stem, the result will still be monosyllabic, as seen in (7). Thus the WSC (if it can be presumed to apply at this stage) would block the selection of the -r allomorph, leading to the selection of -Lr. Polysyllabic stems would never select -Lr, because the WSC would not be violated no matter what the form of the affix (8).

(7) koş+ʳ (violates WSC, leading to selection of -Lr allomorph)  
=> koş+Lr => koş-ar

(8) çalış+ʳ (does not violate WSC, so -r allomorph is chosen)  
=> çalış-Hr (epenthesis) => çalış-ir

This analysis must assume some delicate timing: The WSC must apply before V-epenthesis, otherwise the selection of -r would never be blocked. On the other hand, Ito and Hankamer assumed that the suffix-initial V~0 alternation in the nominal paradigm involves deletion of an underlying suffix-initial vowel following a vowel-final stem:

(9) kız + Hm => kız-ım  baba + Hm => baba + m => baba-m

This V->0 would have to happen before the WSC applies, else the WSC would not rule out forms like *be+m:

(10) be + Hm (satisfies WSC) => *be-m (violates WSC)

One might suggest that this embarrassment might be avoided if the V~0 alternation in the nominal paradigm also involves V-epenthesis instead of V-deletion. The following section will take up this possibility, and show that it too leads to an ordering problem, and a worse one.

3. V-epenthesis

Hankamer (2011) has argued that all V~0 alternations in Turkish suffixes reflect a V-epenthesis process instead of a V-deletion process. He argued that the V-epenthesis hypothesis leads to a much more satisfactory analysis of the progressive suffix –Hyor, and at the same time accounts for the fact that the V~0 alternation occurs only in suffixes whose shape (-C or –CCV) would violate the constraints on syllable structure if directly affixed to a C-final stem. Thus, according to his proposal, a form such as kız-im arises because the combination of underlying morphemes kız+m would result in an illicit syllable-final cluster, triggering epenthesis of a high vowel that undergoes vowel harmony to yield kız-im. Regarding the –Hyor suffix, he argues that the widely
accepted analysis in which the suffix-initial vowel is deleted, followed by a raising and harmonization of the stem-final vowel, cannot be correct, and that in fact it must be the stem-final vowel that deletes (or else the vowels are merged). His further proposal is that the other V~0 alternations in suffixes are all attributable to vowel epenthesis rather than vowel deletion. He notes that all the productive cases of this V~0 alternation involve suffixes that can be analyzed as single C or a form beginning with a sequence of two Cs. The only time a single C could survive attachment to a C-final stem without epenthesis would be if that C were a stop and the stem-final C a liquid or nasal. The only place in the language where this happens, in fact, is where the causative allomorph -t attaches to a polysyllabic stem ending in a liquid: düzel+CAUS -> düzel-t, *düzel-dir.

If this is correct, there cannot be a difference in the nominal and verbal paradigms such that, for example *re-m is due to V->0 applying before the WSC and *kork-ur is due to 0->V applying after the WSC, as required by the Ito and Hankamer analysis. One could, at this point, assume (following Hankamer 2011) that there is no V->0 and thus no ordering involved in *re-m; but there is still an ordering difficulty. If the WSC is in effect before V-epenthesis, as required in the Ito and Hankamer analysis of the Aorist allomorphy, it should prohibit the selection of the allomorph /-m/ when the stem is a monosyllable:

(11) kız + m (violates WSC) --> kız + Hm (epenthesis) --> kız-ım

In this case, /-m/ is the only allomorph, so the WSC should lead to ineffability. It does not, so in the nominal paradigm the WSC would have to apply after epenthesis (or at any rate it could not apply at VI).

The next section will show that the proposed account of Aorist allomorphy fails independently of this ordering paradox, so we will soon be able to escape from it.

4. Aorist allomorphy

To the difficulties outlined in the previous two sections can be added an even more grievous one: the WSC is a word size constraint. It cannot constrain combinations of morphemes that are smaller than the size of a word. If it did, be-m-i and ye-n-di would be ungrammatical, because *be-m and *ye-n are illicit (as words). But these forms are OK, for the speakers we are concerned with. Given this, it is impossible for the WSC to decide the allomorphy of the Aorist, since the form taken by the Aorist suffix does not depend on whether it ends a word or not:
The form of the Aorist suffix, once chosen, does not vary with later suffixation. If it were determined by the WSC, we would expect

(13) *kork-r-um

instead of

(14)  kork-ar-ım

Thus the form of the Aorist suffix cannot be governed by the WSC. In fact, as I argue in Hankamer (in prep), the most economical analysis of the Turkish Aorist is that its -Lr allomorph is idiosyncratically selected by a majority of monosyllabic roots. The -r allomorph appears elsewhere. Thus the only factor in the selection of the Aorist allomorph (in cases not involving the Negative) is root-conditioned allomorphy of a fairly common type: roots of a particular phonological shape tend to select a certain allomorph, but exceptionally some roots of that shape select a different one. This is comfortably boring.

5. Nature of the WSC

What the above discussion shows is that the WSC, while it cannot be viewed as applying at the level of Vocabulary Insertion, must be sensitive (exquisitely sensitive, it appears) to three kinds of information: first, the phonological shape of the form (at least, how many syllables it has); second, how many morphemes it has; and third, whether the form is a word. The third bit of information is the most striking, but perhaps it should not be. There are various phonological phenomena, such as word-stress placement, vowel harmony, and final stop devoicing, which are sensitive to the word as a domain. The parallel in syntax would be root clause phenomena, where certain syntactic processes apply only in root clauses.

How, then, does the WSC interact with the rest of the grammar in a framework like DM? First, it clearly cannot be pre-VI, on the assumption that pre-VI structures contain no phonological information. Could it hold at VI? (Recall, this is what it would have to do to play the allomorph-selecting role envisioned by Ito and Hankamer.) The answer here is no, if Hankamer (2011) is correct in his argument that -C~-VC suffixes are underlyingly -C. For then the affixation of a -C suffix to a monosyllabic root would immediately result in ungrammaticality most of the time:
(15) bez + m => bez-m => bez-Hm => bez-im
  cloth POSS1S  0->V    VH

^  VIOLATES WSC

So on this view, in which the suffix-initial V~0 alternation is due to V-epenthesis, at least V-epenthesis must be given a chance to apply before the WSC is enforced.

This means, of course, that at least in any case where V-epenthesis is involved, the WSC cannot play a role in allomorph selection. In fact, it looks like the WSC is never a factor in allomorph selection. If it were, we would expect to find situations like the following:

(16) HYPOTHETICAL MORPHEME

A morpheme like the Causative, with two allomorphs,
one a single consonant and the other an entire syllable;

Such that when the morpheme attaches to a monosyllabic root,
the syllabic allomorph is selected if the result is a word,
but if there is further suffixation, the single consonant
is chosen.

I have searched, but I don't think there are any such morphemes in Turkish.

6. Two exceptions

There appear to be exactly two exceptions to the WSC, the Aorist forms of the CV-root verbs ye- 'eat' and de- 'say': their Aorist forms are ye-r, de-r, which should be ungrammatical according to the WSC. İnkelas and Orgun suggest (p. 772) that these forms, while historically derived, are synchronically portmaneaus. I come to essentially the same conclusion in Hankamer (in prep). In DM terms, this means that these two roots idiosyncratically fuse with the Aorist pre-VI, so that at the point of VI they are monomorphemic.

An alternative would be to assume that there is a V-> (or VV -> V) operation that is ordered after the WSC, but that is exactly the operation that Hankamer (2011) argued did not exist, and it is not clear what purpose it would serve, other than to fix these two forms.
7. Conclusion

The main conclusion of this paper is that in a DM framework of assumptions the WSC cannot apply at the point of VI, and therefore cannot be involved in allomorph selection. It appears to be a straightforward surface wellformedness filter, which is sensitive to three kinds of information: whether a form is monosyllabic, whether it is monomorphemic, and whether it is a word. The last condition seems to be the morphological equivalent of root-clause sensitivity in syntax.

Footnotes
Orgun and Inkelas (1992), in a very careful empirical study, found that some speakers do not find (1-3) ungrammatical, but that all speakers of Istanbul Turkish observe a slightly different WSC, to the effect that a minimum (polymorphemic) word has two moras. In this paper I will only be concerned directly with the polysyllabic WSC of Ito and Hankamer, though I think the general conclusions will extend to other WSCs.

Inkelas and Orgun (1995) make the WSC (and another one, which they motivate and call the Bimoraic constraint) parts of Level 1 phonology, in a level-ordered theory. I will not argue about the level-ordering theory of Inkelas and Orgun, because I am primarily interested in how the WSC (and phenomena like it, such as the BMC) operate in a general theory of morphosyntax and morphophonology, such as Distributed Morphology.

Ito and Hankamer consider several alternative formulations, but that variation will not be relevant to the present study.

There are actually at least two more, but these occur only in construction with the negative, and so will not have any relation to the WSC.

H represents an underspecified High vowel subject to the four-way backness and roundness harmony; L represents an underspecified Low (unround) vowel subject to backness harmony only.

Orgun and Inkelas found groups of speakers for whom be-m-i is ungrammatical, which they took to support their level-ordering hypothesis. I have no idea what to say about those speakers, but it is clear that allomorph selection is not involved. The only possible allomorph of the POSS1S morpheme after a Vowel is -m.

An instructive parallel is provided by the Causative morpheme, which has (regular) allomorphs -dHr and -t. The -t allomorph is found after polysyllabic stems ending in a liquid, -dHr elsewhere. This means, of course, that a word consisting of a monosyllabic root + the -t allomorph will always be ungrammatical: öl-dür ("kill [it]"), * öl-t. But it would not do to attribute this to the WSC, because further suffixation does not change the allomorph choice: öl-dür-dü, *öl-t-tü.

References:


Hankamer (in prep) "Turkish Aorist Allomorphy"

Inkelas, Sharon and Cemil Orhan Orgun (1995) "Level Ordering and Economy in the Lexical Phonology of Turkish", Language 71.4: 763-793.

Ito & Hankamer (1989) "Notes on monosyllabism in Turkish", in Ito & Runner, eds., Phonology at Santa Cruz. Syntax Research Center working papers, UCSC.


Orgun, Cemil Orhan and Sharon Inkelas (1992) "Turkish Prosodic Minimality", Sixth International Conference on Turkish Linguistics, Anadolu University.