

Derivational suffixes as cues to stress position in reading Greek

Aikaterini Grimani

National and Kapodistrian University of Athens

Athanassios Protopapas

National and Kapodistrian University of Athens and University of Oslo

In Press, Journal of Research in Reading

Author Note

Aikaterini Grimani, Department of Philosophy and History of Science, National and Kapodistrian University of Athens, Greece; Athanassios Protopapas, Department of Philosophy and History of Science, National and Kapodistrian University of Athens, Greece, and Department of Special Needs Education, University of Oslo, Norway.

A preliminary report of this study was presented at the 16th Annual Meeting of the Society for the Scientific Study of Reading in Boston, MA, June 2009.

Aikaterini Grimani is now at the Karolinska Institute, Institute of Environmental Medicine, Intervention and implementation research Unit (IIR), Stockholm, Sweden.

Correspondence concerning this article should be addressed to Athanassios Protopapas, Department of Special Needs Education, P.O. Box 1140, Blindern, 0318 Oslo, Norway. E-mail: athanasios.protopapas@isp.uio.no

Abstract

Background. In languages with lexical stress reading aloud must include stress assignment. Stress information sources across languages include word-final letter sequences. Here we examine whether such sequences account for stress assignment in Greek and whether this is attributable to absolute rules involving accenting morphemes or to probabilistic lexical information.

Methods. Pseudowords were constructed to not resemble particular words and were suffixed with derivational morphemes associated with specific stress patterns, to be read aloud, presented either without a stress diacritic or with a diacritic congruent or incongruent with the suffix. Morphemes differed in whether or not there were stress competitors ending in the same letter sequences.

Results. Stress was assigned consistent with the suffix in the absence of the diacritic, more so when there were no stress competitors in the lexicon.

Conclusions. Results suggest a lexically-based probabilistic mechanism taking into account pattern distributions rather than absolute rules based on morphological accenting.

Keywords. Lexical stress; stress assignment; morphology; probabilistic cues; Greek

Highlights

What is already known about this topic

- Stress assignment in reading can exploit various information sources
- Word endings are statistically associated with stress patterns
- Certain morphemes impose their stress patterns on the words they derive

What this paper adds

- Word-final letter sequences affect stress assignment
- The existence of a diacritic trumps the word-ending effect
- The word-ending effect is weaker when stress competitors exist in the lexicon

Implications for theory, policy or practice

- Multiple sources of information are operative in stress assignment
- A dominance hierarchy seems to govern the influence of stress information sources
- Statistical associations seem a more likely explanation than absolute rules

Derivational suffixes as cues to stress position in reading Greek

In languages with lexical stress, in which stress position can vary and contributes to lexical identity, stress assignment is a necessary component of successful reading aloud. Presumably, each lexical entry includes stress information, otherwise it could not be pronounced correctly. But this does not imply that stress assignment in reading must be based on the lexically stored information only. Indeed, studies have shown that additional sources of stress information are taken into account, observed most clearly in pseudoword reading. For example, orthographic stress information in the form of a special diacritic affects stress assignment when available, such as in the Greek language (Protopapas, 2006; Protopapas, Gerakaki, & Alexandri, 2006, 2007; Protopapas & Gerakaki, 2009). Dominant or default patterns also affect stress assignment in languages such as Italian and Greek (Colombo, 1992; Burani, Paizi, & Sulpizio, 2014; Sulpizio & Colombo, 2013; Protopapas, 2006; for reviews see Protopapas, 2016; Sulpizio, Burani, & Colombo, 2015).

Indirect lexical effects on stress assignment in reading have also been documented, both from individual words and from word groups. Specifically, in Greek, pseudowords preferentially receive the stress patterns of similar words when closely resembling them (i.e., when they share most of their letters), across ages (Protopapas et al., 2006, 2007; the effect being strongest in the early elementary grades; Protopapas & Gerakaki, 2009). In Italian, pseudowords preferentially receive the majority stress patterns of word groups that are similar (i.e., written with the same letters) in their final two syllables (termed “stress neighborhoods”; e.g., Burani & Arduino, 2004; Burani et al., 2014; Colombo, 1992; see Sulpizio et al., 2015, for review). Further removed from specific words, significant statistical associations between stress positions and orthographic patterns, particularly word beginnings and (especially) endings and have

been documented in several languages (Arciuli, Monaghan, & Seva, 2010; Monaghan, Arciuli, & Seva, 2016; Ševa, Monaghan, & Arciuli, 2009), making up an additional, probabilistic source of stress assignment information. These associations appear to be informative in stress assignment processes, at least in English (Arciuli & Cupples, 2006; Arciuli et al., 2010; Wade-Woolley & Heggie, 2015). Indeed, word beginnings and endings associated with particular stress patterns have been used to model stress assignment in reading multisyllabic English words in both connectionist and rule-based computational implementations (Arciuli et al., 2010; Rastle & Coltheart, 2000).

However, in most of these languages there are no specific cues to stress (i.e., no diacritic) and no general default pattern. The question then arises: would these weak statistical associations affect actual stress assignment behavior when reading in a language with strong and specific sources of stress information? Although one might reasonably expect that reliable and specific sources would dominate stress assignment behavior and swamp minor statistical associations, there are also reasons to expect otherwise, as Greek readers have been found to not rely entirely on the stress diacritic despite its reliable presence and excellent cue validity (Protopapas, 2016). This question is of major theoretical importance because it concerns the long-debated notion of probabilistic vs. rule-based cognitive systems. If sensitivity to reliable but relatively minor statistical associations can be demonstrated in the face of well-established alternatives then the appeal of positing absolute rules (and thereby also defining “exceptions” to these rules be “irregular”) may be diminished.

Greek offers the opportunity to address this question in conjunction with a related morphological hypothesis coming from linguistics. Specifically, on the one hand, there are documented weak but significant statistical associations between initial and final letter groups and stress positions (Monaghan et al., 2016). On the other hand, there

is a family of derivational suffixes that, according to linguistic analysis, carry their own stress pattern (being either accented or unaccentable) and impose it on the derived words they form, overriding any stress patterns that may be associated with the roots they attach to (Revithiadou, 1999). For example, the suffix -άδα /'a.ða/ combines with nouns or adjectives to form nouns, such as πορτοκάλι /por.to.'ka.li/ “orange” → πορτοκαλάδα /por.to.ka.'la.ða / “orange juice”. Similar suffixes exist in English and have been found to modulate stress assignment preference in pseudowords (Wade-Woolley & Heggie, 2015). Note that accenting is an absolute, formal property of suffixes, which cannot be diluted or trumped by other properties or constraints, and is not subject to probabilistic influences from other lexical items or related processes.

These two hypotheses, that is, statistical associations and morphological accenting, both suggest that word-final letter sequences can control, or at least contribute to, stress assignment in reading in Greek, as in other languages. However, they can also lead to competing predictions, taking into account their formal properties. Specifically, if critical letter sequences are recognized as accenting morphemes and if this morphological source of stress assignment information is typically exploited when reading, then we should observe full compliance with the suffix stress pattern, even in pseudowords, as long as the pseudowords are appropriately suffixed. This amounts to a morphological accenting hypothesis for stress assignment. A crucial property of this hypothesis is its absoluteness, as formal linguistic rules are posited to apply strictly (not probabilistically) whenever their relevant domain is matched. The suggestion that reading aloud words (and even pseudowords) is based on morphemic parsing is consistent with the findings of Wade-Woolley and Heggie (2015), who found that English pseudowords suffixed with accenting morphemes are preferred when stressed in accordance with the suffix. It is also consistent with the model of Rastle and Coltheart

(2000), which is based on the identification of specific prefixes and suffixes in English that are associated with stress patterns via hard-coded rules. It is further supported by evidence in Italian (Burani, Arduino, & Marcolini, 2006; Burani, Marcolini, De Luca, & Zoccolotti, 2008), which is relatively similar to Greek in orthographic transparency and morphological complexity. Therefore, if words (and pseudowords) are routinely parsed into roots and suffixes in the course of processing for reading aloud (cf. morpho-orthographic segmentation; see, e.g., Beyersmann et al., 2016, and references therein), it is reasonable to hypothesize that the stress properties of roots and suffixes are available to the (adult) skilled reader and may be operative in directing stress assignment.

In contrast, if word-final letter sequences are only probabilistically associated with stress patterns, irrespective of their status as morphemes or other linguistic units, and if this statistical information is taken into account, then we should observe stress assignment consistent with the pattern typically associated with the letter sequence—but only to the extent that it dominates alternative patterns. That is, if words exist that end in the same letter sequences but have different stress patterns, then pseudowords ending in these letter sequences would be assigned stress less systematically than pseudowords ending in letter sequences that only appear in words with a single stress pattern. In other words, the presence of stress competitors (and their type and token frequency) would be expected to affect the position and reliability of stress assignment. This can be termed the probabilistic stress assignment hypothesis, consistent with related findings in English (Arciuli & Cupples, 2006; Arciuli et al., 2010). It is closely related to stress consistency effects observed in Italian stress neighborhoods because those effects are also thought to arise from knowledge of distribution of stress patterns over words with common word-final segment sequences (e.g., Sulpizio, Arduino, Paizi, &

Burani, 2013; Sulpizio & Colombo, 2013). However, our probabilistic stress assignment hypothesis differs from stress neighborhood effects, as they are traditionally defined, in that it is not a priori restricted to a prespecified domain such as the last two syllables of the words. Rather, letter sequences of any size may be potential cues to the extent that they are systematically associated with a particular stress pattern. The cumulative effect of multiple, size-varying, and potentially overlapping cues, would then be observable in stress assignment behavior on unfamiliar pseudowords, in which there is no fixed lexical knowledge to dominate the probabilistic cues and mask their influence.

Greek is a language with a relatively transparent orthography (estimated at 95% at the grapheme-phoneme level for the reading direction; Protopapas & Vlahou, 2009), in which stress is obligatorily marked with a diacritic. Previous studies in Greek have documented effects of three sources of stress assignment information: lexical (i.e., knowledge of the word), orthographic (i.e., diacritic), and default (i.e., the penultimate syllable) (Protopapas, 2016). Of these, the lexical source dominates responses to pseudowords already from Grade 2, while the diacritic gains strength gradually as reading expertise develops. The default pattern results in a weak stress assignment tendency, evident only in the absence of other sources. To examine the relative strength of morphological or probabilistic stress assignment information in skilled Greek readers, in the present study we presented pseudowords with accenting suffixes without a diacritic or with a diacritic congruent or incongruent with the suffix. The hypothesized morphological and probabilistic source would be most clearly evident in the absence of a diacritic, whereas its relative strength, compared to the diacritic, can be determined by stress assignment preferences in the incongruent condition. Operation of the default source would be evident in a preference for penultimate-syllable stress.

Method

Participants

In total, 57 adult native speakers of Greek (18–56 years old; 32 female) participated in this experimental procedure voluntarily, without monetary or other compensation. Most were students (25–30 years old). Each responded to one of three experimental lists only (19 participants per list).

Stimuli

A list of 123 presumed stress-assigning derivational suffixes, 1–3 syllables long, were selected from Greek grammar sources (Babinotis, 2006; Holton, Mackridge, & Philippaki-Warburton, 1999; Ralli, 2005; Triandafillidis, 1997), including suffixes stressed on all three final syllables. Words with final phoneme sequences matching these suffixes (i.e., including differently spelled homonyms) were then retrieved from IPLR (speech.ilsp.gr/iplr; Protopapas, Tzakosta, Chalamandaris, & Tsiakoulis, 2012). For each suffix, the proportion of words stressed according to the presumed stress pattern of the suffix was calculated. An experimental subset of 44 suffixes were selected, 2–3 syllables long, satisfying the following criteria: present in >15 word types, of cumulative frequency >70 tokens, and with asymmetric stress proportion (>14 for both types and tokens) over stress competitors, including phonological and orthographic stress competitors (i.e., matching final phoneme or letter sequence and stressed on a different syllable). For 34 of these suffixes there were no orthographic stress competitors, that is, no words in the corpus ending in the same letter sequence and stressed on a different syllable. For the other 10 suffixes there were stress competitors (of lower cumulative frequency); for example, the word λησμοσύνη (/lizmo'sini/ “oblivion”) contains the suffix -οσύνη, whereas the word χαρμόσυνη (/xar'mosini/ “joyous”) ends in the same letter sequence but does not contain the same suffix and is

stressed differently. Table 1 lists the distribution and summary metrics of the selected suffixes and stress competitors.

One hundred eighty words 3–5 syllables long containing the selected suffixes were then modified to create pseudoword stimuli, by changing several consonants and vowels in the root while preserving the suffix as well as the length, syllable structure, and bigram frequency of the word. For example, the word *πριγκίπισσα* (/pri'ɟipisa/ “princess”) was changed to *μπλαχόμισσα* /bla'xomisa/. For suffixes with stress competitors, pseudowords were matched not to a single word but to a word pair, including a suffixed and a matched non-suffixed word ending with the same letter sequence. For example, the pseudoword *ρολποσυνη*/rolposini/ was based on the aforementioned *χαρμόσυνη-λησμοσύνη* word pair.

Pseudowords were submitted to a pretest to ensure that they did not resemble any real word. Eleven individuals (none of which participated in the main experiment) were given printed sheets with the 180 items and were asked to write next to each pseudoword a real word that came to mind. Pseudowords were discarded if more than two individuals produced the same word (either the source word or any other word).

A set of 90 experimental items were finally selected. All experimental items were pseudowords, 3–5 syllables long, containing suffixes 2–3 syllables long. Of these, 66 were derived from words with suffixes having no stress competitors, henceforth termed “unanimous” stress items. The remaining 24 were derived from stress-competitor word pairs with and without suffixes, termed “majority” stress items. The experimental suffixes and items are listed in the Appendix. Although there was an imbalance in the number of unanimous and majority items in the experiment, there were no other items fulfilling our strict selection and matching criteria for the majority items, and we preferred the imbalance rather than relaxing the criteria and thereby risking the

interpretability of the results.

Filler items included the 24 stress-competitor word pairs, as well as 90 words and 48 additional pseudowords (from Protopapas et al., 2007) without derivational suffixes, matched in bigram frequency and length, for a total of 138 words and 138 pseudowords (276 stimuli in all, of which only the 90 critical suffixed pseudowords are analyzed below), half of each without a suffix, to discourage attraction of attention to the suffixes. Words were included in the study as fillers, mainly to permit a pilot examination of the effects of omitting or misplacing the diacritic on real words (not reported here). The existence of the words in the stimulus set is not expected to affect the results if the morphological hypothesis is correct, because morpheme accenting is by definition unaffected by context. However, the presence of stress competitor pairs might conceivably affect probabilistic stress assignment, thus enhancing observed effects, by drawing attention to alternative stress patterns, if stress assignment is susceptible to local context. Thus these fillers could potentially act as an amplifier of the predicted differences between the two competing theoretical hypotheses.

All stimuli were presented with or without a diacritic; the diacritic, when present, was either in the same position as in the source word or in a different position, for a total of three stress diacritic conditions: congruent, incongruent, absent. Three experimental lists were created, with an equal number of items from each condition. Each list contained every experimental item in one of the three conditions only.

Procedure

Participants were tested individually in a quiet room, with a small break between 92-trial blocks. They were asked to read aloud the items presented on the screen and were encouraged to proceed as quickly and spontaneously as possible without correcting their responses or otherwise interrupting the procedure. Individual

stimuli were displayed for 1,500 ms in 40-pt white Arial Greek on black background, controlled by DMDX (Forster & Forster, 2003). Six familiarization items preceded the experimental trials. Vocal responses were recorded using a headset-mounted microphone.

Data Analysis

Responses were individually examined using CheckVocal (Protopapas, 2007), a specialized software that displays the waveform and spectrum of the recorded response along with the correct response and RT mark. A native Greek listener went through the recordings twice, once to verify the response time and to judge whether each response was pronounced correctly, and once to judge on which syllable it was stressed.¹ Each valid response was considered correct when all segments were pronounced in the correct order, regardless of stress assignment. Self-corrections were not considered; that is, in cases of multiple responses in the audio file, only the first recorded response was taken into account, whether complete or incomplete, correct or incorrect.

Response times were logarithmically transformed and analyzed using linear mixed-effects models with crossed random effects for participants and items (Baayen, 2008; Baayen, Davidson, & Bates, 2008). Stress assignment proportions were compared using generalized linear mixed-effects models for binomial distributions (Dixon, 2008) via a logit transformation (Jaeger, 2008). Models included near-maximal random structures (Barr, Levy, Scheepers, & Tily, 2013) to the extent allowed by convergence constraints, and were fit with functions `lmer` and `glmer`, respectively, of the `lme4` package (Bates, Maechler, Bolker, & Walker, 2013) in R (R Core Team, 2012).

¹ Only a single listener judged stress position for each response. We do not expect any reliability issues necessitating additional judges because this is a very clear judgment to make in Greek and no disagreements have been observed in pilot and past work.

Unless noted otherwise, contrasts involving categorical variables were deviation-coded, that is, one factor level was assigned the value -1 and the other the value $+1$. This sets the model intercept to the grand mean and estimates coefficients as main effects of the factor relative to that (for more details, see UCLA Statistical Consulting Group, n.d., and Protopapas, 2015). p values for the RT models were calculated based on the Satterthwaite approximation using package `lmerTest` (Kuznetsova, Brockhoff, & Christensen, 2014).

The following factors were used in the analyses, as appropriate: “Stress preference” was a between-item factor with two levels: unanimous vs. majority. “Stress position” was a between-item factor with three levels: antepenultimate, penultimate, and final. “Diacritic condition” was a within-item factor with three levels, corresponding to presentation: diacritic absent vs. congruent with the suffix vs. incongruent with the suffix. “Response stress consistency” was a post-hoc within-item factor with two levels, grouping together responses consistent vs. inconsistent with a cue (i.e., responses consistent in stress pattern with the suffix or with the displayed diacritic). Because there were two stress cues (diacritic and suffix), potentially conflicting, analyses including response stress consistency were carried out separately for response-diacritic stress consistency and response-suffix stress consistency.

Results

Response times

Table 2 shows the RTs for correct responses to each type of item in each diacritic presentation condition. There is no evidence that responses were faster in any condition, for any item type, or when consistent with the suffix or with the diacritic. Indeed, analysis of log RTs with diacritic condition (absent vs. congruent and absent vs. incongruent) and stress preference (unanimous vs. majority) as fixed factors revealed

no significant effects. Specifically, in the analysis including the absent and congruent diacritic conditions, there was no main effect of diacritic condition ($\beta = -0.02$, $t = -1.93$, $p = .058$) or stress preference ($\beta = -0.03$, $t = -1.24$, $p = .220$) and no interaction between the two ($|\beta| < 0.01$, $t = 0.07$, $p = .947$). Similarly, in the analysis including the absent and incongruent diacritic conditions, there was no main effect of diacritic condition ($|\beta| < 0.01$, $t = 0.40$, $p = .695$) or stress preference ($\beta = -0.04$, $t = -1.62$, $p = .108$) and no interaction between the two ($\beta = -0.02$, $t = -0.81$, $p = .420$). Thus the presence of absence of the diacritic did not significantly affect response times to items with either unanimous or majority stress preference.

Next we examined whether RTs were different when responses were consistent or inconsistent with the available cues, namely the stress preference of the suffix and the displayed diacritic. Log RTs were analyzed with response stress pattern (consistent vs. inconsistent with the suffix) and suffix stress preference (unanimous vs. majority) as fixed factors, in the appropriate diacritic condition. In the absent diacritic condition, there was no main effect of response-suffix consistency ($|\beta| < 0.01$, $t = -0.23$, $p = .818$) or suffix stress preference ($\beta = -0.03$, $t = -1.31$, $p = .213$) and no interaction between the two ($\beta = 0.01$, $t = 0.40$, $p = .692$). Similarly, in the incongruent diacritic condition, there was no main effect of response-diacritic consistency ($|\beta| < 0.01$, $t < 0.01$, $p = .998$) or suffix stress preference ($\beta = -0.06$, $t = -1.87$, $p = .069$) and no interaction between the two ($\beta = 0.03$, $t = 0.67$, $p = .506$). The congruent diacritic condition was not analyzed because there were too few responses inconsistent with the suffix and the diacritic.

The lack of effects in response latencies permits us to focus on stress assignment proportions without concerns stemming from differences in difficulty or speed-accuracy tradeoffs.

Stress assignment

Table 3 shows the proportion of stress assignment in the three diacritic conditions as a percentage of total responses, broken down by suffix stress position, including all responses (top) or (segmentally) correct responses only (bottom; percentages not adding up to 100% due to mispronunciations). Incorrect responses are only shown for completeness and were not analyzed further. Across suffixes, taking only correct responses into account, unanimous-stress items were stressed consistent with the suffix 75.7%, 94.0%, and 18.3% in the absent, congruent, and incongruent diacritic condition, respectively; whereas the corresponding percentages for majority-stress items were 63.4%, 93.0%, and 12.1%. The distribution (participant means) of stress assignment responses consistent with the suffix in each diacritic condition, averaged over suffix stress positions, is displayed graphically in Figure 1 (left).

To examine the effect of suffix preference and stress position on stress assignment behavior, the proportion of stress assignment in responses consistent with the suffix (*strsuf*; a binomial variable) was analyzed in models including suffix stress position (*strpos*, with three levels: final, penultimate, antepenultimate; simple-coded) and stress preference (*strpref*, with two levels: unanimous, majority; dummy-coded) as fixed factors, along with random intercepts for participants and items and random slopes of the two fixed factors and their interaction per participant. In R notation, this was coded as

```
strsuf ~ strpos * strpref + ( strpos * strpref | subject ) + ( 1 | item ).
```

In the no diacritic condition there was a significant intercept ($\beta = 1.87$, $z = 10.36$, $p < .001$), indicating stress assignment consistent with the suffix rather than equally distributed. For unanimous stress pairs there was no effect of suffix stress position (final vs. penultimate: $\beta = -0.44$, $z = -0.93$, $p = .351$; antepenultimate vs. penultimate: β

= 0.26, $z = 0.60$, $p = .548$). However, there was an effect of stress preference ($\beta = -1.19$, $z = -3.42$, $p < .001$), indicating that majority-stress items produced fewer suffix-consistent stress assignments than unanimous-stress items. The interaction between stress position and stress preference was not significant (final vs. penultimate: $\beta = 1.20$, $z = 1.69$, $p = .092$; antepenultimate vs. penultimate: $\beta = -0.51$, $z = -0.55$, $p = .583$).

In the congruent diacritic condition, again there was a significant intercept ($\beta = 19.44$, $z = 5.02$, $p < .001$), indicating stress assignment consistent with the suffix and with the diacritic. There was no significant effect of stress position (final vs. penultimate: $\beta = -1.60$, $z = -0.17$, $p = .863$; antepenultimate vs. penultimate: $\beta = 1.97$, $z = 0.22$, $p = .825$) or stress preference ($\beta = 12.59$, $z = 1.11$, $p = .267$) and no interaction between the two (final vs. penultimate: $\beta = -3.59$, $z = -0.25$, $p = .803$; antepenultimate vs. penultimate: $\beta = 4.93$, $z = 0.16$, $p = .872$).

In the incongruent diacritic condition, there was a significant *negative* intercept ($\beta = -2.86$, $z = -8.49$, $p < .001$), indicating stress assignment predominantly inconsistent with the suffix. There was no significant effect of stress position (final vs. penultimate: $\beta = -0.06$, $z = -0.09$, $p = .925$; antepenultimate vs. penultimate: $\beta = -0.69$, $z = -1.37$, $p = .172$) for unanimous-stress items. There was a significant effect of stress preference ($\beta = -1.51$, $z = -3.07$, $p = .002$) and a significant interaction of stress preference with stress position (final vs. penultimate: $\beta = -4.45$, $z = -4.01$, $p < .001$; antepenultimate vs. penultimate: $\beta = 0.18$, $z = 0.15$, $p = .881$).

To examine whether this pattern of results in the incongruent diacritic condition is better accounted for by the position of the diacritic than by the suffix, we fit the same model after replacing the dependent variable with a new binomial variable expressing stress assignments consistent with the diacritic. There was now a significant positive intercept ($\beta = 2.43$, $z = 7.09$, $p < .001$), indicating stress assignment predominantly

consistent with the diacritic. Again, there was no significant effect of stress position (final vs. penultimate: $\beta = -0.43$, $z = -0.81$, $p = .416$; antepenultimate vs. penultimate: $\beta = 0.005$, $z = 0.01$, $p = .992$) for unanimous-stress items. There was a significant effect of stress preference ($\beta = 1.86$, $z = 3.87$, $p < .001$) and a significant interaction of stress preference with stress position (final vs. penultimate: $\beta = 4.99$, $z = 4.65$, $p < .001$; antepenultimate vs. penultimate: $\beta = 0.48$, $z = 0.43$, $p = .665$), due to majority-stress items with a penultimate-stress suffix, which were assigned stress consistent with the diacritic somewhat less reliably. This effect is shown in Figure 1 (right).

Discussion

The results showed that word-final letter sequences significantly affected stress assignment to pseudowords in the absence of a diacritic. In this condition, the mean proportion of responses consistent with the suffix was 76% for unanimous stress items and 64% for majority stress items. This difference was statistically significant. The substantial proportion of responses inconsistent with the suffix and the difference between unanimous and majority stress items suggest that stress assignment is probabilistically associated with item-final letter sequences and not absolutely based on morphological accenting.

Our results for Greek are consistent with studies in English, in which probabilistic associations have been documented between word-final letter sequences and stress patterns and, moreover, behavioral evidence indicates that these associations are operative in stress assignment by children and adult readers (Arciuli & Cupples, 2006; Arciuli et al., 2010; Monaghan et al., 2016; Wade-Woolley & Heggie, 2015), subject to a developmental progression that can be accounted for by a statistical (connectionist) model learning to assign stress by exposure to a word corpus (Arciuli et al., 2010). Although the Greek stress system is very different from the English one in

many ways, including phonologically, phonetically, and orthographically (see, e.g., Protopapas, Panagaki, Andrikopoulou, Gutiérrez Palma, & Arvaniti, in press), it seems that stress assignment behavior in both languages is consistent with cumulative effects of probabilistic cues, based on lexical experience and knowledge, rather than absolute stress assignment rules associated with specific morphemes or other units. In particular, our findings do not support the notion that so-called accenting morphemes are strongly associated with specific stress patterns in the sense of strictly determining stress assignment behavior when reading aloud novel items (pseudowords). Therefore—and in agreement with Arciuli et al. (2010)—we submit that the available empirical evidence is not consistent with a view of stress assignment based on the operation of absolute, suffix-specific rules, arising either from linguistic or from lexical analysis.

Sensitivity to multiple probabilistic cues is hardly a novel idea in language learning and psycholinguistics (Chater & Manning, 2006; Christiansen, 2013). The more general notion of statistical learning has been tied to individual differences in language skill (e.g., Arciuli & von Koss Torkildsen, 2012; Kidd, & Arciuli, 2016), language representations in the brain (e.g., Arciuli, McMahon, & de Zubicaray, 2012), and literacy (e.g., Arciuli & Simpson, 2012; Gabay, Thiessen, & Holt, 2015; Spencer, Kaschak, Jones, & Lonigan, 2015). Although many studies of statistical learning have mainly relied on sequential learning paradigms involving tracking of transitional probabilities, there are also other paradigms involving multiple probabilistic cues that have been linked to individual differences (and difficulties) in language and literacy (see discussion in Krishnan, Watkins, & Bishop, 2016). In particular, there is strong evidence for statistical learning of spelling patterns across ages and languages (e.g., Deacon, Conrad, & Pacton, 2008; Samara & Caravolas, 2014) and there is evidence that children are sensitive to

spelling patterns even before they can spell phonologically (Pollo, Kessler, & Treiman, 2009). In this context, the suggestion that stress assignment in reading operates on the basis of multiple probabilistic constraints fits within a productive and well-supported general theoretical framework for language processing.

In contrast, our findings do not seem to be amenable to modeling approaches akin to the dual route formulation of Rastle and Coltheart (2000), who hard-coded absolute rules specific to particular prefixes and suffixes into their model for reading English, thereby defining “regularly” and “irregularly” stressed items by reference to those rules. Notably, the rules in the model of Rastle and Coltheart (2000) were based on an analysis of word types, that is, on observed probabilistic associations between stress patterns and word beginnings and endings. Moreover, the rules were applied to model behavioral data, again probabilistically associating stress patterns with particular items across participants. To posit a level of theoretically absolute rules, intervening between two levels of probabilistic observations, seems not only unparsimonious but also empirically unjustified. Specifically, the performance of the model incorporating stress assignment rules was 89.7% correct. This seems impressive but it should be judged against the baseline of 83.2% correct that would be achieved simply by stressing every item on the first syllable, due to the great asymmetry in English stress patterns. Thus it seems that the notion of rule-based stress assignment appears weak in English as well as cross-linguistically.

However, our findings cannot be interpreted as entirely precluding a role for morphological processing in Greek. In particular, these findings cannot be taken as evidence against the accenting properties of morphemes as proposed by linguistic analysis. Such properties may be used to account for the observed distribution of stress patterns over derived words in a language, evident in word production, and need not

participate in processing mechanisms such as reading aloud. The processing demands of reading aloud (which involves visual word recognition) may be sufficiently distinct from those of spontaneous word production, or of productive derivation, so that each may depend on partially different representations, at different levels of abstraction and with different formal and statistical properties (cf. Protopapas et al., in press).

The moderation of suffix-consistent stress assignment by the existence of stress competitors, which was observed in our data, may conceivably reflect the operation of two conflicting processes: One process tends to assign stress based on the accenting properties of the suffix while another process may create opposition on the basis of the statistical association between orthographic and phonological (stress) patterns in the lexicon. At present we do not have any specific evidence for the operation of the former (i.e., the morphological) process in Greek reading aloud, which remains to be investigated in future studies. However, such a hybrid approach is theoretically unparsimonious, because the main rationale behind positing a formal rule-based operation is the productivity as observed in highly systematic associations in the data. If the actual empirical situation suggests that patterns are only probabilistically associated, then the rule-based approach loses much of its driving force, as it does not seem to do any theoretically useful work. Therefore, even though our data cannot preclude the possibility of a morphological operation working against, or in combination with, a probabilistic process, it seems theoretically preferable to side with Arciuli et al. (2010) in rejecting this possibility in favor of statistical processing alone.

Our findings suggest a potential reinterpretation of the well-established neighborhood consistency effects typically observed in studies of reading Italian (Sulpizio et al., 2015). Although the notion of stress neighborhoods seems well motivated and empirically supported, it has not been specifically investigated whether

the operative domain must be precisely limited to the last two syllables of the words, excluding the onset of the penult, or whether it may be more flexibly extended or abridged. It is conceivable that smaller or larger word parts may be taken into account in the processing of written words by Italian readers, to reflect the grain size of associations between stress patterns and letter groups as they happen to have formed over the course of evolution of the language, which may not always align with the boundary between the onset and the nucleus of the penultimate syllable. Indeed, the analysis of Monaghan et al. (2016) suggests that letter groups of various sizes are associated with stress patterns in Italian, not necessarily limited to the stress neighborhood domain. Note that freedom from the two-syllable restriction does not imply another fixed domain of a different size. Rather, letter groups of arbitrary sizes may conceivably exert a simultaneous probabilistic influence to the extent they are associated with stress patterns. The cumulative influence of all reliably associated units will then be observable as systematic stress assignment behavior. This proposal is akin to the proposal of learning orthographic patterns and graphophonemic mappings of multiple lengths simultaneously, as a result of systematicities occurring at different grain sizes, leading to sight word reading as well as to sublexical orthographic and decoding knowledge (Ehri, 2005, 2014; see also Apel, 2009, 2011).

Along these lines, a computational model of reading aloud Italian, including an explicit representation of stress patterns but no representation of stress neighborhoods or morphemes, has successfully simulated both stress consistency and morphological constituency effects (Perry, Ziegler, & Zorzi, 2014). In the model, these effects presumably arose from interactions between a connectionist network performing graphophonemic mappings and a representation of arbitrary lexical mappings from orthography to phonology. This finding confirms that explicit representation of special

processing units, although theoretically attractive, is not necessary when probabilistic information is effectively coded indirectly in large-scale representations of lexical and sublexical mappings. Thus, the notion of stress neighborhood, useful as it has been proven to be, may amount to a higher-level statistical association insofar as the orthography-phonology covariation involving stress is typically (but not necessarily always) captured at this grain size. Further empirical investigation, both corpus-based and behavioral, will be required to resolve this issue.

Our data suggest that probabilistic letter cues to stress assignment in Greek are relatively weak, insofar as they were trumped by the diacritic: When unanimous stress items were presented with a suffix-incongruent diacritic, the mean proportion of responses consistent with the suffix was only 18%; for majority stress items, the mean proportion was 12% (in the face of great individual variability, seen in Figure 1, left). However, probabilistic letter cues evidently override the influence from the default pattern, as there were no significant differences in stress assignment proportions consistent with the suffix between items with suffixes stressed on different syllables. If there were a substantial default effect, we should have observed more assignments consistent with the suffix when stressed on the penultimate than when stressed on other syllables, because the penultimate would be supported by both the suffix and the default. There was, however, an effect of suffix stress position for majority stress items presented with an incongruent diacritic. Specifically, an incongruent diacritic was slightly less effective when the suffix was stressed on the penultimate, consistent with a stronger “pull” toward the penultimate. Therefore, when the probabilistic letter cue is weak, due to the existence of stress competitors, the conflict from an incongruent diacritic apparently permits a weak influence of the default pattern to surface.

In a study of stress assignment in pseudoword reading by adult skilled readers,

Protopapas et al. (2007, p. 708) reported the relative strength of stress information sources, when uncontested by other source, as follows: default 60%, lexical 80%, diacritic 90%. Note that this underestimates the typical force of the lexical source because the estimates were derived from pseudowords, which, even though resembling specific words well enough to elicit their activation, were clearly mismatching the words, evidenced by their segmentally correct pronunciation. Therefore in typical word reading the lexicon is probably the dominant source of stress information. Taking into account the findings of the present study, the order of importance of stress assignment information sources for Greek seems to be: lexical > diacritic > final letters > default.

Our conclusions and suggestions may be tempered by the fact that we could only examine a limited set of suffixes, which was unbalanced with respect to the critical distinction between unanimous and majority items. This was an unavoidable consequence of applying criteria meant to exclude suffixes that would be unlikely to produce the sought-after stress assignment effect. That is, suffixes appearing in only a few word types of tokens and suffixes with too many stress competitors might have unfairly disadvantaged the morphological accenting hypothesis. In this sense we can be relatively confident that probabilistic association is more likely than absolute morphological accenting, because the deck was not stacked in the direction of this outcome. However, it remains possible that the small, and largely forced, set of suffixes examined, especially the majority items, may have resulted in a finding that somehow depends on (some of the) particular suffixes rather than on a general psycholinguistic processing principle. We believe that this limitation is not very severe, as our conclusion is consistent with the interpretation of empirical psycholinguistic observations in other studies across languages, even if it is at somewhat odds with certain linguistic analyses.

In conclusion, in the present study we have documented an additional source of

stress assignment information in reading aloud Greek. Specifically, probabilistic word-final information associated with accenting derivational suffixes was found to affect reading behavior, especially in the absence of lexical and diacritic information. Although statistical associations between word endings and stress patterns have been previously documented in other languages, it is not known in most cases whether this kind of information is operative in actual reading processes. Our finding enriches our understanding of stress assignment and suggests that the reading system can be sensitive to any relevant information, weighing each potential source to achieve optimal performance.

References

- Apel, K. (2009). The acquisition of mental orthographic representations for reading and spelling development. *Communication Disorders Quarterly, 31*, 42–52.
doi:10.1177/1525740108325553
- Apel, K. (2011). What is orthographic knowledge? *Language, Speech, and Hearing Services in Schools, 42*, 592–603. doi:10.1044/0161-1461(2011/10-0085)
- Arciuli, J., & Cupples, L. (2006). The processing of lexical stress during visual word recognition: Typicality effects and orthographic correlates. *The Quarterly Journal of Experimental Psychology, 59*, 920–948. doi:10.1080/02724980443000782
- Arciuli, J., McMahan, K., & de Zubicaray, G. (2012). Probabilistic orthographic cues to grammatical category in the brain. *Brain and Language, 123*, 202–210.
doi:10.1016/j.bandl.2012.09.009
- Arciuli, J., Monaghan, P., & Seva, N. (2010). Learning to assign lexical stress during reading aloud: Corpus, behavioral, and computational investigations. *Journal of Memory and Language, 63*, 180–196. doi:10.1016/j.jml.2010.03.005
- Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science, 36*, 286–304. doi:10.1111/j.1551-6709.2011.01200.x
- Arciuli, J., & von Koss Torkildsen, J. (2012). Advancing our understanding of the link between statistical learning and language acquisition: The need for longitudinal data. *Frontiers in Psychology, 3*, 324. doi:10.3389/fpsyg.2012.00324
- Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge University Press.
doi:10.1017/cbo9780511801686
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–

412. doi:10.1016/j.jml.2007.12.005

Beyersmann, E., Ziegler, J. C., Castles, A., Coltheart, M., Kezilas, Y., & Grainger, J. (2016).

Morpho-orthographic segmentation without semantics. *Psychonomic Bulletin & Review*, *23*, 533–539. doi:10.3758/s13423-015-0927-z

Babiniotis, G. (2006). *Μικρό λεξικό της νέας ελληνικής γλώσσας [Small Dictionary of the Modern Greek language]*. Athens: Kentro Lexikologias.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. R package v. 1.0-5. <http://CRAN.R-project.org/package=lme4>.

Burani, C., & Arduino, L. S. (2004). Stress regularity or consistency? Reading aloud Italian polysyllables with different stress patterns. *Brain and Language*, *90*, 318–325. doi:10.1016/S0093-934X(03)00444-9

Burani, C., Arduino, L. S., & Marcolini, S. (2006). Naming morphologically complex pseudowords: A headstart for the root? *The Mental Lexicon*, *1*, 299–327. doi:10.1075/ml.1.2.07bur

Burani, C., Marcolini, S., De Luca, M., & Zoccolotti, P. (2008). Morpheme-based reading aloud: Evidence from dyslexic and skilled Italian readers. *Cognition*, *108*, 243–262. doi:10.1016/j.cognition.2007.12.010

Burani, C., Paizi, D., & Sulpizio, S. (2014). Stress assignment in reading Italian: Friendship outweighs dominance. *Memory & Cognition*, *42*, 662–675. doi:10.3758/s13421-013-0379-5

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and

acquisition. *Trends in Cognitive Sciences*, 10, 335–344.

doi:10.1016/j.tics.2006.05.006

Christiansen, M. H. (2013). Language has evolved to depend on multiple-cue integration.

In R. Botha & M. Everaert (Eds.), *The evolutionary emergence of language:*

Evidence and inference (pp. 42–61). Oxford University Press.

doi:10.1093/acprof:oso/9780199654840.003.0003

Colombo, L. (1992). Lexical stress effect and its interaction with frequency in word

pronunciation. *Journal of Experimental Psychology: Human Perception and*

Performance, 18, 987–1003. doi:10.1037/0096-1523.18.4.987

Deacon, S. H., Conrad, N., & Pacton, S. (2008). A statistical learning perspective on

children's learning about graphotactic and morphological regularities in spelling.

Canadian Psychology/Psychologie Canadienne, 49, 118–124. doi:10.1037/0708-

5591.49.2.118

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory*

and Language, 59, 447–456. doi:10.1016/j.jml.2007.11.004

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies*

of Reading, 9, 167–188. doi:10.1207/s1532799xssr0902_4

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading,

spelling memory, and vocabulary learning. *Scientific Studies of Reading*, 18, 5–21.

doi:10.1080/10888438.2013.819356

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with

millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35,

116–124. doi:10.3758/BF03195503

Holton, D., Mackridge, P. & Philippaki-Warburton, I. (1999). *Γραμματική της ελληνικής*

γλώσσας [Grammar of the Greek language]. Athens: Patakis.

- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences, 19*, 117–125. doi:10.1016/j.tics.2014.12.010
- Gabay, Y., Thiessen, E. D., & Holt, L. L. (2015). Impaired statistical learning in developmental dyslexia. *Journal of Speech, Language, and Hearing Research, 58*, 934–945. doi:10.1044/2015_JSLHR-L-14-0324
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446. doi:10.1016/j.jml.2007.11.007
- Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development, 87*, 184–193. doi:10.1111/cdev.12461
- Krishnan, S., Watkins, K. E., & Bishop, D. V. (2016). Neurobiological Basis of Language Learning Difficulties. *Trends in Cognitive Sciences, 20*, 701–714. doi:10.1016/j.tics.2016.06.012
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models. R package version 2.0-11. <http://CRAN.R-project.org/package=lmerTest>
- Monaghan, P., Arciuli, J., & Seva, N. (2016). Cross-linguistic evidence for probabilistic orthographic cues to lexical stress. In J. Thomson & L. Jarmulowics (Eds.), *Linguistic rhythm and literacy* (pp. 215–236). Amsterdam: John Benjamins. doi:10.1075/tilar.17.10mon
- Perry, C., Ziegler, J. C., & Zorzi, M. (2014). CDP++.Italian: Modeling sublexical and supralexical inconsistency in a shallow orthography. *Plos ONE, 9*, 4. e94291. doi:10.1371/journal.pone.0094291

- Pollo, T. C., Kessler, B., & Treiman, R. (2009). Statistical patterns in children's early writing. *Journal of Experimental Child Psychology, 104*, 410–426.
doi:10.1016/j.jecp.2009.07.003
- Protopapas, A. (2006). On the use and usefulness of stress diacritics in reading Greek. *Reading and Writing: An Interdisciplinary Journal, 19*, 171–198.
doi:10.1007/s11145-005-4023-z
- Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39*, 859–862. doi:10.3758/BF03192979
- Protopapas, A. (2015). An introductory explanation of contrast coding in R linear models. Retrieved from <https://rpubs.com/protopap/84177>.
- Protopapas, A. (2016). From diacritics to the mental lexicon: Where is the stress? In J. Thomson & L. Jarmulowics (Eds.), *Linguistic rhythm and literacy* (pp. 237–264). John Benjamins. doi:10.1075/tilar.17.11pro
- Protopapas, A., Gerakaki, S., & Alexandri, S. (2006). Lexical and default stress assignment in reading Greek. *Journal of Research in Reading, 29*, 418–432.
doi:10.1111/j.1467-9817.2006.00316.x
- Protopapas, A., Gerakaki, S., & Alexandri, S. (2007). Sources of information for stress assignment in reading Greek. *Applied Psycholinguistics, 28*, 695–720.
doi:10.1017/S0142716407070373
- Protopapas, A., & Gerakaki, S. (2009). Development of processing stress diacritics in reading Greek. *Scientific Studies of Reading, 13*, 453–483.
doi:10.1080/10888430903034788
- Protopapas, A., Panagaki, E., Andrikopoulou, A., Gutiérrez Palma, N., & Arvaniti, A. (in press). Priming stress patterns in word recognition. *Journal of Experimental*

- Psychology: Human Perception & Performance*. doi:10.1037/xhp0000259
- Protopapas, A., Tzakosta, M., Chalamandaris, A., & Tsiakoulis, P. (2012). IPLR: An online resource for Greek word-level and sublexical information. *Language Resources & Evaluation*, 46, 449–459. doi:10.1007/s10579-010-9130-z
- Protopapas, A., & Vlahou, E. L. (2009). A comparative quantitative analysis of Greek orthographic transparency. *Behavior Research Methods*, 41, 991–1008. doi:10.3758/BRM.41.4.991
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Ralli, A. (2005). *Μορφολογία [Morphology]*. Athens: Patakis.
- Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language*, 42, 342–364. doi:10.1006/jmla.1999.2687
- Revithiadou, A. (1999). *Headmost Accent Wins: Head Dominance and Ideal Prosodic Form in Lexical Accent Systems*. LOT Dissertation Series 15 (HIL/Leiden Universiteit). Holland Academic Graphics, The Hague.
- Samara, A., & Caravolas, M. (2014). Statistical learning of novel graphotactic constraints in children and adults. *Journal of Experimental Child Psychology*, 121, 137–155. doi:10.1016/j.jecp.2013.11.009
- Ševa, N., Monaghan, P., & Arciuli, J. (2009). Stressing what is important: Orthographic cues and lexical stress assignment. *Journal of Neurolinguistics*, 22, 237–249. doi:10.1016/j.jneuroling.2008.09.002
- Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and Writing*, 28, 467–490.

doi:10.1007/s11145-014-9533-0

- Sulpizio, S., Arduino, L. S., Paizi, D., & Burani, C. (2013). Stress assignment in reading Italian polysyllabic pseudowords. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 51–68. doi:10.1037/a0028472
- Sulpizio, S., Burani, C., & Colombo, L. (2015). The process of stress assignment in reading aloud: Critical issues from studies on Italian. *Scientific Studies of Reading*, *19*, 5–20. doi:10.1080/10888438.2014.976340
- Sulpizio, S., & Colombo, L. (2013). Lexical stress, frequency, and stress neighbourhood effects in the early stages of Italian reading development. *The Quarterly Journal of Experimental Psychology*, *66*, 2073–2084. doi:10.1080/17470218.2013.785577
- Triandafillidis, M. (1997). *Νεοελληνική γραμματική [Modern Greek Grammar]*. Athens: Organismos Ekdoseos Didaktikon Vivlion.
- UCLA Statistical Consulting Group (n.d.). Regression with SAS, Chapter 5: Additional coding systems for categorical variables in regression analysis. Retrieved from <http://www.ats.ucla.edu/stat/sas/webbooks/reg/chapter5/sasreg5.htm>.
- Wade-Woolley, L., & Heggie, L. (2015). Implicit knowledge of word stress and derivational morphology guides skilled readers' decoding of multisyllabic words. *Scientific Studies of Reading*, *19*, 21–30. doi:10.1080/10888438.2014.947647

Appendix

List of stimuli and suffixes

Experimental items (suffixed pseudowords). νοτσαθειψα, θεπαζοευσα, αβοσπισιμο, ντεδισκοσιμο, ολαμεσημο, πτοννοκασημο, θναξιμο, ομποννεξιμο, φετωρισσα, πεκισσα, γαντεπουλο, φεζοραπουλο, ετογκλισμα, θαεφασμα, ελοφασμα, επαζοτητα, κροβυντοτητα, νοκυτητα, χνωκυτητα, αχθυτητα, εμορμασματα, ονωδισματα, γουλιδισματα, κιντομενος, χνορομενος, γοδακομενος, ολτουμενος, γολουμενος, φοθαγκοτερος, ζεφολοτατος, γαρθολευω, σηκνεβιζω, ιφλεπαδα, ογλετηρι, ζειλατηρι, κονοχηρας, χαμοριτσα, θοντολοχαδες, θενετουχος, φειβεας, θωγκαιας, σλαποριστας, ισκοσμενος, μπρωλατεος, δοπορεωδης, ρολποσυνη, λεαμποσυνη, χομεροσυνη, νιβδοσυνη, ζακτοτητων, καλεγκοτητων, μεδοροτητων, μεκωμενος, ικνωμενος, ποιχωμενος, ενεμπωμενος, ερβοδρωμενοι, σχορβατωμενοι, μομαιντωμενοι, οσπονωμενοι, οδρενισμος, ερεγκατισμος, ισκολισμός, χοντορικασμος, αντανιγενες, οχνιγενες, νολεραγενες, οντραναιδης, γαποταειδης, ζινολαιδης, ιλατικός, οντοχτικός, μοτσατικός, βοντοτικός, βεθατικός, θαοδριστικός, ναεδιστικός, βετσεντιστικός, τεσιςτικός, κοσχοριστικες, ιβατιστικες, ναμεγιστικες, μοκοριστικες, γοντεμπιστικες, ζαπατικως, πεδαπατικως, ραεπατικως, χαμποζατικός, σταπαγαλατικός, μεντοχτατικός.

Table A1

List of suffixes

Suffix	Number of syllables	Stress position	Suffixed words		Stress competitors	
			Type frequency	Token frequency	Type frequency	Token frequency
<i>A. Suffixes without stress competitors</i>						
-ισμός	2	1	574	24946		
-ασμός	2	1	99	3583		
-ειδής	2	1	50	176		
-γενές	2	1	23	277		
-εύω	2	2	104	6033		
-ίζω	2	2	182	9039		
-άδα	2	2	133	30957		
-τήρι	2	2	30	1183		
-τήρας	2	2	46	1622		
-ίτσα	2	2	31	547		
-άδες	2	2	256	22385		
-ούχος	2	2	40	1604		
-ίστας	2	2	42	953		
-τέος	2	2	18	128		
-ώδης	2	2	103	1374		
-εψα	2	3	48	485		
-ευσα	2	3	34	74		
-σιμο	2	3	113	4674		
-σημο	2	3	23	1680		
-ξιμο	2	3	32	675		
-ισσα	2	3	65	1269		
-πουλο	2	3	22	225		
-ισμα	2	3	266	9317		
-ασμα	2	3	101	5568		
-ατικός	3	1	155	3513		
-ατικώς	3	1	20	165		
-οτήτων	3	2	90	6076		
-ότητα	3	3	539	69897		
-ύτητα	3	3	16	3439		
-ίσματα	3	3	140	3478		
-άσματα	3	3	63	3781		
-ούμενος	3	3	135	2762		
-ότερος	3	3	167	2119		
-ότατος	3	3	74	257		
<i>B. Suffixes with stress competitors</i>						
-τικός	2	1	822	18079	31	123
-έας	2	2	90	16113	5	317
-αίας	2	2	66	4513	4	291
-σμένος	2	2	406	5946	1	8

-ιστικός	3	1	177	1193	5	8
-ιστικές	3	1	206	6185	14	62
-οσύνη	3	2	60	4165	3	54
-ωμένος	3	2	269	2809	9	119
-ωμένοι	3	2	245	2549	10	63
-όμενος	3	3	268	6087	4	85

Table 1

Summary metrics of selected suffixes and their stress competitors

Number of syllables	Stressed syllable	Derivational suffixes			Stress competitors		
		Number of Suffixes	Mean types	Mean tokens	Number of suffixes	Mean types	Mean tokens
2	3	9	78.2	2663.0	0	–	–
2	2	14	110.5	7314.1	3	3.3	205.3
2	1	5	313.6	9412.2	1	31.0	123.0
3	3	8	175.3	11477.5	1	4.0	85.0
3	2	4	166.0	3899.8	3	7.3	78.7
3	1	4	139.5	2764.0	2	9.5	35.0

Note: Stress syllable counted from the end (1=final, 2=penultimate, 3=antepenultimate). Types refer to the number of word forms with these suffixes (or final letter sequence); tokens refer to the number of occurrences of these word forms in the corpus.

Table 2

Response times (ms) in the three diacritic conditions for each type of item, in total and broken down relative to suffix preference and diacritic position

Response stress pattern	Unanimous stress items						Majority stress items					
	Absent		Congruent		Incongruent		Absent		Congruent		Incongruent	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
All	952.6	306.5	936.3	273.3	967.8	293.7	942.7	312.8	910.0	261.3	925.2	298.3
<i>Relative to suffix:</i>												
Consistent	954.9	310.7	936.1	273.0	988.3	278.1	957.3	308.4	912.3	262.9	961.6	381.7
Inconsistent	943.4	289.8	945.9	292.6	963.5	296.8	915.4	320.0	845.1	209.2	920.2	285.4
<i>Relative to diacritic:</i>												
Consistent			936.1	273.0	961.2	292.7			912.3	262.9	920.2	285.4
Inconsistent			945.9	292.6	992.7	296.8			845.1	209.2	961.6	381.7

Note: Response times shown for correct (i.e., segmentally accurate) responses only.

Table 3

Proportion of stress assignment (%) in the three diacritic conditions as a function of suffix stress position

Table 3a. All responses												
Suffix stress	Absent				Congruent				Incongruent			
	NA	1	2	3	NA	1	2	3	NA	1	2	3
<i>Unanimous stress items</i>												
1	3.0	71.7	7.8	17.5	6.4	91.7	0.8	1.1	5.0	17.5	54.6	23.0
2	2.9	7.1	77.6	12.4	1.3	0.8	95.3	2.6	4.5	46.1	19.5	30.0
3	5.7	0.8	16.6	77.0	3.7	0.0	1.6	94.7	6.6	40.5	34.7	18.1
<i>Majority stress items</i>												
1	1.0	71.3	2.4	25.4	2.9	94.3	0.5	2.4	1.9	6.2	0.0	91.9
2	2.6	1.1	57.9	38.4	3.2	0.0	91.1	5.8	2.6	10.5	18.9	67.9
3	1.8	0.0	45.6	52.6	1.8	0.0	3.5	94.7	7.0	0.0	82.5	10.5

Table 3b. Correct (i.e., segmentally accurate) responses only												
Suffix stress	Absent			Congruent			Incongruent					
	1	2	3	1	2	3	1	2	3			
<i>Unanimous stress items</i>												
1	61.8	7.5	16.1	80.9	0.8	1.1	13.0	51.2	19.7			
2	6.1	72.1	10.8	0.5	86.8	2.4	40.0	15.8	26.3			
3	0.8	13.1	69.4	0.0	1.2	84.0	36.6	31.6	13.6			
<i>Majority stress items</i>												
1	67.5	2.4	23.4	89.0	0.0	2.4	5.3	0.0	86.6			
2	0.5	52.6	34.7	0.0	83.2	4.2	10.5	17.4	62.6			
3	0.0	43.9	52.6	0.0	1.8	89.5	0.0	82.5	10.5			

Note: Stress positions in syllables, counted from the end: 1=final, 2=penultimate, 3=antepenultimate. NA=stress not assigned.

Figure Captions

Figure 1. Left, proportion of stress assignment responses consistent with the suffix in each diacritic condition, averaged over suffix stress positions. Right, proportion of stress assignment responses consistent with the diacritic in the incongruent diacritic condition, for each suffix stress position. In both panels, boxes enclose the second and third quartiles; the thick link indicates the median; error bars extend to the full range.

