Lexical and default stress assignment in reading Greek

Athanassios Protopapas

Institute for Language & Speech Processing

Svetlana Gerakaki and Stella Alexandri

University of Athens

Greek is a language with lexical stress that marks stress orthographically with a special diacritic. Thus, the orthography and the lexicon constitute potential sources of stress assignment information in addition to any possible general default metrical pattern. Here, we report two experiments with secondary education children reading aloud pseudo-word stimuli, in which we manipulated the availability of lexical (using stimuli resembling particular words) and visual (existence and placement of the diacritic) information. The reliance on the diacritic was found to be imperfect. Strong lexical effects as well as a default metrical pattern stressing the penultimate syllable were revealed. Reading models must be extended to account for multisyllabic word reading including, in particular, stress assignment based on the interplay among multiple possible sources of information.

Reading aloud entails the production of spoken words on the basis of visual stimuli. Even when words are read in isolation, such as in a list of word items, exclusive attention to the segmental level misses important factors that are part of the complete phonological specification. Perhaps the most salient of these factors concerns word stress, at least for stress-assigning languages. Current reading theories are not clear on where stress assignment information is derived from, and by what mechanism. For example, stress may be orthographically specified, but it may also be assigned by rule (based on phonological properties), by default (on a fixed position) or derived from the lexical representation (Colombo, 1992; Gutiérrez Palma, 2003; Schiller, Fikkert & Levelt, 2004; cf. Rastle & Coltheart, 2000).

The relevance of each option depends on the language: in languages with fixed stress, that is, where stress falls on the same syllable on every word (e.g. on the first or final syllable of every word, as in Finnish and French, respectively; Hirst & di Cristo, 1998), stress assignment in reading is not an issue. In languages with lexical stress, that is, where

stress may vary in position and contributes to lexical identity, stress may be assigned on a lexical or a sub-lexical basis, or a combination of these. For example, in Spanish, Italian and Portuguese, there is a 'regular' stress assignment, determined in part by syllable weight. Deviations from regularity are either marked orthographically or determined lexically (Colombo, 1992; Gutiérrez Palma, 2003; Nunes, Roazzi & Buarque, 2003). Thus, in these languages, there is evidence for alternative information sources affecting stress assignment during reading.

Greek offers a particularly interesting case for studying stress assignment in reading because it always marks stress position in the orthography with a special diacritic. From a phonological standpoint, there are no known restrictions on stress position other than the constraint that stress must fall on one of the last three syllables (Malikouti-Drachman & Drachman, 1989). Stress position is a lexical property, taking into account morphological type. That is, conjugations and declinations may affect stress position (Botinis, 1998; Malikouti-Drachman, 2002), because certain classes of inflectional morphemes ('accenting suffixes') determine stress placement in interaction with lexically assigned stress from word roots (Revithiadou, 2004).

In Greek spelling, contemporary rules dictate that every word with more than one syllable must bear a stress diacritic on the vowel of its stressed syllable (Petrounias, 2002). Greek words with two or more syllables written without a stress diacritic are thus considered misspelled, even though stress assignment can usually be guessed successfully from the phoneme sequence (Protopapas, 2006). Extending and complementing previous studies in Italian and Spanish, Greek allows investigation of stress assignment free from the structural (phonological) constraints that interact with default placement in those languages.

The existence of the diacritic as an obvious source of information does not imply that it is the preferred source. There must be a computational cost associated with decoding the diacritic from print and constructing an appropriate metrical frame properly aligned with the phonemic sequence. This cost should be compared with that of retrieving a stored metrical pattern from the lexical and morphological specification, in which it is presumably included. The lexicon is thus a second potential source of information. A third potential source is 'default' assignment. The hypothesis that Greek has a preferred stress assignment position on the penultimate is based on phonological (Malikouti-Drachman, 2002) and developmental (Kappa, 2002) considerations.

Our long-term goal is to understand stress assignment as a sub-process of reading, and to help improve cognitive models of reading by paying attention to the underlying mechanism(s). Although we are interested in the reading of words, it would be problematic to use words in studies at this early stage, because it would not be possible to manipulate the salience of lexically derived information. If stress can be based on lexical activation, then once words are activated stress is fully determined, so it would be difficult to study potential orthographic (i.e. diacritic) or structural (i.e. default) contributions. One could restrict testing to stress-ambiguous word pairs (i.e. segmentally identical words differentiated only by stress) but then lexical factors such as frequency could still affect selection. We have thus used non-word stimuli in order to be able to manipulate the availability of lexical and orthographic information.

Well-practised strategies that a reader uses for word reading can be expected to be applied to non-words. The success of these strategies will depend in part on their applicability and in part on their efficiency. If words are typically read such that stress assignment is based on lexical activation, then derivation of metrical patterns from the

lexicon will be highly practised and efficient. When a non-word resembles a word sufficiently (i.e. enough to activate the lexical entry for the word), it will be stressed by analogy to the activated word regardless of the position (or the existence) of the diacritic. According to this *lexical* hypothesis, the pseudo-word stimulus αποκελεσμα, which resembles the real word αποτέλεσμα (meaning 'result'), should be pronounced with stress on the antepenultimate, despite the lack of a written diacritic.

When a non-word does not activate any lexical items sufficiently, then the lexical strategy will not apply and stress assignment failures should be observed. Thus, the non-word εποτεγετζα, also derived from, but not resembling, αποτέλεσμα, should not be consistently stressed on the antepenultimate. Failures of lexical access can take the form of either random stress assignment or preferential assignment on a particular syllable. In the latter case, we would have evidence for a default strategy, operative in the absence of specific stress position indicators. If εποτεγετζα is usually pronounced with penultimate stress, a metrical pattern not attributable either to a diacritic (which is not presented) or to a word (because of dissimilarity), then a default non-lexical cause can be invoked.

If, on the other hand, in regular word reading words are typically read such that stress assignment is based on the written diacritic, then decoding of the diacritic will be highly practised, hence efficient, and because this strategy can also be used with non-words it is expected that non-words will always be stressed on the syllable indicated by the diacritic. A finding that non-words are not reliably stressed where the diacritic indicates would therefore be inconsistent with the existence of an efficient decoding strategy for word stress assignment based on the diacritic. That is, any time εποτέγετζα (presented with the diacritic as shown) is pronounced with stress not on the antepenultimate, this 'stress error' must be interpreted as a failure to decode the diacritic. If there is a tendency for such errors to be made towards a specific syllable, this again would constitute evidence for a default non-lexical cause, which can become active in cases of (lexical or orthographic) assignment failure.

In a recent study of reading skill assessment in Greek (Protopapas, 2006), seventh-grade schoolchildren were found to make many stress assignment errors when reading pseudo-words but not when reading real words. The great majority of stress errors were made towards the penultimate syllable, suggesting a default metrical pattern in the form of a word-final trochee. Following up, here we report experiments revealing the effects of alternative sources of stress assignment, using non-word materials specifically constructed to manipulate and contrast orthographic and lexical information.

Experiment 1

The first experiment was designed to test the hypothesis that school-age children, when reading non-words, activate and use metrical patterns based on words in their mental lexicon. If this hypothesis is correct, then non-words resembling specific words should be stressed by analogy to the words they resemble, whereas non-words not resembling specific words would be stressed either randomly or by reference to a default stress pattern. Moreover, if stress assignment is a result of lexical activation, then non-words resembling frequent words might be stressed by analogy to these words more often than non-words resembling infrequent words, because frequent words can be expected to be activated more easily and more strongly than infrequent words.

The design of this study involves three groups of pseudo-words constructed by modifying real words. Group 1 includes pseudo-words based on high-frequency words that were minimally changed; Group 2 on high-frequency words that were changed beyond recognition; and Group 3 on low-frequency words that were minimally changed. These pseudo-words were presented without a stress diacritic to be read aloud by the participants. Lack of diacritic implies lack of explicit orthographic information for stress assignment.

If Group 1 items are pronounced with stress on the same position on which the words they were derived from were stressed, and Group 2 items do not follow this pattern, this is strong evidence that lexical information, via similarity to the words, contributes stress assignment information. Further, if Group 1 items are stressed on the same position as their source words more than Group 3 items, this would be consistent with the hypothesis that stress assignment information is mediated by gradual lexical activation, which we expect to be sensitive to word frequency. Finally, any consistent trends in stress assignment for Group 2 items, for which both orthographic and lexical information is lacking, would be consistent with the existence of a default metrical pattern for stress assignment.

Method

Participants

Thirty-seven children from grades 7 to 9 participated, including 13 boys and 24 girls, 12–15 years old (mean age 13.3, *SD* 0.9 years). Participation was voluntary and no compensation was provided.

Stimuli

Ninety-two real words three to five syllables long, with 0 to two consonant clusters, formed the initial set, including low-frequency and high-frequency items. Written frequencies were obtained from the Hellenic National Corpus (HNC; at http://hnc.ilsp.gr/en; Hatzigeorgiu et al., 2000), which contains 34 million words from a collection of literary, legal and journalistic texts. Here, 'low-frequency' refers to items with counts between 0.1 and 24 per million words, and 'high-frequency' to 139–607 per million. Each word was then turned into a non-word by changing one or more phonemes so that the result remained phonotactically acceptable. For consonants, phoneme changes were made on one or more of the three main phonetic dimensions: place of articulation, manner of articulation and voicing. To quantify the phonetic difference, a total change score was calculated for each word by adding the number of changed dimensions for each changed consonant, plus two for each changed vowel; added or removed consonants contributed three points.

About half of the high-frequency words were changed considerably, to become unrecognisable, and received change scores 5–10; the rest were changed minimally (on a single segment) and received change scores 1–3. As an example, the word αποκέλεσμα (pronounced /apoˈtɛlɛzma/) was minimally changed to form αποκελεσμα (/apoˈcɛlɛzma/) by a single consonant place change (/t/ to /c/), receiving a phonetic difference score of 1. The same word was also used to construct εποτεγετζα (/ɛpoˈtɛjɛdza/), by several changes

(vowel a → ε, place and manner in l → j, manner in z → dz, and deletion of m), receiving a score of 8.

The resulting non-words were submitted to a pre-test, in which 10 university students were presented with the printed list and asked to write down, for each item, a word that comes to mind. Minimally changed non-words were discarded if two or more persons failed to produce the exact original word, or if even one person produced a word stressed on a different syllable. Highly changed non-words were discarded if any words at all were produced for them. There remained 15 minimally changed non-words from high-frequency words (Group 1), 14 highly changed non-words from high-frequency words (Group 2) and 15 minimally changed non-words from low-frequency words (Group 3). Table 1 shows the characteristics of the item groups. The full set of 44 non-words is listed in the Appendix (Table A1).

The 44 items were printed on a single sheet of A4 paper in two columns, in Times Roman 14pt, in a randomised order (the same for all participants), without stress diacritics.

Procedure

Children were asked to read aloud the pseudo-words. They were encouraged to read as quickly as possible without making mistakes. Stress or stress diacritics were not mentioned at all. Each child was tested individually in a quiet room at the school. Reading was tape-recorded and timed using a hand-held stopwatch. Both transcription and timing were subsequently verified from the recording.

Analyses

Phonetic accuracy in reading the pseudo-words was examined by counting segmental errors and analysing their probability of occurrence by group, length, complexity and original word stress position.

In order to examine the central question of stress assignment behaviour, we created and tested two specific contrasts: (a) the *lexical index* (LexI) was formed, by subtraction, from the mean number of times pseudo-words were stressed on the syllable on which the corresponding original word was stressed, minus the mean number of times stressed on any other syllable. LexI quantifies the tendency to follow the stress pattern of the original (source) words. Thus, a LexI of 1.00 indicates 100% stress assignment identical to the original word; 0.00 indicates no preference and -1.00 indicates a complete mismatch between non-word stress assignment and original word stress position. LexI is expected

Table 1. Characteristics of the three groups of non-words used in Experiments 1 and 2 (group means, with corresponding standard deviations in parentheses).

Group	Experiment	N	Word frequency	Phonetic difference	Stress position	Number of syllables	Consonant clusters
1	1	15	301 (126)	1.4 (0.6)	2.2 (0.8)	4.0 (0.8)	0.6 (0.7)
2	1	14	280 (99)	8.0 (1.4)	2.4 (0.7)	4.1 (0.8)	0.6 (0.6)
3	1	15	5 (7)	1.2 (0.4)	2.0 (0.8)	4.0 (0.8)	0.8 (0.7)
4	2	46	52 (118)	13.7 (5.3)	2.0 (0.8)	4.2 (0.7)	1.8 (0.8)

Note: Stress position indicates the stressed syllable of the source word, counting from the end (1 = final). Frequency (raw counts) refers to the written frequency of the source word, from the Hellenic National Corpus.

to be the highest when lexical influence is the strongest, that is, in Group 1 items, and 0 when no lexical influences are expected, that is, in Group 2. A difference in LexI between Groups 1 and 3 would indicate frequency effects on lexical influence. (b) The *penultimate index* (PenI) was formed, by subtraction, from the mean number of times pseudo-words were stressed on the penultimate minus the mean number of times stressed on any other syllable. PenI quantifies the tendency to use a penultimate stress pattern. Thus, a PenI of 1.00 indicates 100% stress assignment on the penultimate; 0.00 indicates equal proportions of penultimate and other and -1.00 would indicate that the item was never stressed on the penultimate. PenI is expected to be the highest when orthographical and lexical sources of information are missing, that is in Group 2 items, and the lowest when lexical sources are the strongest, that is, in Group 1.

All statistical testing was performed both with averaging first across non-words (subjects analysis) and with averaging first across participants (items analysis). Differences were considered statistically significant when both tests (indexed by 1 and 2, respectively) exceeded the usual criterion of $p < .05$. Occasional tests of particular interest meeting a less stringent criterion of $p < .10$ are mentioned as 'marginally significant'.

Results and discussion

Segmental accuracy

Participants read the 44 non-words on average in 118 seconds (*SD* 49 seconds, range 61–259 seconds). Excluding four missed responses, each response was rated for phonetic accuracy, receiving one point for each incorrect phoneme produced. In 3×3 analysis of segmental errors by stress position of the original word and group, there was no significant main effect of stress position, $F_1(2, 72) = 1.192$, $p = .310$; $F_2(2, 35) < 1$, and no interaction with item group, $F_1(4, 144) < 1$; $F_2(4, 35) < 1$. There was a significant effect of number of syllables on segmental errors, $F_1(2, 72) = 4.697$, $p = .012$; $F_2(2, 35) = 3.337$, $p = .047$, and a corresponding linear trend, $F_1(1, 36) = 8.152$, $p = .007$; $p_2 = .014$, in the absence of a quadratic trend ($F_1, F_2 < 1$). There was neither a main effect of number of consonant clusters, $F_1(2, 72) = 2.280$, $p = .110$; $F_2(2, 35) = 1.632$, $p = .210$, nor an interaction between clusters and number of syllables, $F_1(4, 144) = 1.995$, $p = .98$; $F_2(4, 35) = 1.112$, $p = .366$.

Therefore, because item group and word stress position did not affect the phonetic accuracy of reading aloud the non-words, any differences in stress assignment behaviour arising from these factors cannot be attributed to general difficulty in pronouncing the non-words.

Stress assignment

Table 2 shows the distribution of stress assignment responses in each condition, for each stress position of the original words. Note that in Groups 1 and 3, the majority of items are stressed according to the corresponding original words whereas in Group 2 there is a preponderance of penultimate stress assignment. Table 3 shows the values of the lexical and penultimate indices for each condition. Every value listed in Table 3 differs significantly from 1.0 in *t*-test by both subjects and items analysis, whereas only the indicated cases differ significantly from 0.0.

Table 2. Proportion of times in which stress was assigned on each syllable in Experiment 1, grouped by original word stress position (in syllables, counting from the end), separately for each pseudo-word group.

Group	Word stress	Stress assignment		
		1	2	3
1	1	.79	.15	.05
	2	.02	.92	.05
	3	.00	.12	.88
2	1	.20	.62	.16
	2	.03	.76	.20
	3	.08	.47	.45
3	1	.63	.25	.11
	2	.01	.91	.08
	3	.01	.23	.76

LexI differs significantly between Groups 1 and 2, $t_1(36) = 16.89$, $p < .0005$; $t_2(27) = 4.46$, $p < .0005$, consistent with the hypothesis that only non-words resembling words would be assigned the same stress pattern as the words. PenI differs significantly between Groups 1 and 2 only by participants, $t_1(36) = 8.31$, $p < .0005$; $t_2(27) = 1.15$, $p = .260$, providing only partial support to the hypothesis that non-words not resembling words would be disproportionately assigned penultimate stress. Both indices differ significantly only by participants between Groups 1 and 3, LexI: $t_1(36) = 5.53$, $p < .0005$; $t_2(28) = 1.79$, $p = .085$; PenI: $t_1(36) = 3.73$, $p = .001$; $t_2(28) < 1$, thus failing to support the hypothesis that the frequency of the original word would affect the influence of its stress pattern on the reading of the non-word.

However, the inclusion of items based on words stressed on the penultimate syllable obscures the distinction between the indices, because such items are expected to be stressed on the penultimate both lexically and by the hypothesised default. For this reason, Table 3 also shows the indices calculated over the subset of items based on words not stressed on the penultimate. LexI is hardly affected in Groups 1 and 3 by the exclusion of penultimate-based items, whereas PenI is rendered negative, indicating that items in these conditions were not stressed on the penultimate unless the original word carried penultimate stress.

Table 3. Average (by participant) lexical (LexI) and penultimate (PenI) indices for Experiment 1.

Group	All items		Excluding penultimate	
	LexI	PenI	LexI	PenI
1	.80	.10[1]	.76	− .30
2	.21	.43	− .01[1]	.32[1]
3	.65	.20[1]	.54	− .14[1]

[1]Statistically indistinguishable from .00 by t-test.
Note: Left columns: including responses to all items; Right columns: excluding responses to items based on words stressed on the penultimate.

Comparing these modified indices between Groups 1 and 2, and despite the smaller number of items included, significant differences are found for both (LexI: $t_1(36) = 16.81$, $p < .0005$; $t_2(16) = 4.82$, $p < .0005$; PenI: $t_1(36) = 11.48$, $p < .0005$; $t_2(16) = 3.77$, $p = .002$), consistent with the hypothesis that resemblance to a word causes the lexical stress pattern to be applied. In the comparison between Groups 1 and 3, however, the difference by items remains non-significant (LexI: $t_1(36) = 5.50$, $p < .0005$; $t_2(17) = 1.75$, $p = .099$); PenI: $t_1(36) = 4.42$, $p < .0005$; $t_2(17) = 1.63$, $p = .121$, indicating either that the frequency manipulation is not robust over items or that more items are needed if the trends apparent in Table 3 reflect real effects.

The robust difference in LexI found between Groups 1 and 2 indicates that lexical knowledge is used, when available, to affect stress assignment in reading non-words. In Group 2, when the potential confound with items derived from penultimate-stress words is removed, LexI becomes indistinguishable from 0.0, confirming the absence of lexical information. However, we cannot conclude that lexical information can be the *sole* source of stress assignment information in reading non-words without a diacritic, even when the non-words obviously and strongly resemble particular frequent words, because LexI was always statistically distinguishable from 1.00.

In the absence of lexical information, the hypothesised default metrical pattern of penultimate stress seems to apply in most cases, as shown by the positive PenI. However, PenI was also statistically distinguishable from 1.00 in every case, including in Group 2 with the items not resembling any particular words. One possible interpretation is that other factors, not related to the default pattern, may take over for certain items in the absence of lexical information. Therefore, despite clear evidence in favour of the default stress pattern for Greek, it remains unclear exactly when this pattern is applied and what other factors might be contributing.

Experiment 2

In Experiment 1, stimuli were presented without diacritics, even though standard Greek spelling always includes the diacritic. Even though no participant complained about the lack of diacritics, and in fact writing without diacritics is not uncommon among school-age children, it remains possible that an unnatural way of processing the pseudo-words may have given rise to unnatural stress assignment. Therefore, in this experiment, pseudo-words were presented with stress diacritics. We concentrated on testing the default pattern against the diacritic, using pseudo-words constructed to resemble no real words, thus diminishing lexical influences as much as possible. A new set of pseudo-words were created, with higher phonological complexity, to ensure that reading errors would occur.

There was only one item group in this experiment, presented with stress diacritics, for the children to read aloud. The measure of interest was the distribution of stress assignment position as a function of the position of the diacritic. Any deviation from the position indicated by the diacritic, considered a 'stress error', would indicate insufficient reliance on, or imperfect processing of, the diacritic. Moreover, a consistent tendency for stress errors to occur towards the penultimate syllable would constitute evidence for the operation of a default penultimate pattern even in the presence of reliable orthographic information from the diacritic.

Method

Participants

Thirty-three children from grades 7 to 9 participated, including 11 boys and 22 girls, 12–15 years old (mean age 13.2, *SD* 0.8 years). All but three had also participated in Experiment 1. Participation was voluntary and no compensation was provided.

Stimuli

Seventy-two real words three to five syllables long, with one to three consonant clusters, formed the initial set. Written frequencies of these items, derived from HNC, ranged between 0.1 and 741 per million. Each word was turned into a non-word by changing several phonemes so that the result was unrecognisable yet phonotactically acceptable. Phonetic change scores, calculated as for the items used in Experiment 1, ranged between 6 and 27.

The resulting non-words were submitted to a pre-test, in which 10 university students were presented with the printed list and asked to write down, for each item, a word that comes to mind. Non-words were discarded if any words at all were produced in response to them. Thus, there remained 46 items (listed in the Appendix, Table A2), whose average characteristics are listed in Table 1 as Group 4. These items were printed on a single sheet of A4 paper in two columns, using Times Roman 14pt typeface, in randomised order (the same for all participants), with stress diacritic in the same position as in the corresponding original words.

Procedure

The procedure was identical to that of Experiment 1.

Results and discussion

Segmental accuracy

Participants read the 46 non-words on average in 130 seconds (*SD* 54 seconds, range 71–284 seconds). Each response was rated for phonetic accuracy, receiving one point for each incorrect phoneme produced. In one-way analyses of segmental errors, there was a significant effect of number of syllables, $F_1(2, 64) = 8.610$, $p < .0005$; $F_2(2, 43) = 5.845$, $p = .006$, with a corresponding significant linear trend, $F_1(1, 32) = 10.487$, $p = .003$; $F_2(1, 43) = 9.763$, $p = .003$, and no quadratic trend. The main effect of number of consonant clusters was significant by subjects only, $F_1(2, 64) = 3.963$, $p = .024$; $F_2(2, 43) = 2.114$, $p = .133$, but the corresponding linear trend was significant in both analyses, $F_1(1, 32) = 5.894$, $p = .021$; $F_2(1, 43) = 4.225$, $p = .046$, in the absence of a quadratic trend. There was no significant effect of stress position on segmental errors, $F_1(2, 64) = 1.535$, $p = .223$; $F_2(2, 43) = 1.630$, $p = .208$. Thus, segmental errors were increased with higher phonological complexity (more syllables or more clusters), as intended, but they did not depend on stress position.

Table 4. Proportion of times in which stress was assigned on each syllable in Experiment 2, grouped by position of the printed stress diacritic (in syllables, counting from the end).

Diacritic position	Stress assignment		
	1	2	3
1	.754	.184	.063
2	.027	.896	.078
3	.006	.172	.822

Stress assignment

Responses in which a pseudo-word was produced with stress assigned on a syllable different from the one indicated by the stress diacritic were considered stress errors. In one-way analyses of variance of stress errors, there was no significant main effect of number of syllables, $F_1(2, 64) = 1.115$, $p = .334$; $F_2(2, 43) < 1$, or number of consonant clusters ($F_1, F_2 < 1$), indicating that phonological complexity did not affect stress errors. In contrast, there was a significant effect of stress position on stress errors, $F_1(2, 64) = 9.021$, $p < .0005$; $F_2(2, 43) = 7.771$, $p = .001$. The linear trend of stress position narrowly missed statistical significance, $F_1(1, 32) = 4.155$, $p = .050$; $F_2(1, 43) = 3.668$, $p = .062$, but the quadratic trend was significant, $F_1(1, 32) = 14.489$, $p = .001$; $F_2(1, 43) = 11.874$, $p = .001$, supporting the interpretation that significantly fewer errors were made on items stressed on the penultimate than on items stressed on other syllables.

Table 4 shows the distribution of stress assignment responses for each position of the stress diacritic. Off-diagonal proportions are stress errors. Their proportion, about 10–25%, is comparable with the performance reported in Protopapas (2006) for a different pseudo-word list with similar properties, and consistent with the interpretation that the printed stress diacritic alone is insufficient to determine stress (necessitating lexical or other information for correct word reading).

In 3×3 analysis of stress error proportions by stress diacritic position and (incorrect) stress assignment position, there was a significant main effect of both diacritic position, $F_1(2, 64) = 9.021$, $p < .0005$; $F_2(2, 43) = 7.771$, $p = .001$, and stress assignment position, $F_1(2, 64) = 20.749$, $p < .0005$; $F_2(2, 86) = 34.690$, $p < .0005$, as well as a significant interaction between the two, $F_1(4, 128) = 19.903$, $p < .0005$; $F_2(4, 86) = 17.475$, $p < .0005$. The interaction is obviously expected, as different diacritic positions offer different opportunities for stress errors (e.g. with the diacritic on the final syllable, it is not possible to make an error on the final). The main effect of stress error position, however, indicates that the overall proportion of errors was not uniform over the three syllables.

As Table 4 shows, error rates towards the penultimate, for diacritic positions on the antepenultimate or final, are around 17–18%, whereas error rates towards the final and antepentultimate are all less than 8%. In direct comparisons of error proportions by paired-samples t-tests, separately for each diacritic position, highly significant differences were found between penultimate and final (for diacritic on the antepenultimate: $t_1(32) = 4.949$, $p < .0005$; $t_2(14) = 6.081$, $p = .005$) and between penultimate and antepenultimate (for diacritic on the final: $t_1(32) = 4.128$, $p < .0005$; $t_2(14) = 3.336$, $p = .005$) but only marginally by items between antepenultimate and final (for diacritic on the penultimate: $t_1(32) = 3.203$, $p = .003$; $t_2(14) = 2.033$, $p = .060$). The higher proportion of stress errors towards the penultimate is consistent with the hypothesis that the penultimate is the default syllable for stress assignment.

General discussion

In a stress-assigning language such as Greek, the explicit assignment of a stress pattern to a word is a necessary step in reading the word aloud, and perhaps in silent reading as well. We have examined the potential of lexical information and of a default metrical pattern to affect stress assignment, using non-word stimuli with varying resemblance to words (to control lexical activation) and manipulating the presence of the written diacritic (to control orthographic information).

Resemblance of the pseudo-word stimuli to words was found to affect stress assignment strongly, thus confirming the availability and use of a lexical route for stress assignment. The importance and automatic involvement of this lexical route is appreciated more clearly in its involvement with non-word stimuli, for which it should properly not apply were it not for similarity-based activation. The observed effects suggest that lexical derivation of the diacritic is an efficient, well-practised strategy, and therefore a likely primary source of stress assignment information in normal word reading. The results also confirmed the existence of a default metrical pattern stressing the penultimate syllable, seen both in fewer deviations from the orthographically marked syllable for items marked on the penultimate, as well as in more 'stress errors' towards the penultimate than towards other syllables.

Factors increasing phonological complexity affected segmental errors but not stress errors, whereas the converse was true for stress position. This pattern is consistent with independent processing of segmental and suprasegmental information. Evidence that stress assignment constitutes a process distinct from segmental assembly in reading comes from speech production studies. For example, Roelofs and Meyer (1998) have found that, in the context of a common segmental onset, syllable structure alone was not an effective prime whereas number of syllables and stress position were. More recently, Schiller et al. (2004) reported a regularity effect for Dutch, which they interpreted as being consistent with a non-lexical stress assignment strategy for the predominant stress pattern.

Reading models have all but ignored stress assignment (Black & Byng, 1986; Duncan & Seymour, 2003), typically dealing only with monosyllabic words, with the exception of a recent extension by Rastle and Coltheart (2000) of the DRC model, an implementation of the dual-route theory. A dual-route strategy seems appropriate in accounting for a general tension between lexical and non-lexical sources of stress assignment information, as it accounts for the tension between addressed and assembled phonology in the segmental specification. This extension sought to incorporate (a) stress information in the lexical specification, and (b) a set of rules accounting for the phonological constraints on stress assignment that apply in the English language. Much effort was devoted to particular rules affecting stress assignment non-lexically, further complicated by the requirement to handle vowel reduction, a stress-related phonological process in English.

In languages such as Italian, Spanish and Greek, however, the main issue in stress assignment would not be the specification of the structural rules, because the rules are either simple (as in Spanish and Italian) or non-existent (as in Greek). The availability of additional sources of information, such as the diacritic and inflectional morphology, and the different nature of the default metrical pattern, suggest that extension of DRC to other languages will not be straightforward. On the other hand, connectionist models of reading have not yet accounted for words with more than one syllable and thus cannot handle lexical stress (e.g. Harm & Seidenberg, 2004). Clearly, if stress assignment is an

important and necessary step in reading aloud, then cognitive models of reading must be extended to include it.

In conclusion, our experiments have provided evidence in favour of the hypothesis that lexical and non-lexical sources of information are available and actively taken into account when reading aloud. Our use of non-word stimuli, which enabled manipulation of lexical information, necessitates additional work with real words before strong conclusions for word reading can be confidently stated. Even though it is reasonable to expect that well-practised decoding and retrieval routines, based on word reading experience, would also predominate with non-word stimuli, it remains to be tested empirically whether the interplay between the various competing sources of stress assignment information is comparable for words and non-words. Future investigations should further clarify the role of each individual source of stress assignment information in reading, focusing on the most ecologically relevant task of reading words in coherent text context and connecting to other suprasegmental processes affecting phrase-level intonation.

Acknowledgements

References

Black, M. & Byng, S. (1986). Prosodic constraints on lexical access in reading. *Cognitive Neuropsychology*, 3, 369–409.

Botinis, A. (1998). Intonation in Greek. In D. Hirst & A. di Cristo (Eds.), *Intonation systems: A survey of twenty languages*. (pp. 288–310). Cambridge: Cambridge University Press.

Colombo, L. (1992). Lexical stress and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 987–1003.

Duncan, L.G. & Seymour, P.H.K. (2003). How do children read multisyllabic words? Some preliminary observations. *Journal of Research in Reading*, 26, 101–120.

Gutiérrez Palma, N. (2003). El acento léxico y su función en el reconocimiento de palabras escritas en adultos y en niños [Lexical stress and its role in written word recognition by adults and children]. *Unpublished PhD Thesis*, Universidad de Granada, Spain.

Harm, M.W. & Seidenberg, M.S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, 111, 662–720.

Hatzigeorgiu, N., Gavrilidou, M., Piperidis, S., Carayannis, G., Papakostopoulou, A., Spiliotopoulou, A. et al. (2000, 31 May–2 June). Design and implementation of the online ILSP corpus. In *Proceedings of the Second International conference of language resources and evaluation (LREC)* (Vol. 3, pp. 1737–1740). Athens, Greece.

Hirst, D. & di Cristo, A. (1998). A survey of intonation systems. In D. Hirst & A. di Cristo (Eds.), *Intonation systems: A survey of twenty languages*. (pp. 1–44). Cambridge: Cambridge University Press.

Kappa, I. (2002). On the acquisition of syllabic structure in Greek. *Journal of Greek Linguistics*, 3, 1–52.

Malikouti-Drachman, A. (2002). Greek phonology: A contemporary perspective. *Journal of Greek Linguistics*, 2, 187–243.

Malikouti-Drachman, A. & Drachman, G. (1989). Τονισμός στα ελληνικά [Stress in Greek]. *Studies in Greek Linguistics*, 9, 127–143.

Nunes, T., Roazzi, A. & Buarque, L.L. (2003). Learning to mark stress in written Portuguese. *Revue de Linguistique*, 22, 99–108.

Petrounias, E.V. (2002). Νεοελληνική γραμματική και συγκριτική ανάλυση, Τόμος Α: Φωνητική και εισαγωγή στη φωνολογία [Modern Greek grammar and comparative analysis, Vol A: Phonetics and introduction to phonology]. Thessaloniki: Ziti.

Protopapas, A. (2006). On the use and usefulness of stress diacritics in reading Greek. *Reading & Writing: An Interdisciplinary Journal*, 19(2), 171–198.

Rastle, K. & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and non-words. *Journal of Memory and Language*, 42, 342–364.

Revithiadou, A. (2004). A comparative study of Greek-English stress. Invited talk presented at the Generative approaches to contrastive linguistics (GACL1): The grammars of English and Greek in comparison. Department of English Studies, University of Cyprus, Nicosia, 26–28 November.

Roelofs, A. & Meyer, A.S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 922–939.

Schiller, N.O., Fikkert, P. & Levelt, C.C. (2004). Stress priming in picture naming: An SOA study. *Brain and Language*, 90, 231–240.

Address for correspondence: Athanassios Protopapas, Institute for Language & Speech Processing, Artemidos 6 & Epidavrou, GR-15125 Maroussi, Greece. E-mail: *protopap@ilsp.gr*

Appendix: Lists of stimuli

Table A1. Properties of the non-words (and source words) used in Experiment 1.

Group	Non-word	Pronunciation	Original word	Freq-uency	Phonetic difference	Stress position	Number of syllables	Consonant clusters
1	αποκελεσμα	apoˈcɛlɛzma	αποτέλεσμα	385.5	1	3	5	1
1	γιμονται	ˈjimodɛ	γίνονται	166.7	1	3	3	0
1	δηραδη	ðiraˈði	δηλαδή	607.1	1	1	3	0
1	εξωδεριχων	ɛksoðɛriˈkon	εξωτερικών	288.5	2	1	5	1
1	επιφειρησεις	ɛpifiˈrisis	επιχειρήσεις	204.1	1	2	5	0
1	θεωρειδαι	θɛoˈriðɛ	θεωρείται	139.0	2	2	4	0
1	χρογραμμα	ˈkroɣrama	πρόγραμμα	266.5	1	3	3	2
1	περιοχη	pɛrioˈci	περιοχή	353.2	1	1	4	0
1	πεχιπου	pɛˈçipu	περίπου	402.3	3	2	3	0
1	προχειλενου	prociˈlɛnu	προχειμένου	237.2	2	2	4	1
1	συγκεχρινενα	sijɛkriˈnɛna	συγκεχριμένα	167.3	1	2	5	1
1	τερευταια	tɛrɛˈftɛa	τελευταία	412.4	1	2	4	0
1	φερισσοτερο	fɛriˈsotɛro	περισσότερο	380.5	2	3	5	0
1	φρισχεται	ˈfriscɛtɛ	βρίσχεται	324.5	1	3	3	2
1	χρειαλεται	xriˈalɛtɛ	χρειάζεται	183.8	1	3	4	1
2	απαχεροται	apaˈcɛrotɛ	αναφέρεται	133.2	7	3	5	0
2	εποτεγετζα	ɛpoˈtɛjeʣa	αποτέλεσμα	385.5	8	3	5	0
2	εσαρεθο	ɛˈsarɛθo	ανάμεσα	268.9	9	3	4	0
2	ετζωτηραχων	ɛʣotiraˈkon	εξωτερικών	288.5	9	1	5	0
2	θρογλαμετα	θroˈɣlamɛta	προβλήματα	374.4	7	3	4	2
2	νερισου	nɛˈrisu	περίπου	402.3	5	2	3	0
2	νυγχαχρομινα	nigakroˈmina	συγκεχριμένα	167.3	8	2	5	1
2	πρειοραται	priˈoratɛ	χρειάζεται	183.8	7	3	4	1
2	πρινειρονου	priniˈronu	προχειμένου	237.2	9	2	4	1
2	σατσαυω	saˈtsavo	πιστεύω	148.2	10	2	3	0
2	σλεταχα	zlɛtaˈxa	σχετιχά	244.2	6	1	3	1
2	τηλαυταχα	tilaˈftaka	τελευταία	412.4	9	2	4	1
2	φερασοτιχο	fɛraˈsotiko	περισσότερο	380.5	9	3	5	0
2	χουλεζιστο	xuˈlɛzisto	τουλάχιστον	288.5	9	3	4	1
3	βυλλωσια	viloˈsça	φυλλωσιά	0.2	1	1	3	1
3	ζυχαρια	zikaˈrja	ζυγαριά	3.8	2	1	3	1
3	χαραθι	kaˈraθi	καλάθι	18.2	1	2	3	0
3	χλωμιζεται	kloˈmizɛtɛ	κλονίζεται	1.7	1	3	4	1
3	χουφαδαχι	kufaˈðaki	χουβαδάχι	0.1	1	2	4	0
3	λαιμαλγα	ˈlɛmalɣa	λαίμαργα	0.6	1	3	3	1
3	οιχοδομημα	igoˈðomima	οικοδόμημα	6.4	1	3	5	0
3	ορημερις	orimɛˈris	ολημερίς	0.6	1	1	4	0
3	προαγγερος	proˈajeros	προάγγελος	1.4	1	3	4	1
3	προτιαγραφες	protiaɣraˈfes	προδιαγραφές	24.5	2	1	5	2
3	στοχαστρο	ˈstokastro	στόχαστρο	8.9	1	3	3	2
3	συγκυφερνητη	sijifɛrˈniti	συγκυβερνήτη	0.4	1	2	5	1
3	συνοιγισμος	sinijiˈzmos	συνοιχισμός	0.4	2	1	4	1
3	τενετουργια	tɛnɛturˈjia	τελετουργία	1.3	1	2	5	1
3	υιοδεσια	ioðɛˈsia	υιοθεσία	1.3	1	2	5	0

Notes: Stimuli and original words in Greek orthography, pronunciation in the International Phonetic Alphabet. Non-word pronunciation includes stress assignment marking (stress mark precedes stressed syllable), by analogy to the original words; orthographic presentation of the non-word stimuli omits stress diacritics, consistent with presentation in Experiment 1. Frequency refers to occurrences (per million words) of the original word in the Hellenic National Corpus. Stress position refers to the original word, in syllables, counting from the end (1 = final).

Table A2. Properties of the non-words (and source words) used in Experiment 2.

Non-word	Pronunciation	Original word	Freq-uency	Phonetic difference	Stress position	Number of syllables	Consonant clusters
ετραδόχνια	ɛtraˈðoknia	οικογένεια	67.4	14	3	5	2
οκωπτισά	okoptiˈsa	ερωτικά	6.3	10	1	4	1
σπαλεστοχή	spalɛstoˈci	πολιτική	741.1	12	1	4	2
δεαρδαμπά	ðɛarðaˈba	διαδρομή	34.7	10	1	4	1
φτιψάχου	ftiˈpsaxu	κειμένου	30.5	14	2	3	2
νιοψαρτασμάρος	ɲopsartaˈzmaros	διαμερισμάτων	3.4	23	2	5	3
πραχλεσταρτιχές	praklɛstartiˈcɛs	κληρονομιχές	0.9	23	1	5	4
λαχεχαρμίο	lacɛxarˈmio	πολυφωνία	2.0	19	2	5	1
ξιφταμοχερτό	ksiftamoçɛrˈto	μετωνυμική	0.0	27	1	5	3
φαχταψοταρί	faxtapsotaˈri	μεσολαβητή	3.6	25	1	5	2
εξόρεσλη	ɛˈksorɛsli	εξέγερση	14.8	7	3	4	2
μοσοχλοτής	mosoxloˈtis	μεταβλητός	0.3	11	1	4	1
κεσμαχρηντός	cezmakriˈdos	χαμαρωτός	0.3	12	1	4	2
βαρνεσχίχτορο	varnɛˈscixtoro	γενικότερα	49.5	21	3	5	3
σοστούχα	soˈstuxa	συστάδα	0.3	6	2	3	1
ναστεόλμιχνα	nastɛˈolmixna	διχαιώματα	97.9	18	3	5	3
σεναξιτροδόν	sɛnaksitroˈðon	συντηρητιχών	6.9	18	1	5	2
ορχελιχτάστε	orcɛliˈktastɛ	επιχαλούνται	7.6	24	2	5	3
δενοτροτεχό	ðɛnotrotɛˈko	δημοχρατιχά	6.3	10	1	5	1
πλαρυμπόναν	plariˈbonan	περιμένουν	47.8	10	2	4	1
σμονεχρόα	zmonɛˈxroa	μοναρχία	2.1	10	2	4	2
ηντοχλωχή	idokloˈçi	αναχωχή	3.4	8	1	4	1
απαστεβέχι	apastɛˈvɛci	εμπιστοσύνη	40.3	14	2	5	1
χεφταψίνετο	çeftaˈpsinɛto	νομιμότητα	12.6	26	3	5	2
ρεζαπόφεσμα	rɛzaˈpofɛzma	δημοψήφισμα	15.8	15	3	5	1
πλησμόσταο	pliˈzmostao	δημόσιο	138.0	13	3	4	3
αξόβλοξη	aˈksovloksi	εξέλιξη	156.1	9	3	4	3
οντέπλοτα	oˈdɛplota	ενότητα	39.1	11	3	4	1
πτορχανού	ptorxaˈnu	ψυχιχού	5.2	11	1	3	2
εχταστεήφρας	ɛktastɛˈifas	υποχρεώσεις	47.3	16	2	5	2
ανησλάπεις	aniˈslapis	απαντήσεις	45.5	13	2	4	1
οχέπλη	oˈcɛpli	επάνω	36.6	11	2	3	1
πραχτεοφρίσο	praktɛoˈfriso	περιορίζω	0.2	16	2	5	3
σπόλωρας	ˈspoloras	τέσσερις	137.9	12	3	3	1
λοσόσθαση	loˈsosθasi	μετάθεση	5.7	12	3	4	1
μετόσφαλη	mɛˈtosfali	μετάφραση	30.4	9	3	4	1
σχονετάπανε	sxonɛˈtapanɛ	σχηματίζονται	4.5	13	3	5	1
οποσταγάς	opostaˈɣas	επιταγής	9.0	9	1	4	1
πταρασβότος	ptaraˈzvotos	παραβάτης	0.6	10	2	4	2
χλαξεδές	klaksɛˈðɛs	χρονιχές	4.8	14	1	3	2
φεχτονίξοσι	fɛktoˈniksosi	χατανάλωση	30.1	16	3	5	2
ψέρτανο	ˈpsɛrtano	χίνδυνο	69.5	11	3	3	2
επασλόβουν	ɛpaˈslovun	αναλάβουν	21.9	10	2	4	1
ταπλαχή	taplaˈci	τελιχά	290.6	10	2	3	1
επαπλιψός	ɛpapliˈpsos	εποπτιχός	0.1	8	1	4	2
προχλίο	proˈxlio	βιβλίο	136.4	10	2	3	2

Notes: Stimuli and original words in Greek orthography, pronunciation in the International Phonetic Alphabet. Non-word pronunciation includes stress assignment marking (stress mark precedes stressed syllable), by analogy to the original words; orthographic presentation of the non-word stimuli includes stress diacritics, consistent with presentation in Experiment 2. Frequency refers to occurrences (per million words) of the original word in the Hellenic National Corpus. Stress position refers to the original word, in syllables, counting from the end (1 = final).