Stress Typicality Effects in Native and Non-native Speakers of English

Joanne Arciuli and Linda Cupples
Speech, Hearing and Language Research Centre, Macquarie University, Australia
E-mail: joanne@holly.com.au, linda.cupples@mq.edu.au

ABSTRACT

This study investigated the processing of lexical stress in native and non-native speakers of English. Specifically, we examined stress typicality effects (where typicality is defined on the basis of grammatical category) in disyllabic words using two on-line tasks. Our grammatical classification experiment showed an overall effect of stress typicality in non-native speakers but no overall effect in native speakers (although some native speakers with high error rates did show a significant effect). Our onset gating experiment revealed significant stress typicality effects in both groups of speakers. These findings are discussed with reference to previous research.

INTRODUCTION

In English, most disyllabic words exhibit trochaic stress; however, typical patterns of lexical stress vary across disyllabic nouns and verbs. While most disyllabic nouns exhibit trochaic stress (e.g., ‘patron’) most disyllabic verbs exhibit iambic stress (e.g., ‘prevent’) [5,6]. Further evidence comes from homographs, like ‘incense’ and ‘refuse’ (trochaic stress is typically assigned to the nominal meaning, and iambic stress to the verbal meaning) [7]. Based on these statistical occurrences in English, trochaic nouns and iambic verbs can be considered to be ‘typical’ while iambic nouns and trochaic verbs can be considered ‘atypical’.

Many early studies that examined the processing of lexical stress (not stress typicality *per* se) failed to find significant effects of lexical stress in native speakers of English [8,9,10,11]. Interestingly, a direct investigation of stress typicality also reported nonsignificant effects during spoken word recognition [12]. In a grammatical judgement task native speakers of English heard items in the context of “to…” or “the…” and had to decide whether what they heard was grammatical or not. The results showed no effect of stress typicality (i.e., “the apple” was no easier than “the cigar” and “to await” was no easier than “to borrow”).

In other research, a different grammatical classification task was used to investigate effects of stress typicality [4]. In this task native and non-native speakers were presented with individual items and were asked to decide whether each one was a noun or a verb. The authors found no effect of stress typicality in native speakers but did find a significant effect in non-native speakers. They suggested that *task difficulty* might play a role in the processing of lexical stress. Indeed, the authors showed that their non-native speakers exhibited a higher overall error rate than their native speakers. However, they did not carry out any statistical analyses to demonstrate that error rate *per se* was related to the size of the stress typicality effect.

A more recent study has shown that *both* native and non-native speakers of English utilise lexical stress information in identifying spoken words [1]. During cross-modal priming and forced-choice identification experiments both groups of speakers showed effects of lexical stress (although the study did not examine stress typicality *per se*). In attempting to explain this apparent conflict with previous research, the authors noted that previous studies might not have found significant effects of lexical stress in native speakers because they included *less sensitive tasks*.

Interestingly, the processing of lexical stress in native speakers has also been examined using the onset-gating paradigm [2,3]. The same 12 disyllabic and 18 trisyllabic items were included in three conditions (onset only, onset plus duration and onset plus duration plus prosody). While the results showed no difference in isolation times (a measure of speed of recognition) between the first two conditions, there was a significant difference between these conditions and the third condition. The authors concluded that the provision of prosodic information (in the absence of segmental information) can speed recognition during onset-gating.

In view of these mixed findings we set out to provide further empirical data on the effects of stress typicality in native and non-native speakers of English. Specifically, in Experiment 1, we wanted to replicate the findings of a previous study that had used grammatical classification [4] and to look more closely at the role of task difficulty (as indicated by overall error rate) in the processing of lexical stress by native speakers. In particular, we wondered whether some native speakers (those with high error rates) might show significant effects of stress typicality during grammatical classification. In Experiment 2, we examined stress typicality effects using the onset-gating paradigm. As stress effects had been found previously using this task [2,3], we hypothesised that onset-gating might be a more sensitive paradigm that could detect typicality effects in both native and non-native speakers.

 ISBN 1-876346-48-5

EXPERIMENT 1

In this experiment we used the same grammatical classification task used in a previous study [4] with a different set of items.

Method

Participants. A total of 40 students took part in the experiment (23 native speakers and 17 non-native speakers).

Materials. A set of 120 unambiguous disyllabic nouns and verbs was selected for the experiment. Of the 60 nouns included, 30 had typical trochaic stress (e.g., 'wizard') and 30 had atypical iambic stress (e.g., 'terrain'). Of the 60 verbs included, 30 had typical iambic stress (e.g., 'invent') and 30 had atypical trochaic stress (e.g., 'punish'). Typically and atypically stressed words were closely matched in terms of base frequency (67.5 per million vs. 67 per million respectively; $F < 1$). Stimuli were recorded using a female speaker and resulting files were normalized and presented in 44.1KHz 16bit DAT format. Tokens were edited and trimmed where sound levels were below 40dB below average SPL, with a linear fade from t=0 to t=50 msecs before and after the utterance. There was no significant difference in the average duration of typically and atypically stressed items (787 msecs and 802 msecs respectively; $F < 1$).

Procedure. The experiment was run using DMDX software. Participants were instructed that they would hear individual words and would be prompted for a response to each word by a message on the computer screen. Approximately half of the participants were instructed to press the left mouse button if the word was a noun and the right mouse button if the word was a verb (presentation of each wav was immediately followed by the phrase "Noun or Verb?" for these participants). The other half were instructed to press the right mouse button if the word was a noun and the left mouse button if the word was a verb (these participants were presented with the message "Verb or noun?").

There was a practice set of 16 items followed by the experimental set (presented in a uniquely randomised order for each participant). Inter-stimulus delay was 1000 msecs (after a response) or 5000 msecs (if no response was provided).

Results

In Experiments 1a and 1b we used repeated measures in our subject analyses and between items measures in our item analyses.

Experiment 1a – Native Speakers

Response Times. Means are provided in Table 1a. While there was a trend indicating a typicality advantage, the data showed no significant difference between typically stressed words (trochaic nouns and iambic verbs) and atypically stressed words (iambic nouns and trochaic verbs) (both $Fs < 1$). There were no other significant effects (all $Fs < 1$).

	Nouns	Verbs	Overall
Typical Stress	758 (9.0)	740 (8.1)	749 (8.6)
Atypical Stress	769 (12.3)	760 (10.3)	765 (11.3)

Table 1a. Mean Grammatical Classification Times and Percent Errors (in parentheses) in Experiment 1a (N = 23).

Error Rates. The means provided in Table 1a show that participants were more accurate in classifying typically stressed words than atypically stressed words, a difference that was significant in the subject analysis ($F1\ (1,22) = 9.13$, $p < .01$), but did not generalise to items ($F2\ (1,116) = 1.92$, $p > .10$). There were no other significant effects (all $Fs < 1$).

A correlational analysis indicated a relationship ($r\ (23) = .50$, $p < .05$) between overall error rate and the size of the stress typicality effect (computed by subtracting the mean error rate for typically stressed words from that for atypically stressed words). When we looked at the data from the 11 participants who had exhibited overall error rates above the median (above 10%) we found a significant main effect of stress typicality ($F1\ (1,10) = 16.38$, $p < .005$; $F2\ (1,116) = 3.19$, $p = .08$). Typically stressed words elicited fewer errors than atypically stressed words.

Experiment 1b – Non-native Speakers

Response Times. As can be seen in Table 1b, there was a trend for typically stressed words to elicit faster reaction times than atypically stressed words in the group of non-native speakers, but this difference was not significant (both $Fs < 1$). There were no other effects that were significant by subjects and items.

	Nouns	Verbs	Overall
Typical Stress	862 (16.7)	878 (17.5)	870 (17.1)
Atypical Stress	910 (24.3)	869 (24.3)	890 (24.3)

Table 1b. Mean Grammatical Classification Times and Percent Errors (in parentheses) in Experiment 1b (N = 17).

Error Rates. The data in Table 1b show that non-native speakers were more accurate in classifying typically stressed words than atypically stressed words, a difference that was significant by subjects and items ($F1\ (1,16) =$

22.03, $p < .01$; F2 (1,116) = 5.4, $p < .05$). There were no other effects that were significant by subjects and items.

EXPERIMENT 2

In this experiment we used onset-gating to investigate stress typicality effects. In line with previous research we also manipulated presentation condition. However, we included two rather than three conditions: onset only and onset plus duration plus prosody.

Method

Participants. A total of 70 students took part in the experiment (34 native and 36 non-native speakers).

Materials. Experimental items consisted of 40 disyllabic words: 20 nouns (10 with typical trochaic stress and 10 with atypical iambic stress) and 20 verbs (10 with typical iambic stress and 10 with atypical trochaic stress). Typically and atypically stressed words were matched on initial phoneme (all consonants) and did not differ significantly in terms of base frequency ($F < 1$). Typically stressed words had an average of 3.8 phonological neighbours while atypically stressed words had an average of 4.3 neighbours ($F < 1$). There was also no significant difference between the average spoken frequency of neighbours (CELEX database) across typically and atypically stressed words (F (1,33) = 3.41, $p > .05$). Our analysis of uniqueness points is provided in the results.

Stimuli were recorded using a female speaker. The resulting files were normalised and presented in 44.1KHz 16bit DAT format with a linear fade from t=0 to t=50 msecs before the utterance. For each word in each condition, the gate size began at 100 msecs (because of a 50 msec linear fade before each utterance) and increased by increments of 50 msecs until the entire word was presented. For the onset only condition (silenced), onsets were presented with the remainder of the word replaced by silence. For the onset plus duration plus prosody condition (filtered), onsets were presented with a filtered version of the remainder of the word (using a low-pass filter with a cut-off value of 325Hz). This type of filtering provides prosodic information in the absence of segmental information. The average number of gates used in the presentation of typically and atypically stressed words was 12 and 11 respectively.

For each condition, the sequence of files for each word was recorded on a CD with a 6 sec pause between each file (to allow time for responding). There were two random orders of presentation for each condition (the sequence of files for each word was, of course, presented in fixed order). The commencement of the sequence of files for each new word was signalled using a male recording of the numbers 1-40.

Procedure. Each participant heard all 40 words (presented as a sequence of files with increasing onsets) in one of two random orders (R1 vs. R2) for one of two conditions (silenced vs. filtered). After the presentation of each gate, participants had to write down what they thought the word could be (they were provided with a detailed response sheet listing the number of gates they should expect for each word).

Results

Response sheets were examined to determine the gate (in msecs) at which each participant correctly identified the word and did not subsequently change his/her decision (isolation point). There were no significant effects of randomisation (R1 vs. R2) so the results reported below do not include this variable. Subject analyses included presentation condition (silenced vs. filtered) as a between subjects variable and stress typicality and grammatical status as repeated measures. Item analyses included presentation condition as a repeated measure, stress typicality and grammatical status as between items variables and number of gates as a covariate. Where participants failed to identify a word correctly (after presentation of the final gate) we replaced these missing values (7.5% of the data) with the value of the final gate.

Experiment 2a – Native Speakers

Isolation Points. Mean isolation points (and standard deviations) for native speakers are presented in Table 2a.

	Presentation Condition			
	Silenced		Filtered	
	Stress Typicality			
	Typical	Atypical	Typical	Atypical
Nouns	309.7	337.8	307.2	345.6
	(6.0)	(7.9)	(5.7)	(7.4)
Verbs	307.2	360.9	333.6	362.5
	(6.7)	(5.6)	(6.3)	(5.2)
Overall	308.4	349.4	320.4	354.0
	(5.4)	(4.8)	(5.1)	(4.5)

Table 2a. Mean Isolation Points and Standard Errors (in parentheses) in Experiment 2a (N = 34).

As expected, our analysis showed a main effect of stress typicality (F1 (1,32) = 136.78, $p < .001$; F2 (1, 35) = 8.12, $p < .01$). Typically stressed trochaic nouns and iambic verbs were recognized faster than atypically stressed iambic nouns and trochaic verbs. There were no other effects that were significant by both subjects and items.

Experiment 2b – Non-native Speakers

Isolation Points. Mean isolation points (and standard deviations) for non-native speakers are presented in Table 2b. Like the native speakers, results from our non-native speakers demonstrated a significant main effect of stress typicality (F1 (1, 34) = 96.93, $p < .01$; F2 (1,35) = 10.68, $p < .005$). Again, typically stressed words were recognized faster than atypically stressed words. There were no other effects that were significant by both subjects and items.

 ISBN 1-876346-48-5

	Presentation Condition			
	Silenced		Filtered	
	Stress Typicality			
	Typical	Atypical	Typical	Atypical
Nouns	353.2	414.7	371.8	429.1
	(13.7)	(16.2)	(14.4)	(17.1)
Verbs	351.3	397.1	382.9	401.5
	(13.9)	(10.5)	(14.7)	(11.1)
Overall	352.2	405.9	377.4	415.3
	(12.4)	(12.1)	(13.1)	(12.8)

Table 2b. Mean Isolation Points and Standard Errors (in parentheses) in Experiment 2b (N = 36).

Uniqueness Points. Using the CELEX database we restricted our search to disyllabic words with the same stress pattern as the target item. We determined the uniqueness point for each item by finding the phoneme that provided the point of separation from other words and dividing this value by the total number of phonemes in the word. So, for "success", the uniqueness value was .83 (the fifth phoneme provided separation from other words such as "succinct" so we divided 5 by 6). There was no significant difference in the average uniqueness values of typically and atypically stressed words (.83 and .85 respectively; $F < 1$).

DISCUSSION

The results of Experiment 1a and 1b provide a replication of previous research [4]. While native speakers of English did not show an overall effect of stress typicality during grammatical classification, non-native speakers did. In Experiment 1a we provided statistical evidence (via correlational analysis) to support the notion that error rates were related to the size of the stress typicality effect *within* our group of native speakers. Further, we demonstrated that a subgroup of native speakers (those with high error rates) did show significant effects of stress typicality.

The results of Experiment 2b provide further evidence of a stress typicality effect in non-native speakers. Interestingly, and in contrast with previous research, Experiment 2a provides evidence of an overall stress typicality effect in native speakers as well. In this sense, it seems that the onset-gating task is more 'sensitive' than other tasks. An unexpected finding from Experiment 2 was that participants did not recognize words more quickly when onsets were presented with prosodic information (filtered condition) than when onsets were presented alone (silenced condition). This finding is in contrast with previous studies [2,3]. While there are substantial methodological differences between this study and previous studies (our study was on-line, participants made written responses and we used a larger set of only disyllabic items), further research is needed to clarify this issue.

REFERENCES

[1] Cooper, N., Cutler, A., Wales, R. "Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners", *Language and Speech*, in press.

[2] Wingfield, A., Lindfield, K., Goodglass, H. "Effects of age and hearing sensitivity on the use of prosodic information in spoken word recognition", *Journal of Speech, Language, and Hearing Research, 43,* 915-925, 2000.

[3] Lindfield, R., Wingfield, A., Goodglass, H. "The contribution of prosody to spoken word recognition", *Applied Psycholinguistics, 20*, 397-407, 1999.

[4] Davis, S. M., Kelly, M. H. "Knowledge of the English noun-verb stress difference by native and nonnative speakers", *Journal of Memory and Language, 36,* 445-460, 1997.

[5] Sereno, J. A. "Stress pattern differentiation of form class in English", *Journal of the Acoustical Society of America, 79,* S36, 1986.

[6] Kelly, M. H., Bock, J. K. "Stress in time", *Journal of Experimental Psychology: Human Perception and Performance, 14,* 389-403, 1988.

[7] Sherman, D. "Noun-verb stress alternation: An example of lexical diffusion of sound change in English", *Linguistics, 159,* 43-71, 1975.

[8] Slowiacek, L. "Effects of lexical stress in auditory word recognition", *Language and Speech, 33,* 47-68, 1990.

[9] Small, L., Simon, S., Goldberg, J. "Lexical stress and lexical access: Homographs versus nonhomographs", *Perception and Psychophysics, 44*, 272-280, 1988.

[10] Cutler, A. "*Forbear* is a homophone: Lexical prosody does not constrain lexical access", *Language and Speech, 29*, 201-220, 1986.

[11] Bond, Z., Small, L. "Voicing, vowel and stress mispronunciations in continuous speech", *Perception and Psychophysics, 34*, 470-474, 1983.

[12] Cutler, A., Clifton, C. "The use of prosodic information in word recognition", In H. Bouma, & D.G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes*. (pp.183-196). Hillsdale: Erlbaum, 1984.