Metrical rhythm in speech planning: priming or predictability

Jason A. Shaw

MARCS Institute, University of Western Sydney

J.Shaw@uws.edu.au

Abstract

We investigated how repeating metrical templates, i.e., stress priming, and the contextual predictability of stress affect behavior in a word naming task. We measured naming latencies and the phonetic duration of words produced under four different conditions (congruent stress, low predictability; congruent stress, high predictability; incongruent stress, low predictability; incongruent stress, high predictability). Results showed strong effects of metrical predictability on both naming latency and word duration but no effect of stress congruency on either variable. The relevance of these results to theories of speech production and phonological representation are discussed.

Index Terms: stress priming, metrical rhythm, timing, speech production

1. Introduction

English is rhythmically organized to avoid consecutively stressed syllables, i.e., stress class [1]. Stress clash avoidance shows up in speech corpora as well as in experimental tasks [2]. For example, English speakers are more likely to assign stress to the second syllable of a disyllabic nonce word when it is preceded in a sentence by a stressed syllable than when it is preceded by an unstressed syllable [2]. This behavior suggests an online computation of stress that is sensitive to, at least, local metrical context. Whether stress assignment is sensitive to more global metrical patterns has been largely unexplored.

The requirement that syllables are parsed into feet constitutes an efficient mechanism of stress clash avoidance that is strictly local. Psycholinguistic evidence for metrical templates, such as disyllabic iambs and trochees, is available in the form of speech errors patterns, which tend to preserve metrical position, and experimental paradigms that have demonstrated stress priming, most notably the "implicit form priming" paradigm [3]. In this paradigm word pairs are memorized and one word is used to cue the production of the other. Words that share the number of syllables and the position of stress have been shown to facilitate naming latencies. Evidence such as this has been taken as support of the hypotheses that metrical templates are independent from speech segments. In the theory of [4], for example, metrical templates specifying syllable count and stress location are activated and associated to phonemes in the process of speech production.

Particularly for a language like English that has both disyllabic trochees and disyllable iambs, merely activating a metrical template in speech production will not result in a bias toward alternating stress or repeating foot structures as long as the level of activation returns to baseline after selection. If, however, the activation level of metrical templates does not return to baseline after selection, then residual activation could constitute a soft bias towards repeated metrical templates. In other words, it may be faster to prepare an iamb preceded by another iamb because the metrical template can remain active across productions. The residual activation of metrical templates could potentially account for observed stress priming effects, such as those in [5]. In that study, Italian speakers were asked to read words and nonce words one at a time from a computer screen. When target words with non-dominate stress (antepenultimate stress) were preceded by nonce words with dominate stress (penultimate), Italian speakers made a large number of stress errors and showed slower reaction times on correct trials than when targets were preceded by words with congruent stress. [5] argue that these results support a sub-lexical mechanism of stress assignment that can be primed through repeated activation.

In this study, we test this hypothesis for English speakers in a modified version of [5]'s paradigm. In addition to testing congruent and incongruent stress primes, we also manipulated the predictability of stress position in target words. The target word and the immediately preceding prime were embedded in larger metrical patterns. The metrical pattern of the words flanking the target predicted the stress position of the target to different degrees across conditions. If residual activation of metrical templates facilitates speech preparation, then stress congruency should influence target production time across predictability conditions. It is also possible, however, that effects that appear to be attributable to local prime-target congruency are driven, or heavily influenced, by metrical predictability. If this is the case, then we expect that global predictability as opposed to local stress congruency will have the dominate influence on reaction times.

In addition to naming latency, we also report the effect of stress congruency and metrical predictability on the phonetic duration of words. Studies of stress priming typically do not analyze the phonetic properties of the words produced, but phonetic duration may provide a sensitive measure of metrical expectancy/priming and reveal relationships between the planning and production of speech.

2. Experimental methods

2.1. Design

The experiment had participants reading both primes and target words aloud from a computer screen. Words were presented one at a time, but the sequences of words followed fixed metrical patterns which depended on condition. The design is summarized in Table 1. The congruency of stress between the target word and the immediately preceding prime was a within-subjects factor in the experiment. The overall metrical predictability of the word lists, as conditioned by the stress pattern in three words preceding the target, the primes, was a between-subjects factor. The string of six letters in each cell of Table 1 indicates the type of words in the hextuple for that condition. The letters correspond to the following word types: "T" = trochee, "I" = iamb, "R" = Random. The target iamb is given in bold font.

Word presentation was organized according to hextuples, sequences of six words. The 5th word of each hextuple was the main word of interest. These words were always iambs (e.g., *cadét*, *paráde*, *ravíne*, *sedán, lapél*). The three words preceding these targets, the primes, were composed of different mixtures of iambs and trochees, depending on the condition. The first and the sixth word of each hextuple always drew randomly from lists of iambs and trochees.

In the "congruent" condition, the prime immediately preceding the target contained stress on the second syllable, the same position as the target iamb. In the "incongruent" condition, the prime immediately preceding the target had stress on the first syllable.

The metrical predictability of the entire sequence of primes and targets was a between subjects factor. For the high metrical predictability group, the target iamb always came after either three consecutive iambs or three consecutive trochees. For the high predictability group, all of the primes were iambs in the congruent condition and trochees in the incongruent condition. For the low metrical predictability group, the order of primes in the congruent condition was trochee, trochee, iamb (TTI). In the incongruent condition, the order of the primes was trochee, iamb, trochee (TIT). These patterns, when embedded between two random items, the 1st and 6th items of hextuples, amount to a less predictable stress pattern than repeating the stress patterns across the primes.

Stress position of the prime (within subjects)	Metrical predictability (between subjects)	
	High	Low
Incongruent	R TTT **I** R	R TIT **I** R
Congruent	R III **I** R	R TTI **I** R

Table 1: *Target words, shown as letters in boldface, were always iambs (I) in the fifth word of a hextuple. The metrical pattern of the the primes, the three underlined words, differed across conditions.*

2.2. Materials

All words in the experiment were disyllabic monomorphemes of shape CVCVC. There were 25 iambic primes, 25 trochaic primes, and 10 iambic target words. Iambic and trochaic primes were matched for initial consonants. The entire list of stimuli is provided in Table 2.

2.3. Participants

Participants were 47 native speakers of English recruited from the UWS community including the psychology student pool. Psychology students received course credit for participation. Other participants were paid $10. 24 participants were assigned to the low metrical predictability group and 23 participants were assigned to the high metrical predictability group.

2.4. Procedure

The experiment was conducted inside of a sound-attenuated booth at MARCS Institute, University of Western Sydney. Participants were seated in front of a computer monitor and two microphones. One microphone was connected to a signal response box containing a voice key. This microphone was used to register participant responses. The other microphone was used to record participant responses. Participants were instructed to read words from the screen as quickly as possible. In order to keep a fairly constant distance between the participant and the microphones, paper handprints were taped to the table beside the microphones. Participants were instructed to keep their hands on the paper handprints throughout the experiment.

Iambic Primes		Trochaic Primes		Targets
baroque	legit	ballad	lettuce	caress
begin	police	beckon	possum	forget
behead	pomade	bonnet	parrot	gazette
belong	panache	bosom	palace	Japan
forbid	parade	fidget	panic	lapel
forgive	piquet	felon	picket	meringue
gazelle	raccoon	gallop	reckon	morass
cadet	ravine	carob	relish	detach
commit	relief	caret	russet	finesse
canal	receipt	cabin	ribbon	pipette
carafe	saloon	canon	salad	
cassette	sedan	cosset	Saturn	

Table 2: *Stimulus items*

Before beginning the experiment, participants completed a practice session. During the practice session participants read each of the sixty words (10 targets + 25 trochaic primes + 25 iambic primes = 60 words) in the experiment one time in random order. The purpose of the practice was to (1) make sure that participants produced the target words with the intended stress pattern (2) familiarize participants with the words to minimize against possible repetition priming effects and (3) make sure that participant responses triggered the voice key. During the practice session, the experimenter remained with the participant. After the practice session, the experimenter left the booth, and participants began the experiment.

Each trial began with a 400 ms display of a crosshair. After the cross hair, a word was displayed in the center of the computer screen. A wave file time-locked to the presentation of the word on the screen was opened to record participant responses. Responses were recorded through an Edirol soundcard directly to the hard drive using one microphone. The other microphone was connected to a voicekey which triggered the end of the trial. The experiment was implemented in E-Prime 2.0.

Words were displayed following the above trial procedure according to the hextuple structure in Table 1. Congruent and incongruent hextuples were presented at random. A total of 48 hextuples (24 congruent and 24 incongruent) were presented to each participant. Target iambs, trochaic primes, and iambic primes were drawn randomly from the word lists described in the materials. The first and sixth members of the hextuples, the random items, were pulled from either the trochaic prime list or the iambic prime list.

2.5. Data processing

The experiment produced a total of 2256 target iamb trials for analysis, each of which was recorded in a separate wav file. The wav files were opened individually in PRAAT. Three temporal intervals were measured: (1) reaction time, the duration from the start of the file (which was time-locked to stimulus presentation) to the onset of speech (2) the duration of the first syllable and (3) the duration of the second syllable. A research assistant, who was a native speaker of Australian English, coded each response for the position of stress, first syllable or second, and for any abnormality in the recording including speech errors. Speech errors or other abnormalities were identified in 391/2256 sound files (17.3%). Of these, the most common error, 65% of all errors, involved a shift in stress from the second syllable to the first syllable. In this paper, all atypical productions including speech errors were excluded from analysis. In addition, words with reaction times greater than two standard deviations from the mean were excluded from analysis. This eliminated an additional 80 trials, 4% of the data. The remaining 1785 tokens were included in the analysis.

Reaction times were recorded by the voice key. They were also measured individually by inspection of the acoustic signal in PRAAT. The hand-measured reaction times showed fewer outliers and lower variability than the reaction times recorded by the voice key. We report only the hand-measured reaction times here; however, both sets of reaction times show the same overall pattern of results.

3. Results

Figure 1 shows the mean naming latency across conditions. The bars indicate reaction time for target words preceded by iambs (white bars) and by trochees (striped bar) in low metrical predictability (left side) and high metrical predictability (right side) conditions. Naming latencies were slightly faster when targets were preceded by iambs than when targets were preceded by trochees. This difference, although it is in the direction of a priming effect, was very small. In contrast, there was a large difference in reaction times across metrical predictability conditions. Reaction times were faster for both prime conditions in the high metrical predictability condition than in the low metrical predictability condition.

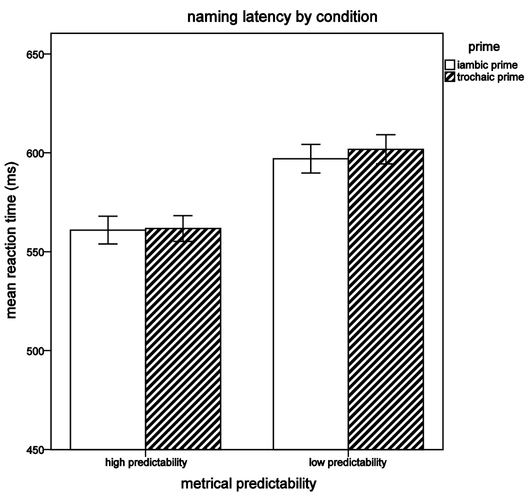

Figure 1: *Reaction times were faster in the high predictability condition than the low predictability condition, but the stress of the prime (iamb/trochee) did not have a significant effect.*

To assess the reliability of the effects shown in Figure 1, a repeated measures ANOVA with prime condition {congruent, incongruent} as a within-subjects factor and metrical predictability {low, high} as a between-subjects factor. The effect of metrical predictability was statistically significant [$F(1,45)$ =9.33, $p < .01$]. Neither the effect of congruency [$F(1,45) < 1$] nor the interaction between congruency and predictability [$F(1,45) < 1$] were significant.

In addition to naming latency, we also investigated the effect of the two conditions, congruency and predictability, on 1st syllable duration, 2nd syllable duration, and total duration. The same pattern of results was obtained for each of these three measures of phonetic duration. We therefore report results only for total duration here. Figure 2 shows the total duration of target iambs in each condition. The white bars show the duration in the congruent condition (iambic targets preceded by iambic primes), and the striped bars show word duration in the incongruent condition (iambic targets preceded by trochaic primes). The bars on the left side of the figure show word duration in the high metrical predictability condition and the bars on the right side of the figure show word duration in the low metrical predictability condition.

Figure 2 shows that the target words were substantially shorter in the high predictability condition than in the low predictability condition. The effect of prime congruency on duration is very small. In the low metrical predictability condition, target iambs were slightly longer when preceded by trochees than when preceded by iambs. The high metrical predictability condition showed the opposite pattern. Target words were slightly shorter when preceded by trochees than when preceded by iambs. A repeated measures ANOVA on total word duration was conducted. Again, prime congruency was a within-subjects factor and metrical predictability was a between-subjects factor. Only the effect of metrical predictability was significant [$F(1,45)$ =10.53, $p < .01$]. Again, the effect of prime congruency [$F(1,45)$ =1.86, $p < .179$] and the interaction between prime congruency and metrical predictability [$F(1,45)$ =2.55, $p = .117$] were not significant.

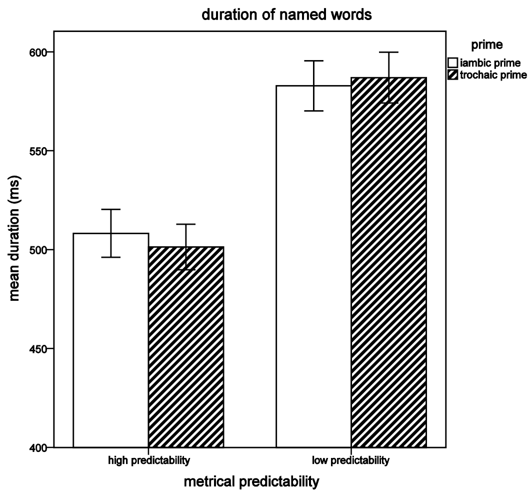

Figure 2: *Prime type did not have a significant effect on word duration, but words were produced with shorter duration in the high predictability condition than in the low predictability condition.*

4. Discussion

This is the first study to our knowledge to investigate both stress congruency and metrical predictability together. It is also the first to investigate these influences on both onset latency and the phonetic duration of words. The main result is that metrical predictability and not stress congruency has a strong effect on naming latency. We also found that metrical predictability and not stress congruency influences the duration of words. We first discuss the reaction time results and then move on to the duration result.

The effect of metrical predictability on naming latency suggests that speakers track metrical patterns implicitly and use the experienced patterns to make predictions about the stress of upcoming words. Sensitivity to implicit metrical patterns could potentially account for speaker behavior in stress assignment and may play a role in lexical access in speech production. The global metrical pattern leading up to the nonce word, as in [2], may bias stress placement towards the more likely metrical pattern. In this case, the local avoidance of stress class or lapse may be a narrow slice of a more global sensitivity to metrical rhythm. Lexical access in speech production may be guided by the metrical rhythms that emerge organically over the course of speech production.

The failure of stress congruency to affect reaction times in either predictability condition may suggest that stress in English is not like other aspects of lexical representations that can be readily primed. The term "lexical stress" encodes the assumption that information that is not predictable by rule must be stored in the lexicon [6]. However, there are now a much wider range of theoretical options available to modern psycholinguistics. First, there are more recent proposals in generative phonology that account for the distribution of stress without requiring that stress be stored in the lexicon (e.g., the theory of lexically indexed constraints in [7]) and proposals that specify stress even when it is predictable (e.g., phonetically detailed representations in [8]). Second, there is a broad trend towards identifying probabilistic as opposed to deterministic rules. This trend has transformed "exceptions" to deterministic rules into well-behaved probabilistic patterns [9], and there is a growing body of literature demonstrating that speaker behaviour is sensitive to probabilistic patterns.

The assumption that stress is stored in the lexicon just when it is not predictable by rule is a relic of the perspective that rules are deterministic and exceptions are listed. The dominant theoretical trend is now that rules are probabilistic and exceptions are too. The sensitivity of local stress priming to predictability based on more global metrical patterns supports this view. Against the backdrop of these more recent theoretical developments, it should be less surprising that local stress priming is a somewhat fragile phenomenon. In online speech production, it may be more important to track the metrical patterns that emerge from speech so that these can be used to guide lexical selection in speech production. Note, that it is not then necessary to store stress in an abstract metrical template with a lexical item. It is only necessary to index lexical items to the metrical contexts they occur in.

The second main result of this study was that metrical predictability has a significant influence on the duration of words. The phonetic duration of both the first and the second syllable were shorter in the high predictability condition than in the low predictability condition. This pattern means that in the condition in which naming latencies were faster, word durations were also shorter. It is not clear from this study whether the shorter word durations are related in principle to the reaction times or to the metrical predictability manipulation. If planning involves a type of internal simulation, it may be possible to relate longer phonetic targets to slower naming latencies. Alternatively, the result could potentially be derived from phonological competition in a cascading model of speech production [e.g., 10]. Increased competition from trochee competitors in the low metrical predictability condition (relative to the high metrical predictability condition) may both delay the onset of production and increase the phonetic duration of the words (because competition continues even after the onset of production).

To sum up, this was the first study to investigate the effects of stress congruency and metrical predictability on naming latency and phonetic duration. We found a strong effect of metrical predictability on both variables and no effect of stress congruency. We considered the broader role that metrical predictability may play in speech planning and explored possible theoretical explanations for the dual patterning of naming latency and word duration. The pattern of results raises theoretical questions about the mental representation of stress and the architecture of the human speech production system. Further studies of the relation between speech planning and speech production will be necessary to resolve these issues.

5. Acknowledgements

The author would like to thank Jaydene Elvin for assistance organizing participants and segmenting speech files, members of the Cornell University phonetics lab, particularly Sam Tilsen, attendees of the PoRT workshop in Glasgow, and three anonymous reviewers for comments that have improved this research. Errors and oversights are all my own. The research was supported in part by ARC grant #DE120101289.

6. References

[1] Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8*, 249-336.

[2] Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 389-403.

[3] Roelofs, A., & Meyer, A. S. (1998). Metrical Structure in Planning the Production of Spoken Words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(4), 922-939.

[4] Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences, 22*, 1-75.

[5] Colombo, L., & Zevin, J. (2009). Stress Priming in Reading and the Selective Modulation of Lexical and Sub-Lexical Pathways. *PloS one, 4*(9), e7219.

[6] Halle, M. (1959). *The Sound Pattern of Russian*. The Hague: Mouton.

[7] Pater, J. (2000). Nonuniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology, 17*, 237-274.

[8] Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 101-139). New York: Mouton de Gruyter.

[9] Plag, I. (2006). The variability of compound stress in English: structural, semantic and analogical factors. *English Language and Linguistics, 10*(1), 143-172.

[10] Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes, 21*(6), 649-683.