DE GRUYTER MOUTON
DOI 10.1515/lp-2013-0007 —— Laboratory Phonology 2013; 4(1): 159–189

Jason A. Shaw

The phonetics of hyper-active feet: Effects of stress priming on speech planning and production

Abstract: This study employs a stress priming paradigm to investigate sensitivity to metrical structure in speech planning and production in Australian English. Target words with iambic stress were preceded by primes with either congruent or incongruent stress and also embedded in metrical contexts biased towards either persistent foot types (Experiment 1) or variable foot types (Experiment 2). Both naming latency, the time from stimulus presentation to the onset of speech, and phonetic patterns showed sensitivity to metrical manipulations. The paradigm produced stress errors, iambic targets produced as trochees, and variation in vowel formants and syllable duration as a function of metrical context. Patterns in the reaction time data indicated sensitivity to global metrical biases calculated over feet. When the metrical bias was toward persistent feet, iambs were produced more quickly in the congruent stress context. When the metrical bias was reversed, iambs were produced more quickly in the incongruent stress context. This pattern of results supports a speech production model that represents metrical structure and allows competition at the metrical level to influence phonetic variability.

Jason A. Shaw: MARCS Institute, University of Western Sydney. E-mail: j.shaw@uws.edu.au

1 Introduction

What is the relationship between formal descriptions of language and the cognitive mechanisms engaged in online speech production? We address a specific instance of this question in the domain of metrical feet. Foot structure is a valuable device in phonological theory, contributing to the formal description of distributional patterns and the domain of phonological and phonetic processes in a wide range of languages, e.g., Slave (Rice 1987), Japanese (Poser 1990), Chinese (Duanmu 1995), Sesotho (Demuth 1994), and Uspanteko (Bennett and Henderson 2011). In English alone, besides its role in conditioning stress patterns and related vowel alternations (e.g., Hayes 1981), foot structure has been described as the domain for phonetic patterns such as flapping (Kiparsky 1979) and aspiration (Jensen 2000).

Models of speech production are on the whole less committed to foot-like units. Some have set aside metrical structure entirely by focusing on monosyllabic words to resolve other factors that influence planning times, such as neighborhood density (Gordon and Dell 2001). Speech production models that incorporate foot-like units include coupled oscillator models that have used foot-sized oscillators to capture patterns of phonetic duration (O'Dell and Nieminen 1999; Barbosa 2007; Saltzman et al. 2008) or phonetic variability (Tilsen 2009a) and models that incorporate prosodic templates specifying syllable count and location of stress, such as WEAVER (Roelofs 1997). These models predict that metrical structure can potentially influence the time it takes to plan utterances (e.g., Roelofs and Meyer 1998) or the phonetic duration of syllables (e.g., O'Dell and Nieminen 1999). However, data that look simultaneously at these two aspects of speech production, planning time and phonetic variability, are scant and, in the domain of metrical structure, entirely absent. This study seeks to fill that gap, providing a look at how varying the local and global metrical context influences reaction times in word naming and the resulting phonetics.

Our main result is that, in word naming, the metrical context of target words has a substantial influence on phonetic variables in addition to influencing the time required to initiate speech production and the frequency of errors. These data demonstrate a clear need for rhythmic structure in models of speech planning and production. Secondly, the codependence of phonetic parameters and reaction times on the main experimental manipulation suggests that models of speech production planning should be constrained both by planning times and by patterns of phonetic variability.

1.1 Stress priming

A variety of experimental paradigms have been developed to investigate the role of metrical structure in speech planning. In the "implicit form priming paradigm", participants are asked to memorize word pairs (Meyer 1990, 1991). One word of the pair is used to cue the production of the other word. In one of a series of experiments using this paradigm, Roelofs and Meyer (1998) tested for stress priming by asking participants to memorize tri-syllabic word pairs that either matched in stress or did not match in stress. Although this paradigm produced facilitory effects for various aspects of shared structure, shared metrical templates alone did not facilitate naming times. This null result of stress priming is actually a prediction of their computational model because metrical templates are accessed in parallel to segments and at about the same speed (Roelofs and Meyer 1998). In the implicit form priming paradigm, repeating metrical templates speeds reaction times only if segments are also repeated. This result is consistent with models of speech production that do not initiate speech movement until segments are associated with a metrical template (Levelt et al. 1999)

Another experiment investigating stress priming used auditory primes in picture naming (Schiller et al. 2004). The auditory primes either matched or mismatched the stress pattern of target words to be named. If processing the auditory prime raises the activation of a segment-independent metrical template, then, in principle, the activation of that metrical template could facilitate reaction time in naming. However, this task also did not show effects of stress congruency. Although auditory primes facilitated naming latencies when the primes were semantically related to the pictures, there was no effect of stress congruency. One possible reason for the null result discussed by Schiller et al. is that the stimuli were not blocked according to target stress pattern. Disyllabic target words with stress on either the initial or the final syllable were randomly presented with either congruent or incongruent primes. This design ensured that the predictability of target word stress was minimized. Subsequent studies have placed target words in more predictable environments and shown positive stress priming effects.

In a design that did not mask the metrical predictability of target words, Colombo and Zevin (2009) demonstrated facilitory effects of stress congruency in Italian. In their design, participants read tri-syllabic words and nonce words from a computer screen one at a time. A sequence of five primes with either congruent or incongruent stress was followed by a target, which was also read aloud. They showed that nonce primes with dominant stress (on the penult) caused a high number of stress errors on real word targets with non-dominant stress (on the antepenult). Reaction times were faster for targets with non-dominant stress when preceded by stress congruent primes than when preceded by stress incongruent primes. Since the sequence of five consecutive prime words with the same stress was not disguised, it may have served to predict the stress of the target. The predictability of stress on target words may have contributed to the faster reaction times found in the congruent stress condition.

Further evidence that stress priming is possible and that predictability augments stress priming effects can be found in Sulpizio et al. (2012). Sulpizio et al. flashed prime words on a screen for 63 ms before presenting trisyllabic Italian target words to be read aloud. The prime words had either congruent or incongruent stress. In the first experiment, trials were blocked by the stress position of the target word. Targets with dominant stress were presented in a separate block from targets with non-dominant stress. Within a block, all target words had the same stress pattern. Results showed a significant effect of stress congruency on both target types. In Experiment 2, the target types (dominant and non-dominant) were mixed in the same block to alleviate the effects of target stress predictability. Reaction times were significantly slower in the mixed block, and, although the effect of stress congruency was still significant, the effect size was greatly reduced. For non-dominant targets the mean difference in reaction time between congruent and incongruent primes narrowed from 37 ms in Experiment 1 to 14 ms in Experiment 2. In this type of experiment, stress congruency has the largest influence on reaction times when target stress is predictable.

Taken together, the results on stress priming reviewed above indicate that both stress congruency, dictated by local stress context, and stress predictability, dictated by more global metrical patterns, can have an influence on planning time in speech production. How these factors influence the phonetic parameters that actuate stress is unknown. None of the studies above report phonetic measurements of target words. Conceivably, the phonetic expression of stress is independent of metrical template selection in speech planning. If so, then no relation between metrical context and the phonetics of stress is expected. Alternatively, since stress can influence the duration of syllables and the quality of vowels, uncertainty about metrical structure might translate into phonetic variability in these dimensions.

1.2 Experiment overview

The experimental paradigm was designed to elicit stress errors and also variation in reaction time on the same set of target words. The language investigated was Australian English. Two experiments using the same paradigm are reported. Both experiments manipulated the congruency of stress between iambic target words and preceding primes. Iambic targets were chosen because iambs are the non-dominant stress pattern in English, and past work (reviewed above) has shown greater priming effects for target words with non-dominant stress. Across experiments the global metrical pattern in the stimulus items was manipulated. This was done by controlling the metrical pattern of prime words that were not local to the target. In Experiment 1, the global metrical pattern introduced a bias toward persistent foot types. In Experiment 2, the global bias was for variable foot types.

Patterns of stimulus presentation in the two experiments are summarized in Table 1. Unbeknownst to participants, stimulus presentation was organized into sequences of six words, or hextuples. All hextuples began and ended with a word sampled randomly from the combined list of iambic and trochaic primes, described in the materials section below. The target word was the fifth word of the

Table 1: The structure of stimulus presentation across experiments and conditions.

Structure of the stimulus (hextuples)	Exp1 Persistent foot types		Exp 2 Variable foot types	
	Congruent	**Incongruent**	**Congruent**	**Incongruent**
(1) random	random	random	random	random
(2) prime	iamb	trochee	trochee	trochee
(3) prime	iamb	trochee	trochee	iamb
(4) prime	iamb	trochee	iamb	trochee
(5) target	**iamb**	**iamb**	**iamb**	**iamb**
(6) random	random	random	random	random

hextuplem and the prime words were the 2nd, 3rd, and 4th words of the hextuple. Accordingly, the words in all hextuples (both congruent and incongruent hextuples across Experiment 1 and Experiment 2) followed the general structure of random(1), prime(2), prime(3), prime(4), target(5), random(6).

The purpose of the random words flanking prime-target sequences was to disguise the structure of the stimulus. To the participant, words simply appeared one at a time at equal time intervals. Post-hoc questioning indicated that participants were unaware of any pattern in stimuli presentation. The generic structure of word presentation is shown in the first column of Table 1. Columns two through five show how this generic structure was instantiated in each condition of each experiment. What varies across conditions is the mix of iambs and trochees contained in the prime words, words two, three, and four of the hextuples.

In Experiment 1, all three primes in the congruent condition were iambs and all three primes in the incongruent condition were trochees. This pattern introduces a mild experiment-wide bias towards persistent, or repeating, foot types. For example, given a trochee target, it is slightly more likely in Experiment 1 that the next word will also be a trochee than that it will be an iamb. Visa versa, given an iamb, it is more likely that the next word will be an iamb than that it will be a trochee in Experiment 1. Experiment 2 introduces a different experiment-wide bias. It also has congruent and incongruent hextuples, but the global bias is towards variable, or alternating, foot types. In the *congruent condition* in Experiment 2, the prime word immediately preceding the target word has the *same stress pattern* as the target. In the *incongruent condition*, the prime immediately preceding the target has *different stress* than the target. Where Experiment 2 differs from Experiment 1 is in the metrical structure of the 2nd and 3rd words of the hextuple. Instead of introducing an experiment-wide bias towards persistent foot types, Experiment 2 introduces a bias towards variable foot types. The primes in the congruent hextuple of Experiment 2 (Table 1: column 4) follow a trochee,

trochee, iamb pattern. The primes in the incongruent hextuple of Experiment 2 (Table 1: column 5) follow a trochee, iamb, trochee pattern. The foot structure of the 2nd and 3rd words of the hextuple in Experiment 2 results in an experiment-wide bias towards variable feet meaning that, given a trochee, it is more likely that the next word will be an iamb than that the next word will be a trochee.

Alongside the bias towards persistent (Experiment 1) and variable (Experiment 2) foot types, there is also a soft bias in Experiment 1 towards iambic stimuli. Each experiment contains three primes and a target in the congruent condition and three primes and a target in the incongruent condition. Of these eight words (3 primes + 3 primes + 2 targets), five of them are iambs in Experiment 1. In Experiment 2, exactly half, four of the eight words, are iambs. The increased frequency of iambs in Experiment 1 allows us to evaluate whether target stress predictability influences behavior in word naming.

To sum up, both experiments contain the local congruent/incongruent prime manipulation. They differ, however, in global (experiment-wide) biases. Experiment 1 includes a slight iamb bias and a bias towards repeating foot types. Experiment 2 is biased towards alternating foot types. These are not deterministic patterns. Rather they are tendencies that may nudge expectations towards persistent or variable foot types and, in doing so, lead to systematic differences in reaction time.

2 Method

2.1 Participants

A total of 47 native speakers of Australian English participated in the two experiments: 24 (7 male) participated in Experiment 1 and 23 (4 male) participated in Experiment 2. The subjects were recruited from the University of Western Sydney community and were either paid $10 for participation or received course credit. No participant reported any history of speech or hearing impairment.

2.2 Materials

All words in the experiment were disyllabic monomorphemes of shape CVCVC. There were 10 target iambs, e.g., *gazette, lapel, meringue, pipette;* 25 iambic primes, e.g., *cadet*, *parade*, *ravine*, *sedan;* and 25 trochaic primes, e.g., *caret, palace, ribbon, ballad*. The initial consonant of iambic and trochaic primes was

Table 2: A list of stimulus words used in the experiment: iambic primes (*n* = 25), trochaic primes (*n* = 25), and target words (*n* = 10).

Iambic primes		**Trochaic primes**		**Target words**
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	
	Tibet		tepid	

matched. The list of trochaic primes included 17 nouns, 6 verbs, and 2 adjectives. The list of iambic primes included 18 nouns, 6 verbs, and 1 adjective. A complete list of the stimulus words is provided in Table 2.

2.3 Procedure and apparatus

Participants were seated in a sound attenuated booth in front of two microphones and a computer screen. Tracings of two handprints were taped to the desk. Participants were instructed to place their hands on the hand prints. This ensured that the participants maintained a fairly fixed position from the microphones.

The experiment was implemented in eprime 2.0. Participants were instructed to read words aloud as soon as they appeared on the computer screen. Words were displayed one at a time in the center of the computer screen. The structure of word sequences was entirely opaque to participants. Each trial began with the 400 ms presentation of a crosshair in the center of the screen followed by presentation of a single word. When the participant read the word aloud, one of the microphones triggered a voicekey, which advanced to the next trial. The other microphone recorded the participant response. A wav file time-locked to the onset of presentation of each stimulus word was opened to record participant responses.

After reading instructions, there was a practice section in which all 60 words in the experiment (25 trochaic primes + 25 iambic primes + 10 target iambs) were displayed in random order on the screen. This eliminated the possibility of repetition priming in the early portion of the experiment. A research assistant remained present during the practice session to ensure that participants were able to read all of the words in the experiment and that participant productions of each word successfully triggered the voicekey.

After completing the practice session, the research assistant left the booth and participants moved on to the main body of the experiment. As with the practice, words were presented on the screen one at a time following a 400 ms crosshair. Participants responded to every trial by saying the word on the screen out loud. One microphone triggered the voicekey advancing the trials and the other recorded participant responses. The responses were recorded through an Eidoral external sound card directly to the computer hard drive.

During the experiment, the order of presentation of words followed fixed metrical sequences summarized above in Table 1. Congruent and incongruent hextuples were selected at random until each hextuple was presented 24 times. A total of 48 target words were displayed to each participant. Stimulus presentation was controlled to ensure that at least 2 instances of each target word were presented in each condition to every participant. The two experiments produced a total of 2256 (1152 in Experiment 1) target iamb trials for analysis.

2.4 Data processing

2.4.1 Excluded data

Responses shorter than 300 ms were excluded from all analyses. This resulted in the loss of two tokens from Experiment 1 and six tokens from Experiment 2. In addition, measurements of vowel formants were not possible for a small number of target words because the vowel was either completely devoiced or partially devoiced such that the period of voicing was too short to obtain reliable information about change in vowel formants over time (see measurement procedure below). A total of 92 vowel tokens, 4% of the target words, did not contribute vowel formant measurements. However, these tokens were included in the analyses of reaction time and syllable duration.

2.4.2 Acoustic measurements: reaction time, syllable duration, and vowel formants

Reaction time and syllable duration were measured from wav files recorded during the experiment and examined individually in Praat. Reaction time was defined as the duration from the start of the file, which was time-locked to stimulus presentation, to the onset of speech visible in the waveform. For sonorant-initial target words, such as *lapel*, the onset of speech in the wav file corresponded to the first consonant of the target word. For target words beginning with voiceless stops, such as *pipette*, the onset of acoustic energy in the waveform corresponded to the stop burst.

Syllable duration measurements reflected the assumption that all words were syllabified as CV.CVC, i.e., the syllable boundary comes before the medial consonant. The duration of the first syllable was measured from the onset of acoustic energy in waveform to the onset of the second consonant.

Once the first syllable was delimited in Praat textgrids, formant values for vowel targets were computed using the following procedure. Formant listings were first extracted at 6 ms intervals using LPC analysis in PRAAT (Burg method with a 6 dB per octave pre-emphasis from 50 Hz) and then imported into MATLAB. In MATLAB, formant listings for each syllable were smoothed using a moving average filter spanning five samples (30 ms). This was done to eliminate noise from the formant listings. To enable comparison across male and female participants, formant ratios were then calculated for F1 and F2 by dividing each F1 and F2 sample by the corresponding value for F3 (see Mohanan and Idsardi [2010] for arguments in support of using formant ratios for vowel normalization). Vowel targets were then determined on the basis of the formant ratios, F1/F3 and F2/F3. The first derivative of the formant ratios, sampled at 6 ms across the interval of visible vowel formants, was calculated to obtain the change in formant values per unit time, a formant change velocity trajectory. The zero-crossings in the velocity signal indicated periods of formant stability. The vowel target was taken to be the first zero-crossing in the formant velocity signal from the right boundary of the syllable.[1] This procedure was applied automatically to all target words.

2.4.3 Stress judgments

The position of stress on the target words was judged by three native speakers of Australian English with some phonetic training. The listeners were told that the sound files were from a speech production experiment in which participants may produce words with stress on the first syllable or on the second syllable. They

---

[1] The zero-crossing was taken from the right boundary of the syllable because it allows consistent application of the same measurement criteria regardless of the identity of the first consonant. Zero-crossings taken from the left boundary of the syllable could potentially correspond to a steady-state portion of an initial sonorant consonant.

Table 3: Pairwise kappa statistic for raters.

	Rater A	Rater B	Rater C
Rater A	–	.736**	.796**
Rater B	.736**	–	.815**
Rater C	.796**	.815**	–

were instructed to judge the position of stress as on either the first syllable or on the second syllable. The listeners were provided with 2256 files from Experiments 1 and 2. They were able to listen to the files as many times as necessary to make a determination of stress. Kappa's test of inter-rater reliability was applied pair wise to the three sets of ratings. Pairwise comparisons were significant at the $p < .001$ level and the average *kappa* statistic was close to .8 (0.78), indicating a high level of inter-rater agreement. The pairwise Kappa scores are reported in Table 3.

In addition, one of the raters who judged the position of stress on the target words also judged the position of stress on the prime words, including all iambs and trochees, an additional 6,768 sound files.

3 Results

The key experimental manipulations in this study, stress congruency within experiments and global metrical biases across experiments, potentially influence a number of different dependent variables. Reaction time, error frequency, first syllable vowel formants, and syllable duration are all reported in this section as a function of the experimental manipulations. The main goal is to investigate the relation between the phonetic variables, syllable duration and vowel formant structures, and the indices of speech planning, reaction times, and errors. The specific prediction for the phonetic variables will depend on whether reaction times and errors are influenced by global predictability or by local congruency. This is because the mechanisms involved in local priming are potentially quite different from those involved in computing experiment-wide expectations. We return to this issue in the discussion (Section 4). This section first establishes what aspects of the experimental manipulations influence planning, as indexed by error patterns and reaction times. These will be used to draw inferences about the mechanisms involved in speech planning. After establishing this, we then report results on the phonetic variables and interpret them in light of the mechanisms evidenced to be involved in planning.

3.1 Behavioural indices of speech planning

3.1.1 Stress errors

All participants produced stress on the second syllable of target iambs during the practice session. Under the experimental conditions, productions of target iambs were sometimes judged to have stress errors, i.e., to have stress on the first syllable. Across experiments, stress errors occurred for 11% of target iambs. These errors were more frequent in Experiment 2 (13.6%) than in Experiment 1 (9.0%). Within experiments, roughly equal numbers of errors occurred in congruent and incongruent stress conditions indicating that stress congruency between prime and target had a negligible effect on stress errors. Table 4 summarizes these results, displaying the frequency of stress errors by experiment and by stress congruency.

Figure 1 shows how the 252 stress errors produced in the experiment are distributed across the different target words, conditions and experiments. The most frequently mis-stressed word was *pipette* followed by *lapel*. These two words accounted for more than half of the total stress errors. The target words, *finesse, morass*, *meringue, caress,* and *gazette* were occasionally mis-stressed. Stress errors for *forget* and *detach* occurred only a few times, and *Japan* was never judged to have stress on the first syllable.

Since the main manipulation in the experiment involves the metrical patterns dictated by the foot type of the primes, it is important to check as well to see that primes were produced with correct stress. Overall, the words selected to be primes were more frequent than the words selected to be targets. Higher frequency words were intended to reduce the number of stress errors on primes. However, this was not the case. There were just as many stress errors on primes, 11.3% across experiments, as on targets, 11% across experiments. Table 5 summarizes the distribution of errors across primes, conditions, and experiments.

It is clear from both the stress error data on targets and the stress error data on the primes that experiments one and two differ systematically. In Experiment

Table 4: Frequency of stress errors on target words by condition and by experiment.

Target words	**Exp 1**		**Exp 2**	
	Congruent	**Incongruent**	**Congruent**	**Incongruent**
Stress errors (trial 5)	50 (9%)	51 (9%)	66 (12%)	85 (15%)
% by experiment	9.2%		13.6%	
% across experiments	252/2004 (11%)			

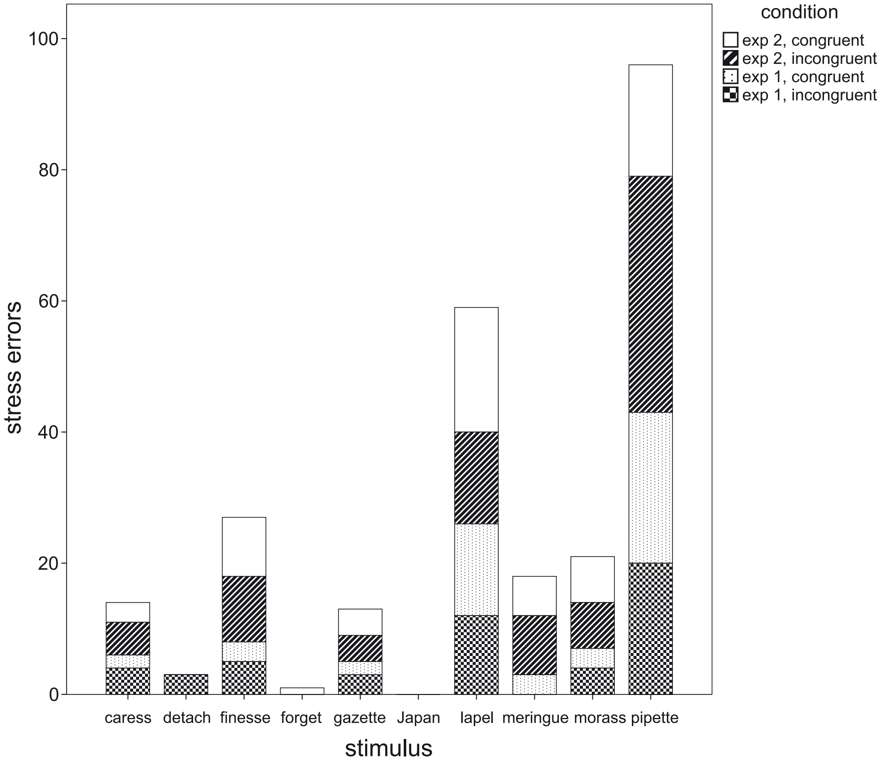

Fig. 1: The number of stress errors for each stimulus item by experiment and condition.

Table 5: Frequency of stress errors on prime words by condition and by experiment.

Prime words	**Exp 1**		**Exp 2**	
	Congruent	**Incongruent**	**Congruent**	**Incongruent**
Prime 1 (trial 2)	72 (13%)	80 (14%)	44 (8%)	45 (8%)
Prime 2 (trial 3)	57 (10%)	77 (13%)	46 (8%)	58 (11%)
Prime 3 (trial 4)	48 (8%)	76 (13%)	46 (8%)	38 (7%)
% by condition	10%	13%	8%	9%
% by experiment	12%		8%	
% across experiments	11.3%			

1, there were more stress errors on trochees than on iambs. This shows up in the comparison of errors on congruent (all iambs) versus incongruent (all trochees) primes. The average error rate for congruent primes (iambs) was 10%, similar to the error rate on target words (also iambs), 9%. The average error rate for incongruent primes (trochees) was greater, 13%. This may be due to the greater number of iambs overall in Experiment 1. Since target words were iambs, and primes were either all iambs (congruent condition) or all trochees (incongruent condition), Experiment 1 had more iambs than trochees overall. Specifically, 5/8 of the stimuli were iambs and 3/8 of the stimuli were trochees. In contrast, Experiment 2 had equal numbers of iambs and trochees, and, in Experiment 2, the reverse pattern was found. There were more errors on iambs than on trochees. This can be seen, first of all, in the high error rates (12% and 15%) on target words (all iambs), and also on prime 2 in the incongruent condition, 11% error rate, which was greater than the trochaic primes, which were all between 7–8%.

3.1.2 Reaction time

3.1.2.1 Experiment 1

Stress judgements were used to partition the reaction times to target words into two categories: iambs, those cases in which the target stress was produced correctly, and trochees, cases of stress errors in which target words were produced with stress on the first syllable. Mean reaction time by stress position and condition for Experiment 1 is shown in Figure 2. Although there is considerably less data for trochees (stress errors) than for iambs, comparison of reaction times is still informative. When target items were produced as trochees (striped bars), they were faster when preceded by trochaic primes (bars on the right) than when preceded by iambic primes (bars on the left). Repeated foot types facilitated reaction times for stress errors. The same effect was not observed for correctly produced stress patterns. Reaction times for targets produced with iambic stress were the same regardless of the stress position of the prime.

To assess the statistical significance of the reaction time patterns displayed in Figure 2, a Linear Mixed Effect Model (see, e.g., Baayen 2008) was fit to the data. The dependent variable was reaction time. Subjects and items were included as random factors.

The main hypothesis was that stress congruency, i.e., whether the target words and the immediately preceding prime word share stress position, influences reaction times. Stress congruency was included as a fixed factor in the model. This factor coded as “1” cases in which the target word and the immediately preceding prime word had stress in the same position, including target

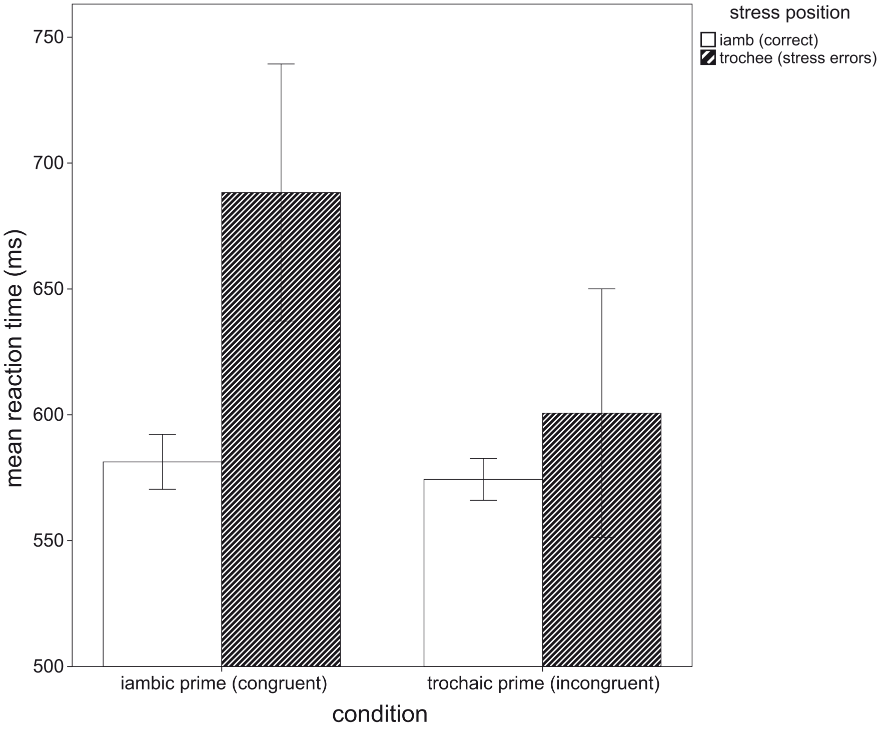

Fig. 2: Mean reaction time by condition in exp 1 for target words produced correctly, as iambs (white bars), and incorrectly, as trochees (striped bars). Error bars indicate 95% confidence intervals.

iambs preceded by iambs and target trochees (stress errors) that were preceded by trochees. The stress congruency factor coded as “2” cases in which the target word and immediately preceding prime had stress in different positions, including target iambs preceded by trochees and trochees (stress errors) preceded by iambs as well as cases of incongruence resulting from errors on the primes, e.g., correctly produced iamb target preceded by a trochee produced errorfully as an iamb. The position of stress was also included as a fixed factor. However, rather than using a binary distinction for stress position, this factor was coded to reflect the level of agreement between the judgments of the three raters. The stress position factor coded the mean rating of stress position for each token. Since interrater reliability was high, most tokens received either a “2” value for unanimous agreement that stress was on the second syllable or a “1” value for unanimous agreement that stress was on the first syllable. For tokens in which two raters judged stress to be on the second syllable and one rater judged stress to be on the first syllable, the value for the stress position factor was 1.67, an average of 2, 2, and 1. This occurred for 22 tokens. Likewise, when only one out of the three raters judged stress to be on the second syllable, a value of 1.33, an average of 2, 1, and 1, was entered for the stress position factor. Stress position therefore codes a quasi-continuous factor reflecting the degree to which three raters agree on the position of stress. In the mean reaction time data shown in Figure 2, it looks as if congruency may have an effect only when target words are produced with initial stress. To assess the reliability of this trend, the interaction between stress congruency and stress position was also included as a fixed factor. A significant interaction would indicate that stress congruency has a differential effect on correctly and errorfully produced targets.

In addition to the main variables of interest, stress position, and stress congruency, the following additional fixed factors were included in the full model: word frequency, grammatical category, trial number, and syllable duration. Word frequency was included because it is known to influence reaction times in word naming and in reading aloud. Trial number was included to assess whether there were any effects of learning, indicated by faster reaction times as the experiment went on, or fatigue, indicated by slower reaction times as the experimented progressed. Since the target words included both nouns and verbs, grammatical category was also included as a fixed factor to assess whether there was a systematic difference in reaction time between these two categories.

Syllable duration was included as a fixed factor in the model for two reasons. First, as described above, reaction times were measured from the acoustic signal. The onset of movement of speech organs precedes the acoustic onset of speech to different degrees depending on the manner of the initial segment (Mooshammer et al. 2012). The onset of acoustic energy for target words that begin with voiceless stops, such as *pipette* and *caress*, does not occur until the release of the consonant. In contrast, the onset of acoustic energy in fricative- and sonorant-initial words follows more closely the onset of articulation. The earlier onset of acoustic energy for fricative- and sonorant-initial words will show up as longer first syllable durations and, complementarily, as shorter reaction times. All else equal, we expect a negative correlation between syllable duration and reaction time to emerge as an artefact of taking both of these measurements from the acoustic signal. The second reason for including syllable duration is that it may provide a measure of metrical competition. If competition between stress patterns is reflected in the phonetics, then syllable duration may predict reaction times. If the trochee competitor exerts influence over the phonetics of iambic targets, we would expect longer syllable duration when reaction times are slower. However, for this effect of syllable duration to show up in the analysis it would have to be

Table 6: Linear mixed effects model of reaction time. '*' indicates $p < .05$; '**' indicates $p < .01$; '***' indicates $p < .001$.

Effect	β	SE(β)	*t*-value	P_{MCMC}
Intercept	619	83.7	7.40	***
Congruency	110	49.6	2.21	*
Stress Position	18.4	40.6	0.45	–
Frequency	−0.10	0.04	−2.41	*
Syllable Duration	−266	96.4	−2.76	**
Grammatical Category	−16.2	13.3	−1.22	–
Trial	−0.04	0.03	−1.25	–
Congruency * Stress Position	−56.9	25.5	−2.23	*

strong enough to overcome the trading relation between syllable duration and reaction time discussed above.

Table 6 summarizes the complete linear mixed effects model of reaction time for Experiment 1. Statistical effects are presented by the estimates of the regression coefficients, β, and by the standard error of β. The *p*-values presented are based on Markov Chain Monte Carlo samples with 1000 simulations. The effect of stress congruency was significant as was the interaction between stress congruency and stress position. This indicates that stress congruency had a differential effect on target words produced as iambs and target words produced as trochees. Stress congruency led to faster reaction times only when target words were produced with stress shifted to the first syllable. In addition to the interaction between stress congruency and stress position, the effects of frequency and syllable duration were also significant. The negative β value for frequency indicates that words of higher frequency were produced more quickly, i.e., with shorter reaction times, than words with lower frequency. The negative β values for syllable duration reflect the trading relation between syllable duration and reaction time. The trend for grammatical category is for verbs to be produced faster than nouns, but this effect does not reach significance. Trial also does not reach significance, but the trend is faster reaction times as the experiment goes on. Lastly, the effect of stress position does not reach significance. Although trochees, which involve stress errors in this experiment, take longer to produce than iambs, the number of trochees is much smaller than the number of items and the variability in reaction time is high.

The reaction time results for Experiment 1 reveal that words with stress errors are produced more slowly than words that are produced correctly. However, the delay caused by errors is attenuated by stress congruency. Errors are produced more quickly when they are preceded by words that share the same stress pattern.

This result occurred in the context of an experiment-wide bias towards persistent foot types. Experiment 2 reverses the bias from persistent feet to variable feet. If participants are sensitive to the global metrical pattern in the stimuli, then the direction of the effect will be reversed from facilitory, in Experiment 1, to inhibitory, in Experiment 2.

3.1.2.2 Experiment 2

Mean reaction time results for Experiment 2 are shown in Figure 3. Overall, reaction times to Experiment 2 were slower than in Experiment 1. As with Experiment 1, the largest effect of stress congruency was on the trochaic realizations of target words. Unlike Experiment 1, however, the effect of congruency is in the opposite direction. In Experiment 1, targets produced as trochees, i.e., speech errors, were produced more quickly when they were preceded by trochaic primes. In Experi-

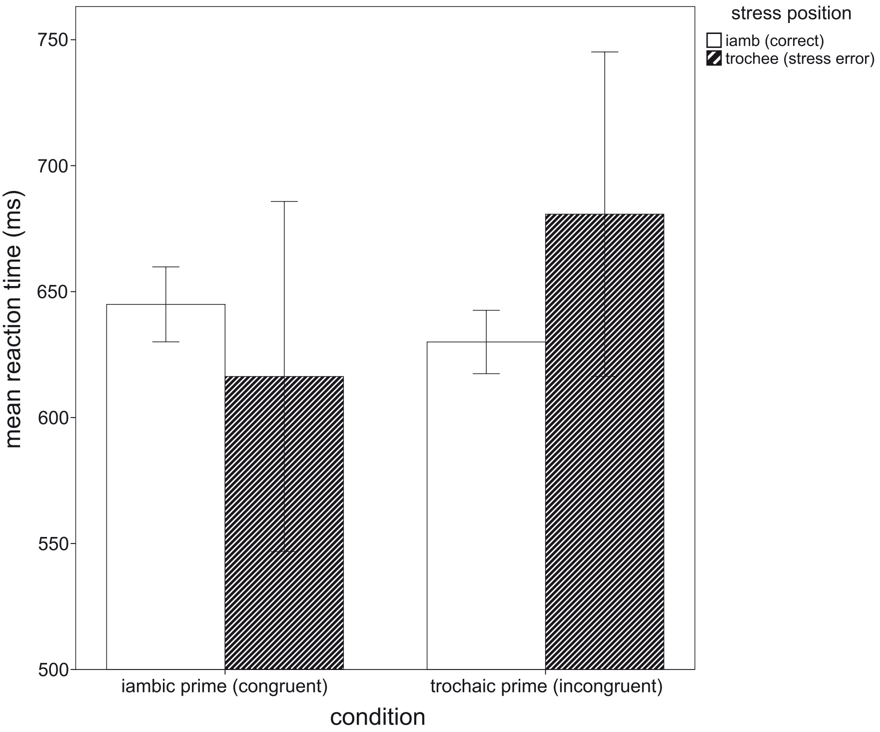

Fig. 3: Mean reaction time by condition in Experiment 2 for target words produced correctly, as iambs (white bars), and incorrectly, as trochees (striped bars). Error bars indicate 95% confidence intervals.

Table 7: Linear mixed effects model of reaction time. '*' indicates $p < .05$; '**' indicates $p < .01$; '***' indicates $p < .001$.

Effect	β	SE(β)	*t*-value	P_{MCMC}
Intercept	818	62.6	13.1	***
Congruency	−15.8	7.90	−2.00	*
Stress Position	−16.9	17.8	−0.95	–
Frequency	−0.20	0.08	−2.36	*
Syllable Duration	−100	104	0.97	–
Grammatical Category	−29.4	29.6	−1.00	–
Trial	−0.30	0.05	−6.25	***

ment 2, trochees were produced more slowly when preceded by trochaic primes. The anti-facilitory effect of stress congruency can also be seen in the reaction times to iambs. Although this effect was smaller, iambs are produced more slowly when they are preceded by iambic primes than when they are preceded by trochaic primes. This result indicates that the global metrical pattern, as opposed to local repetition of foot types, is the dominant force influencing reaction times.

To assess the statistical significance of the reaction time patterns in Experiment 2, again, a linear mixed effects model was fit to the data. The same random effects structure, random effects for both subjects and items, and the same fixed effects, Stress Congruency, Stress Position, Frequency, Syllable Duration, Grammatical Category, and Trial were included. Since the effect of stress congruency influences both iamb and trochee targets, the interaction between Stress Congruency and Stress Position was excluded from the Experiment 2 model.

The linear mixed effects model of reaction time for Experiment 2 is summarized in Table 7. The effect of stress congruency is significant, but its influence on reaction time is in the opposite direction as in Experiment 1. The β values for Congruency in the Experiment 2 model are negative, indicating that reaction times are faster for incongruent prime-target pairs than for congruent prime target pairs. As expected, word frequency is also significant in the Experiment 2 model. As in Experiment 1, stress position and grammatical category are not significant. Syllable duration is also not significant. Lastly, there is a small but highly reliable effect of trial. This indicates that there was a learning effect in Experiment 2. As the trials went on, reaction times got faster.

3.1.2.3 Comparison of Experiments 1 and 2

A comparison of Figures 2 and 3 indicates that reaction times were on average much shorter in Experiment 1 than Experiment 2. To assess the reliability of this difference, a repeated measures ANOVA was conducted on the accurate responses, i.e., target words judged to be iambs. Prime condition {congruent, incongruent} was a within-subjects factor and experiment/metrical pattern was a between-subjects factor. The effect of experiment was statistically significant [$F(1,45) = 9.33$, $p < .01$]. Participants were faster on correct trials in Experiment 1 than in Experiment 2. Neither the effect of congruency [$F(1,45) < 1$] nor the interaction between congruency and predictability [$F(1,45) < 1$] were significant.

Taken together, the results on error patterns and reaction times indicate that the experimental paradigm was successful in scaling the relative activation of foot types. Experiment 1 introduced biases towards iambs and, more broadly, towards persistent foot types. Both of these biases showed up in the data. In line with the iamb bias, there were fewer errors on iambs than on trochees in Experiment 2. Moreover, when there were errors in Experiment 1, iambs produced errorfully as trochees were prepared more quickly when preceded by trochaic primes than when preceded by iambic primes. This reflects the global bias toward persistent feet introduced in Experiment 1. Experiment 2 eliminated the iamb bias and reversed the global bias from persistent feet to variable feet. These manipulations resulted in a reversal of the error patterns and a reversal in the direction of the foot repetition effect. When variable foot types are expected, persistent foot types slow reaction time. These results indicate that speakers are exquisitely sensitive to global metrical patterns computed over feet and that this sensitivity results in pre-activation of expected foot types. We now turn to the phonetics to evaluate how uncertainty about metrical structure may influence the phonetic expression of syllables.

3.2 Phonetic variables

As reviewed in the introduction, a systematic relationship between phonetic variation and planning difficulty, indexed by reaction times and error rates, is not necessarily predicted by current models of speech production. The variation induced by structural priming or anticipatory pre-activation of phonological structure may in principle take place at a level of planning that is disconnected from phonetic form (e.g., Levelt et al. 1999). This study provides the opportunity to investigate co-variation between well-studied indices of speech planning and the fine phonetic details of speech. In this section, we focus on two phonetic variables known to be influenced by metrical structure in English, syllable duration and vowel quality. We first report syllable duration in terms of the experimental variables, then provide an analysis of vowel formant variation, and finish the section with an analysis that considers the relative influences of syllable duration and experimental manipulations (stress congruency, stress predictability) on vowel targets.

3.2.1 Syllable duration

Iambs, such as the target words in this study, are expected to have short first syllables. Trochees, such as those that served as prime words, are expected to have comparatively long first syllables. We have already seen that, in conditions biased towards iambs, speakers produce iambs more quickly. This could potentially be attributable to decreased competition from a trochaic metrical template. Certainly at the level of the speech gesture, there is now substantial evidence that speech errors can involve simultaneous articulation (Goldstein et al. 2007; McMillan et al. 2009; McMillan and Corley 2010). Even when speech errors are not perceptible, competition between segments is detectable at the level of muscle contraction (Mowrey and McKay 1990) and in the fine details of the acoustics (Frisch and Wright 2002). Besides segmental competition, online competition from the lexicon has been argued to persist even after the onset of speech articulation (Goldrick and Blumstein 2006). Generalizing the notion of persistent online competition between phonological structures to metrical feet provides a mechanism capable of accounting for the reaction time results reported above (we return to this in Section 4). It also makes predictions about syllable duration. Competition from a trochaic foot may impinge on an iambic target by lengthening the phonetic duration of the first syllable.

Figure 4 reports the duration of the first syllable of target iambs. Figure 4a shows Experiment 1 (iamb bias). Figure 4b shows Experiment 2 (no bias). Syllable duration is presented separately for stress errors, i.e., words judged to have stress on the first syllable (striped bars), and for words produced correctly, i.e., those judged to have stress on the second syllable (white bars). As expected, first-syllable duration was longer for stress errors in both experiments. Within experiments, there were negligible differences in first-syllable duration across congruency conditions. This is not surprising, since stress congruency also did not affect reaction times or error patterns. Across experiments, however, there were sizable differences in syllable duration. For both accurate and errorful pronunciation, syllable durations were longer in Experiment 2 than in Experiment 1. To assess statistical significance of the syllable duration patterns, a repeated measures ANOVA was conducted on the correct responses, i.e., iambs only. Prime congruency was a within-subjects factor, and experiment/metrical predictability was a between-subjects factor. The effect of metrical predictability was significant [$F(1,45) = 10.53$, $p < .01$]. The effect of prime congruency [$F(1,45) = 1.86$, $p < .179$]

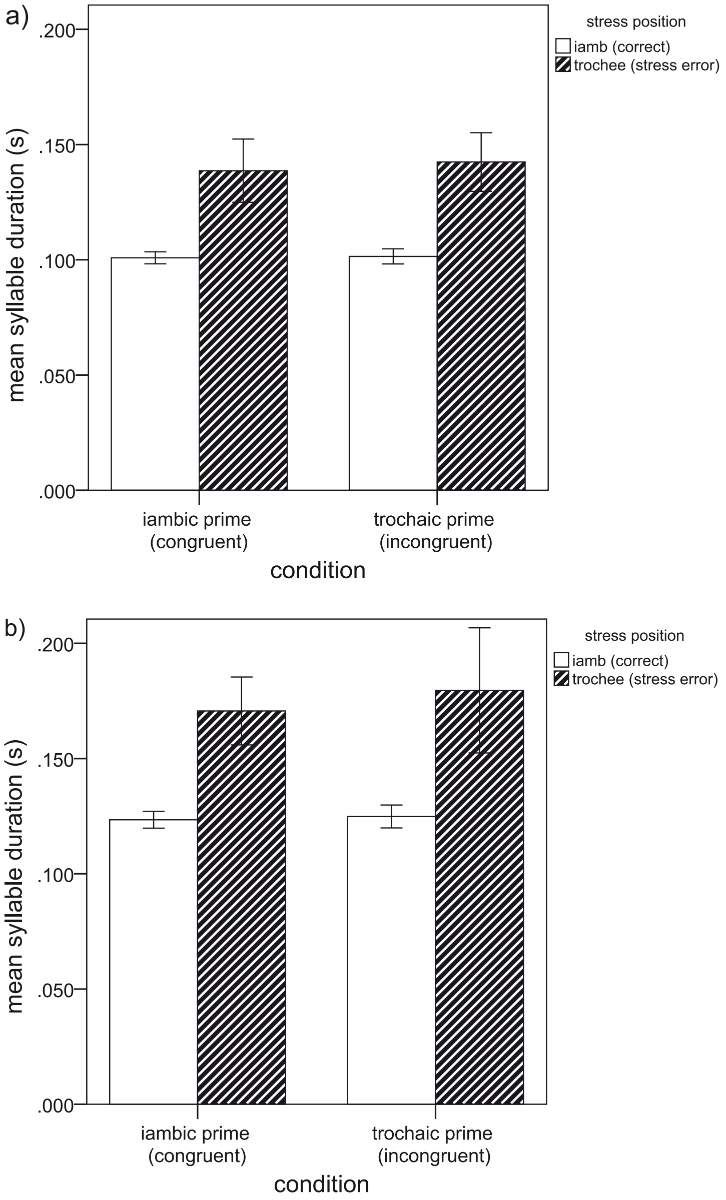

Fig. 4: Average first syllable duration by condition and experiment for target words. Syllable duration was longer in all conditions in Experiment 2, but there was no effect of prime congruency within experiments.

and the interaction between prime congruency and metrical predictability [$F(1,45) = 2.55$, $p = .117$] were not. Speakers produced the first syllable of target iambs faster, i.e., with shorter phonetic duration, in the experiment in which iambs were more likely to occur.

The syllable duration results reflect the metrical pattern manipulation across experiments. While there was no effect of congruency within experiments, there were reliable differences in syllable duration across experiments. This is similar to the reaction time results. On correct responses there were no differences in reaction time or syllable duration within Experiment 1. The change to metrical structure in Experiment 2 resulted in slower reaction times and longer syllables.

The pattern of syllable duration for target iambs supports the metrical competition hypothesis. Correctly produced iambs and targets produced as trochees had shorter first syllables in Experiment 1, where there is a response bias towards iambs, than in Experiment 2. We now turn to vowel quality to evaluate whether changes in metrical pattern across conditions can be observed in vowel formants.

3.2.2 Vowel formants

In order to bring the vowel formant data to bear on the metrical competition hypothesis, we must first consider how formants might change under competition from competing stress patterns. Like other dialects of English, it is typical for unstressed vowels of Australian English to be centralized, or reduced, relative to stressed counterparts (Cox 2012). Vowel centralization can be observed in attested stress alternations for some of the target iambs, e.g., *J*[ə]*pan*~*J*[æ]*panese*, *g*[ɪ]*zette*~*g*[æ]*zetteer*. The basic hypothesis is that vowels influenced by competing metrical templates will take on a quality intermediate between the unstressed vowel and the stressed vowel.

The unstressed vowel in target iambs was either schwa, /ə/, as in *caress*, *forget*, *Japan*, *lapel*, *meringue*, and *morass*, or /ɪ/, as in *detach, finesse*, *gazette*, and *pipette*. These unstressed vowels map to various stressed counterparts. For example, as expected from the *g*[ɪ]*zette*~*g*[æ]*zetteer* stress alternation, the first vowel in *gazette* was produced in stress errors as [æ], while the first vowel in *detach*, also pronounced as [ɪ] when un-stressed, was pronounced under stress as [i]. The realization of schwa vowels, when erroneously stressed, mapped to [æ], [a], [ɜ], or [ɔ], depending on the word. The average formant values across experiments for iambic and trochaic productions of target words are summarized in Table 8. The table shows that the direction of formant change under stress for a given vowel is not uniform. Continuing with the example of [ɪ] in *detach* vs. *gazette*, F1 in *gazette* went up while F1 of the first vowel in *detach* went down.

Table 8: Normalized formant values for each target word produced as both as an iamb and trochee (stress error) across experiments and speakers.

Word		F1	F2	Vowel
detach	iamb	.199	.713	ɪ
	trochee (stress error)	.177	.745	i
finesse	iamb	.192	.708	ɪ
	trochee (stress error)	.210	.659	ɜ
gazette	iamb	.182	.723	ɪ
	trochee (stress error)	.282	.751	æ
pipette	iamb	.182	.711	ɪ
	trochee (stress error)	.206	.699	ɜ
caress	iamb	.229	.707	ə
	trochee (stress error)	.282	.738	æ
forget	iamb	.269	.676	ə
	trochee (stress error)	.150	.619	a
lapel	iamb	.207	.539	ə
	trochee (stress error)	.227	.690	ɜ
meringue	iamb	.229	.663	ə
	trochee (stress error)	.211	.687	ɜ
morass	iamb	.207	.588	ə
	trochee (stress error)	.227	.449	ɔ

If competition at the metrical level induced by the experimental variables influences vowel formants, then we expect less variation in vowel formants in Experiment 1 than in Experiment 2. Experiment 1 introduces a response bias towards iambs which may reduce the level of trochaic competition. If competition results in formant values that are intermediate between the stressed values and the unstressed values, then there will be less formant variation in Experiment 1 than in Experiment 2.

To evaluate the effect of metrical competition on vowel variability in correctly produced iambs, we used Levene's test of equality of variance (Levene 1960). Table 9 reports the mean and standard deviation of normalized formant values across experiments. The table is organized by vowel. The top rows report formant values for words containing [ɪ]. The bottom rows reports formant values for [ə]. The rightmost column reports the results of Levene's test. An asterisk indicates significance at the $p < .05$ criterion. The table shows that the standard deviation of F1 for both vowels is greater in Experiment 2 than in Experiment 1. Levene's test

Table 9: Mean and standard deviation of formant values taken from the first vowel of target words judged to be iambs across experiments. The last column reports Levene's test comparing the variance of formants in Experiment 1 to formants in Experiment 2. The result shows that for both [ɪ] and [ə], F1 is more variable in Experiment 2 than in Experiment 1.

Vowel	Formant	Mean (SD)		Levene statistic
		Exp 1	Exp 2	
ɪ	F1n	.193 (.075)	.185 (.102)	4.23*
	F2n	.698 (.090)	.733 (.078)	.947 (N.S.)
ə	F1n	.242 (.098)	.229 (.108)	5.68*
	F2n	.675 (.122)	.714 (.129)	.141 (N.S.)

demonstrates that these differences are statistically reliable. Differences in F2 variance were not significant. The F1 result supports the predictions of the metrical competition hypothesis. Vowel formants were less variable in Experiment 1, where there was a response bias towards iambs, than in Experiment 2.

The vowel formant patterns reflect the experimental manipulation found to influence reaction time and error patterns. First-syllable vowel formants were sensitive to the level of iamb predictability, showing less variability when iambs were more predictable. Iambs were more likely to occur in Experiment 1 (5/8 of the stimuli) than in Experiment 2 (4/8 of the stimuli). Accordingly, vowel formant patterns for iambs were more stable across conditions in Experiment 1. Under the hypothesis that metrical competition influences vowel formants, the formant stability in Experiment 1 can be understand in terms of a baseline expectancy bias for iambs.

Another key difference between experiments was the bias for persistent foot types (Experiment 1) vs. variable foot types (Experiment 2). We saw in the reaction time data that the effect of local stress congruency reversed depending upon the global context in which local prime-target sequences were embedded. Although Table 8 demonstrates that the formants of Experiment 2 are more variable than Experiment 1, it does not indicate the direction of the influence of stress congruency on formant values. The prediction is that stress congruency will lead to more centralized vowels in Experiment 1 (relative to the incongruent condition). In Experiment 2, the opposite pattern is predicted – because the bias is towards variable feet, the more centralized vowels should occur in the incongruent condition.

To evaluate how formant values are influenced by changes in the global metrical pattern in which target iambs are embedded, we fit a linear mixed effects model to normalized values of the formants. Separate models were fit to F1 and F2 separately, and the data included both iambic and trochaic (stress error) productions of target words. Since we are interested in differences across experiments, the data from both experiments was pooled. Experiment and Congruency were included as fixed factors. The interaction between Experiment and Congruency was coded to test the main hypothesis that, due to differences in the global metrical pattern, congruency has a different effect on formants in Experiment 1 than in Experiment 2.

As we have already observed, stress influences the direction of formant change differentially, depending on the word. To capture this variation, both Subjects and Items were coded as random factors. The fixed factor Vowel coded whether the target vowel was /ɪ/ or /ə/. Stress Position was included as a fixed factor to capture vowel quality differences between trochaic (stress errors) and iambic pronunciations. The different vowel qualities of stressed vowels were not coded for, as this would account for the same variance as Stress Position. Syllable Duration was included as a fixed factor, since formant undershoot can be systematically related to duration (e.g., Lindblom 1963). Finally, Reaction Time was added as a fixed factor to capture the possibility that formant values are related to planning time (e.g., Tilsen 2009b). After fitting a linear mixed effects model with the above factors, factors and interactions that did not account for a significant portion of variance were removed from the model. The final models include the main experimental factors and additional factors that have a significant effect on formant values.

The final model for F1 is summarized in Table 10. Neither of the experimental manipulations, stress congruency, metrical context, nor the interaction between them had a reliable effect on F1. The phonetic variables, vowel identity, syllable duration, and stress position all had significant effects on F1. This result suggests that the value of F1 is not affected directly by the experimental manipulations.

Table 10: Linear mixed effects model of normalized F1. ‘*’ indicates $p < .05$; ‘**’ indicates $p < .01$; ‘***’ indicates $p < .001$.

Effect	β	SE(β)	*t*-value	P_{MCMC}
Intercept	4.26	4.04	1.053	–
Experiment	−0.48	1.42	−0.341	–
Congruency	0.01	1.39	0.007	–
Vowel	7.83	2.05	3.826	***
Syllable Duration	0.07	0.02	4.363	***
Stress Position	2.16	0.86	2.523	**
Vowel * Syllable Duration	−0.04	0.01	−3.398	***
Experiment * Congruency	−0.21	0.89	−0.241	–

Besides the main effects of phonetic variables on F1, there was also a significant interaction between Vowel and Syllable Duration. This interaction indicates that the effect of syllable duration was not uniform across /ɪ/ and /ə/ vowels. Post-hoc correlations between F1 and syllable duration run separately for the two vowels showed that syllable duration is significantly correlated with F1 for /ɪ/ but not for /ə/. This is potentially related to schwa's status as an epenthetic vowel and the minimal influence that co-produced schwa exerts on preceding consonants. The broader picture from the results on F1 is that the interaction between Experiment and Congruency was not found. Moreover, none of the experimental variables showed a significant effect on F1. We now turn to the model of F2, summarized in Table 11.

Table 11: Linear mixed effects model of normalized F2. '*' indicates $p < .05$; '**' indicates $p < .01$; '***' indicates $p < .001$.

Effect	**β**	**SE(β)**	***t*-value**	**P_{MCMC}**
Intercept	548.9	180.2	3.045	**
Experiment	278.9	100.7	2.769	**
Congruency	115.8	75.7	1.531	–
Syllable Duration	2.923	1.006	2.907	**
Stress Position	169.9	47.32	3.591	***
Experiment * Congruency	−95.90	48.46	−1.979	*
Experiment * Syllable Duration	−1.525	0.577	−2.644	**

In contrast to the results for F1, both Experiment and the interaction between Experiment and Congruency had a significant effect on F2. This indicates that stress congruency influenced F2 and that the direction of influence depended on the global metrical context in which target iambs were embedded. In Experiment 1, where there was a bias towards persistent feet, the F2 of target iambs was more centralized in the congruent condition (when preceded by an iamb prime) and more peripheral in the incongruent condition (when preceded by a trochee prime). This was true for both vowels, although the effect was larger for /ə/ (congruent = .665; incongruent = .679) than for /ɪ/ (congruent = .697; incongruent = .699). Experiment 2 showed a different pattern. Although /ə/ was also slightly more central in the congruent condition (congruent = .637; incongruent = .641), the other vowel, /ɪ/, showed the opposite pattern. In Experiment 2, F2 in /ɪ/ became more central in the *incongruent* condition (congruent = .735; incongruent = .730), the opposite of the pattern found in Experiment 1. The significant interaction between Experiment and Congruency on F2 indicates that F2 values, like reaction time (Section 3.1), are sensitive to the global metrical pattern in the stimuli.

Some phonetic factors also had a significant influence on F2, although not to the degree to which they influenced F1. Stress Position and Syllable Duration both had a significant effect on F2. However, as indicated by the significant interaction between Experiment and Syllable Duration, the effect of syllable duration was not uniform across experiments. Post-hoc correlations between syllable duration and F2 by experiment showed that syllable duration and F2 were significantly correlated in Experiment 2 but not in Experiment 1. Lastly, the effect of Vowel, /ɪ/ vs. /ə/, did not have a reliable influence on F2. These vowels tended to be distinguished more robustly in F1.

The over-arching picture from the vowel analysis is that the global metrical patterns in which words are embedded influences the quality of vowels. First, we tested the iamb expectancy bias on F2 variability. Iambs were slightly more probable in Experiment 1 than in Experiment 2. We hypothesized that this difference would lead to more variable formant values in Experiment 2, and this hypothesis was upheld in the data. Specifically, F1 was significantly more variable in Experiment 2 than in Experiment 1. F2 showed equal levels of variability across experiments. We next tested another aspect of the metrical differences across experiments. Experiment 1 was biased towards repeating foot types. Experiment 2 was biased towards repeating foot types. We hypothesized that the metrical bias would interact with congruency. When persistent foot types are expected, as in Experiment 1, more centralized vowels were found in congruent stress environments than in incongruent stress environments. Further, this pattern did not persist in Experiment 2, where the expectancy bias was reversed from persistent foot types to variable foot types. In line with our hypotheses, the distribution of F2 values differed systematically across congruency conditions and this distribution was conditioned by the global metrical pattern.

To sum up the results, each of the dependant measures investigated, stress accuracy, reaction times, vowel formants, and syllable duration, showed systematic variation as a function of the experimental variables. The metrical bias in Experiment 1 towards iambs and towards persistent foot types led to fewer errors on iambs, faster reaction times, less vowel variation, and shorter first syllables than in Experiment 2. Local stress congruency, on the other hand, showed no clear effects in either experiment beyond those attributable to metrical predictability.

4 Discussion

Patterns of speech errors have a long history of informing both theories of phonological representations and models of speech production (see Goldrick 2011 for review). Fromkin (1971) noted that patterns of speech errors collected via listener transcription are remarkably well-behaved from the perspective of the phonological grammar. At the level of representation captured by phonetic transcriptions, speech errors appear to follow phonotactic rules. At a more detailed level of phonetic analysis, studies of speech errors using acoustic or kinematic data have found this not to be the case (e.g., Mowrey and McKay 1990; Goldstein et al. 2007). Rather some errors, at least, seem to involve the simultaneous actuation of multiple gestures. This discovery about segmental errors raises a question about stress errors. Namely, do stress errors involve simultaneous articulation of competing stress patterns?

We hypothesized that, as metrical activation was scaled by manipulations of context, articulation of competing stress patterns would show up in the phonetics in two ways. First, the unstressed first syllable of target iambs would be longer in duration when there was greater trochee competition. Second, the vowel formants of that same syllable, the unstressed first syllable of target iambs, would be more variable when there was a stronger influence from a trochee competitor. We found both of these results. Syllable duration was shorter and vowel formants were less variable in Experiment 1, where the stress pattern of the target word was more predictable, than in Experiment 2.

Both the syllable duration results and the formant variability results are reminiscent of the type of intermediate speech production behaviour found in studies of segmental errors (Mowrey and McKay 1990; Frisch and Wright 2002; Goldstein et al. 2007; McMillan et al. 2009; McMillan and Corley 2010). For example, in a study of segmental errors using acoustic data, Frisch and Wright (2002) showed that speech errors included both categorical substitutions as well as gradient compromises in the phonetic dimensions that instantiate categorical distinctions. They point out that the intermediate phonetic targets in their study can be achieved by extending interactive activation (e.g., Dell 1986) to the articulator level, as in Goldrick and Blumstein (2006). Studies of the kinematics of speech errors have supported this view. Speech errors often involve the simultaneous articulation of segments (McMillan et al. 2009; McMillan and Corley 2010). In some cases, the activations observed in the kinematics do not sound like perceptible errors (Mowrey and McKay 1990; Goldstein et al. 2007). The same can be said for the results of the current study with respect to word stress. The longer syllables and formant variation pushed iambic targets in less predictable stress contexts in the direction of trochees. This phonetic variation may have contributed to the increased number of stress errors, iambic targets judged to be trochees, in Experiment 2.

The patterns of stress errors and related patterns of phonetic variation observed in this study seem comparable to the variation observed in careful phonetic studies of segmental errors. Conceptually, it is a simple step from co-activation of gestures to co-activation of metrical templates. Capturing the intermediate phonetic behaviour conditioned by metrical structure in this study seems to require, minimally, integrating a level of suprasegmental structure capable of capturing stress patterns.

The experimental conditions were successful in scaling the amount of time that participants required to plan and initiate production of target words. Along with changes in reaction time came as well changes in phonetic parameters. Both syllable duration and vowel formants showed patterns of variation that can be followed through reaction time and accuracy results. Errors were more frequent in the conditions with greater vowel variability. Reaction times were faster when syllables were shorter. Such patterns of reciprocal variation can be interpreted in two ways. One possibility is that the whole suite of dependent variables, accuracy, reaction time, syllable duration, and vowel formants, are conditioned independently by the experimental manipulations. However, there is an alternative explanation for co-variation. It could be, instead, that co-variation reflects the mechanism involved in planning and producing speech. We have argued that metrical competition provides a unified account of the co-variation found in this study. Lower uncertainty about target stress raises the activation from competitors. This increases the chance of error and slows reaction time. Due to cascading activation, co-activation of competing metrical patterns leads to greater phonetic variability in the initial vowel. This constellation of changes in accuracy, reaction time, syllable duration, and formant values can be traced back to increased competition created by metrical uncertainty.

5 Conclusion

The representation of metrical structure has received substantial attention in phonological theory, but less work has been done to integrate it into models of speech production. The results of this study indicate that rather intricate metrical patterns can be implicitly learned through the act of producing words. The metrical patterns in the experiment conditioned reaction times, with faster responses coming when target words matched metrical expectation. Stress congruency affected the quality of vowels, both categorically, in the case of stress errors, and gradiently when target words stress did not match expectations. The results suggest that metrical structure plays an active role in online speech production and that metrical uncertainty conditions phonetic variation.

References

Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*: Cambridge: Cambridge University Press.

Barbosa, Plinio A. (2007). From syntax to acoustic duration: A dynamical model of speech rhythm production. *Speech Communication* 49. 725–742.

Bennett, Ryan, & Robert Henderson. 2011. Accent in Uspanteko. Paper presented at the CUNY Conference on Endangered Languages, New York, NY.

Colombo, Lucia, & Jason Zevin. 2009. Stress priming in reading and the selective modulation of lexical and sub-lexical pathways. *PloS one* 4(9). e7219.

Cox, Felicity. 2012. *Australian English: Pronunciation and Transcription*. Cambridge: Cambridge University Press.

Dell, G. S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological review* 93(3). 283–321.

Demuth, Katherine. 1994. On the 'underspecification' of functional categories in early grammars. *Syntactic Theory and First language Acquisition: Cross-Linguistic Perspectives* 1. 119–134.

Duanmu, San. 1995. Metrical and tonal phonology of compounds in two Chinese dialects. *Language* 71. 225–259.

Frisch, Stefan A., & Richard Wright. 2002. The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics 30*. 139–162.

Fromkin, Victoria. 1971. The non-anomalous nature of anomalous utterances. *Language* 47. 27–52.

Goldrick, Matt. 2011. Linking speech errors and generative phonological theory. *Language and Linguistics Compass* 5(6). 397–412.

Goldrick, Matt, & Sheila E. Blumstein. 2006. Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes* 21(6). 649–683.

Goldstein, Louis, Marianne Pouplier, Larissa Chen, Elliot Saltzman, & Dani Byrd. 2007. Dynamic action units slip in speech production errors. *Cognition* 103. 386–412.

Gordon, Jean, & Gary Dell. 2001. Phonological neighborhood effects: Evidence from aphasia and connectionist modeling. *Brain and Language* 79. 21–23.

Jensen, John T. 2000. Against ambisyllabicity. *Phonology* 17. 187–235.

Kiparsky, P. 1979. Metrical structure assignment is cyclic. *Linguistic Inquiry* 10. 421–441.

Levelt, Willem J. M., Ardi Roelofs, & Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioural and Brain Sciences* 22. 1–75.

Levene, Howard. (1960). Robust tests for equality of variances. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling, 2*, 278.

Lindblom, Björn. 1963. Spectrographic study of vowel reduction. *Journal of Acoustical Society of America* 35(11). 1773–1781.

McMillan, Corey T., & Martin Corley. 2010. Cascading influences on the production of speech: Evidence from articulation. *Cognition*, 117(3). 243–260.

McMillan, Corey T., Martin Corley, & Robin J. Lickley. 2009. Articulatory evidence for feedback and competition in speech production. *Language and Cognitive Processes* 24(1). 44–66.

Meyer, Antje S. 1990. The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language* 29(5). 524–545.

Meyer, Antje S. 1991. The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language* 30(1). 69–89.

Mooshammer, Christine, Louis Goldstein, Hosung Nam, Scott McClure, Elliot Saltzman, & Mark Tiede. 2012. Bridging planning and execution: Temporal planning of syllables. *Journal of Phonetics* 40. 374–389.

Mowrey, Richard A., & Ian R. A. MacKay. 1990. Phonological primitives: Electromyographic speech error evidence. *Journal of the Acoustical Society of America* 88, 1299–1312.

O'Dell, Michael, & Tommi Nieminen. 1999. Coupled oscillator model of speech rhythm. Paper presented at the XIVth International Congress of Phonetic Sciences.

Poser, William J. 1990. Evidence for foot structure in Japanese. *Language* 66. 78–105.

Rice, Keren. 1987. Metrical structure in a tone language: the foot in Slave (Athapaskan). *Parasession on Metrical and Autosegmental Phonology, Procedings of the Chicago Linguistic Society* 23. 239–253.

Roelofs, Ardi. 1997. The WEAVER model of word-form encoding in speech production. *Cognition* 64(3). 249–284.

Roelofs, Ardi, & Antje S. Meyer. 1998. Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24(4). 922–939.

Saltzman, Elliot, Hosung Nam, Jelena Krivokapic, & Louis Goldstein. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. Paper presented at the Speech Prosody 2008 conference, Campinas, Brazil.

Schiller, Niels O., Paula Fikkert, & Clara C. Levelt. 2004. Stress priming in picture naming: An SOA study. *Brain and Language 90*(1–3). 231–240.

Sulpizio, Simone, Remo Job, & Cristina Burani. 2012. Priming lexical stress in reading Italian aloud. *Language and Cognitive Processes*. doi: 10.1080/01690965.2011.584820

Tilsen, Sam. 2009a. Multitimescale dynamical interactions between speech rhythm and gesture. *Cognitive Science* 33. 839–879.

Tilsen, Sam. 2009b. Subphonemic and cross-phonemic priming in vowel shadowing: Evidence for the involvement of exemplars in production. *Journal of Phonetics* 37(3). 276–296.