



STRESS SHIFT, STRESS CLASH, AND POLYSYLLABIC SHORTENING
IN A PROSODICALLY ANNOTATED DISCOURSE

Mary E. Beckman, Maria G. Swora,
Jane Rauschenberg, and Kenneth de Jong

Linguistics Department, Ohio State University,
Columbus, Ohio, 43210-1229 USA

ABSTRACT

This paper reports on tests of three different stress-timing effects in a corpus of spontaneous utterances, elicited in a task which prompted naive speakers to produce many tokens of relevant target phrases. Several independent transcriptions of stress and intonation patterns were obtained. The intonation markings were used to control for phrase boundaries in interpreting duration measurements for polysyllabic shortening and stress clash effects. The first, but not the second is robustly demonstrated. The intonation patterns and stress transcriptions were also used to understand stress shift and to show where it will occur. The results show the intricate link between intonation and rhythm, making it crucial to obtain judgments of accent and intonational phrasing in experiments on speech rhythms in languages like English.

1. INTRODUCTION

Three phenomena have long been claimed to support the notion that languages like English differ rhythmically from languages like Japanese in being "stress timed". The first is stress shift; the main stress on words such as *thirteen* and *Chinese* is said to move from final position when the word is in isolation to an earlier position in some contexts to maintain a more regular alternation of stressed and unstressed syllables in a phrase. The second is stress clash; final stressed syllables in words such as *abound* or *peach* are sometimes measured to be longer in the same contexts that trigger stress shift in words such as *Chinese*, making a more regular separation between the onsets of the adjacent stressed syllables, and thus a more constant duration for the stress foot. The third is polysyllabic shortening; the stressed syllables in words such as *sleep* or *sticky* are shorter than the segmentally identical monosyllabic forms *sleep* or *stick*, thus maintaining a more constant duration for the word or stress foot as a whole.

Of these three phenomena, stress shift was noted the earliest, and introspective observation of citation forms of phrases involving it has played a prominent role in the development of metrical phonology -- e.g., [1], [2], [3]. However, there has been little attempt to augment these subjective impressions with controlled observations of actual utterances produced by naive speakers. Cooper and Eady's [4] experiment is the rare exception. Experimental studies on the other two phenomena have been more forthcoming [5], [6], [7], [8], [9], [10]. However, results are somewhat mixed, particularly when the corpus controls for utterance-final lengthening.

In order to examine how the three phenomena might manifest themselves in the more natural and intonationally varied context of spontaneously produced discourse, we had subjects describe an arrangement of labeled pictures that elicited productions of relevant target phrases. For example, pictures of a chest of drawers and a vase, labeled *Chinese dresser* and *Chinese antique*, provided a relevant contrast for stress shift. Pictures of people labeled *Kay Sakis*, *Casey Sakis*, and *K. Sasaki* provided contrasts for stress clash and polysyllabic shortening. This paper reports on a preliminary analysis of the measurements and stress transcriptions from 12 spontaneous dialogues obtained in this way. The results suggest that all three stress-timing effects are heavily dependent on the pattern of intonational phrasing and pitch accents.

2. METHOD

In order to get naturally occurring instances of the structures of interest, we had naive speakers participate in dialogues with pseudo-subjects (who were actually members of the experimental team). Four dialogues were elicited from each speaker. The speaker was seated in a sound booth before a microphone and a board on which there was an arrangement of labeled pictures, and was told to instruct another subject (the pseudo-subject) in building a copy of the arrangement at another table outside the booth. He was instructed to use full sentences, since noise would be randomly introduced into the connection that would interfere with the reception of the instructions. In reality, there was no noise, and the pseudo-subject feigned all instances of misunderstanding in order to elicit repetition. The participants produced utterances like those shown in Fig. 1. (Here, PS is the actual subject and KJ the pseudo-subject.)

-
- 488 PS: To the right of two horses is a picture of
Jody Malone.
- 489 KJ: Let's see... Yeah, OK. Jody Malone, to the
right of two horses. OK.
- 490 PS: To the right of Jody Malone is a picture of a
chinese antique.
- 491 KJ: A picture of a shiney antique?
- 492 PS: To the right of Jody Malone is a picture of a
chinese antique.
- 493 KJ: Chinese antique... OK, got that one.
- 494 PS: Below the chinese antique is a picture of K.
Sasaki.
-

Fig. 1. Sample exchange from dialogue 3.

For each utterance containing a target phrase, two or three kinds of analysis were performed: intonation tran-

scriptions and duration measurements (for all of the phrases), and stress transcriptions (for just the targets for stress shift). There were 4 transcribers for intonation and 5 for stress. Each of the transcriptions was made independently of all the others; no one saw anyone else's transcriptions until after completing the task, and the transcribers who transcribed both stress and intonation did these transcriptions at separate times.

For intonation, the transcribers used displayed F0 tracks to augment their aural impressions of the pitch pattern. Transcriptions were made using Pierrehumbert's notation system [11]. In this system, intonation contours are described as a series of low (L) and high (H) tones. Pitch accents are associated with particular stressed syllables. English has six pitch accents: H*, L*, H*+L, H+L*, L*+H, L+H* (where the * indicates alignment with the stressed syllable). A well-formed intermediate phrase contour consists of one or more pitch accents and a L or H phrase accent, which is not aligned with any syllable but simply follows the last pitch accent in the phrase. The last accented syllable then bears the "nuclear accent" (or "sentence stress"). An intonation phrase is composed of one or more intermediate phrases closed with a L or H final boundary tone. The intonation transcribers ranged in experience from having had a one-week introduction to this system in a phonetics course to having had considerable research experience using it.

Duration measurements were made from the digitized speech files using a waveform editor and standard segmentation criteria (e.g., the first vowel in *fifteen* was measured from the zero crossing for the first period of regular voicing after the [f]).

For stress, a numerical transcription was used, with 1 signifying primary stress, 2 indicating secondary stress, 3 indicating any tertiary stress, and a 0 used for unstressed syllables; all syllables in the target phrases were labeled. None of the 5 stress transcribers had any specific instruction in transcribing stress, although all had taken at least one graduate-level course in phonetics.

3. POLYSYLLABIC SHORTENING AND STRESS CLASH

Table 1 shows the target phrases that we used to test for two of the stress-timing phenomena. Each set of three names allows two contrasts. Comparing vowel lengths in two-syllable first names to those in one-syllable names (e.g., *Casey Sakis* vs *K. Sasaki*) tests for polysyllabic shortening. Comparing monosyllabic names in alternating versus adjacent stress patterns (e.g., *K. Sasaki* vs *Kay Sakis*) tests for stress clash. In the comparison for stress clash, we must control for the fact that two syllables do not count as adjacent if they are separated by an intonational phrase boundary. For polysyllabic shortening, also, we must control for the confounding effects of phrase-final lengthening. We used the intonation transcriptions to control for phrasing, excluding tokens judged by a majority of transcribers to consist of two intermediate or full intonation phrases.

Table 1. Name paradigms

Emma Lone	Casey Sakis	Jody Bolt	Lula Moore
M. Malone	K. Sasaki	Joe DeBolt	Lou L'Amour
M. Lone	Kay Sakis	Joe Bolt	Lou Moore

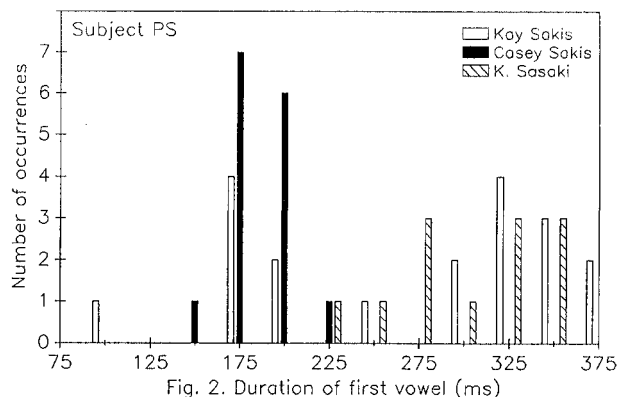


Fig. 2 illustrates our results. It is a histogram of the durations of the vowel [e] in the first syllable of the names *Casey Sakis*, *K. Sasaki*, and *Kay Sakis* produced by subject PS. The figure shows polysyllabic lengthening, but not stress clash. The durations for *Casey Sakis* cluster to the left of the graph, and clearly have a shorter mean than the durations for *K. Sasaki*, but the mean for *Kay Sakis* is shorter rather than longer than that of *K. Sasaki*.

Table 2. Mean durations of first vowels

Subject	LD		MS		PS	
	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N
Joe Bolt	406 (114)	6	218 (56)	9	217 (46)	12
Joe DeBolt	366 (108)	9	192 (52)	6	203 (59)	16
Jody Bolt	232 (78)	16	232 (19)	13	135 (12)	18
Kay Sakis	433 (89)	13	293 (47)	7	262 (87)	19
K. Sasaki	462 (48)	14	298 (31)	9	280 (45)	13
Casey Sakis	252 (23)	8	202 (24)	18	175 (17)	15
Lou Moore	206 (49)	6	176 (46)	22	123 (37)	11
Lou L'Amour	251 (15)	4	205 (51)	6	135 (32)	9
Lula Moore	175 (29)	10	140 (17)	12	104 (17)	8
M. Lone	201 (26)	4	133 (18)	3	214 (42)	5
M. Malone	171 (0)	1	168 (0)	1	243 (48)	4
Emma Lone	145 (16)	5	113 (11)	8	142 (16)	17

ns not significant; *P<.05; **P<.01; ***P<.001

Table 2 shows mean values for all the word sets produced by all four speakers. The values are consistent with the pattern in Fig. 2. Polysyllabic shortening is robustly demonstrated, with the vowels of the target polysyllables being significantly shorter in every case where there were enough tokens without a phrase break to make a meaningful comparison. Stress clash, on the other hand, is not

robustly demonstrated. Even in the few cases where the means differed in the right direction, the difference is not significant.

These results are in keeping with other experiments on the two effects, which generally show polysyllabic shortening to be the more robust and consistent effect. Where stress clash has been demonstrated, the difference in the means is always very small, and some studies (e.g. [4]) have found no difference at all. Polysyllabic shortening, by contrast, is more consistently demonstrated in the literature. For example, Beckman and Edwards [10] found it in all conditions for all subjects except for two subjects in unaccented position at fast rate. In our corpus, the speakers tended to speak rather deliberately, to compensate for the feigned line noise. In keeping with this style of speech, pitch accents were placed on all of the target syllables (and were so transcribed by all 4 transcribers). Perhaps a less deliberate style would have been more conducive to demonstrating stress clash. (It is also possible, however, that some of the tokens included in our means do have intermediate phrase boundaries, which would have been excluded if we had used a more conservative criterion than majority rule.)

4. STRESS SHIFT

There were 4 target phrases for stress shift, shown in Table 3. In each phrase, the target word *fifteen* or *Chinese* precedes another word which either has primary stress on its first syllable (lefthand column) or on its second (righthand column). Metrical theory predicts that the stress can shift to the first syllable of the target word, and is particularly likely to do so in the phrases on the left. For each utterance in the dialogues that contained one of these phrases, we used the stress transcriptions to decide whether the stress was shifted from the second syllable. We also compared the 5 different stress transcriptions to get a measure of the transcriber's confidence in the percept of shifted or unshifted stress.

Table 3. Target phrases for stress shift

Chinese dresser	Chinese antique
Fifteen rabbits	Fifteen antiques

The transcribers seemed to use rather different criteria in deciding what stress number to use for any syllable. For example, only 2 of the transcribers (JR and MB) ever transcribed tertiary stress. Transcriber MS often marked secondary stress on the second syllable in *Chinese* and only twice marked zero-stress for this syllable, whereas EH almost always marked zero-stress here and only once a secondary stress. Despite these apparent discrepancies, the transcriptions were fairly consistent in one way; if we divide them into three categories -- (1) those which clearly indicated unshifted stress by giving the second syllable greater stress than the first, (2) those which showed equal stress on the two syllables, and (3) those which clearly showed shifted stress -- all 5 transcribers agreed in 48% of the tokens of *Chinese*. If we relax the criterion to allow no more than 1 disagreeing transcription, the number of consistent judgements is 82%. The comparable numbers for *fifteen* are 11% and 40%, respectively. We interpret this to mean either that the two syllables of *fifteen* are inherently equally stressed or, more likely, that stress is difficult to perceive on the short first vowel. We will limit the rest of our discussion of stress shift to the more confident judgements of *Chinese*.

Table 4 shows the percentage of stress judgments in *Chinese* that fell into each of the three categories. (The percentages in parentheses will be explained later.) The overwhelming majority of judgments had stronger stress on the first syllable. Moreover, this was true in both phrases. There was a slightly larger percentage of unshifted judgments in the context of *antique* (in keeping with the predictions of metrical theory), but this difference is very small by comparison to the main trend. The primary stress tends to be shifted from the second syllable whether or not the next word has initial stress.

Table 4. Distribution of stress judgments in *Chinese*

stress	dresser	antique	total
unshifted	2%	10% (5%)*	13% (12%)
equal	9%	12% (11%)	30% (18%)
shifted	89%	78% (84%)	69% (70%)

*Numbers in parens exclude tokens with contrast.

This pattern is consistent with Cooper and Eady's results [4]. They used syllable duration and average F0 values as measures of relative stress, and found very little difference between the means for phrases such as *thirteen colleges* versus *thirteen universities*. They conclude from their results that primary stress always falls on the first syllable in these cases. In fact, they even question the notion that stress has shifted, proposing that the primary stress is on the first syllable of *thirteen* even in isolation.

Our results suggest a different conclusion. We note first that although initial stress is a possible alternate isolation pattern for *thirteen* in standard American English, *Chinese* in isolation would sound odd with sentence stress on its first syllable. Second, although shifted stress is the more common judgment, there are judgments of equal or unshifted stress for both *fifteen* and *Chinese* in our corpus. Moreover, some such judgments were obtained from every transcriber, and for utterances produced by every speaker. Thus, while there is a tendency to hear greater stress on the first syllable in these phrases, stress shift is clearly not mandatory.

In what sorts of tokens did transcribers not perceive a shifted stress? One example occurs in line 492 of Fig. 1 above. The speaker PS repeats the phrase *Chinese antique* to correct the pseudo-subject's feigned mishearing as *shiny antique*. In such cases of contrastive correction, a native speaker would usually put the nuclear accent of the sentence on the emphasized word or syllable. In this particular example, the intonation transcriptions suggest that PS has indeed placed the nuclear accent on the second syllable of *Chinese*, consistent with the perception of the isolation stress pattern. All but one of the cases where the majority of transcribers heard unshifted stress were similar; a phrase in the background puts an implicitly or explicitly contrastive emphasis on *Chinese*. When there is no special emphasis on *Chinese*, on the other hand, nuclear accent would be expected to fall elsewhere. At the first reference to a target picture (e.g. line 490 of Fig. 1), the nuclear accent would be expected to fall on the second word of the phrase. In an immediately following locative use (e.g. line 494), the nuclear accent might be expected to fall on the locative preposition, which would be the new information in the phrase.

These patterns of expected nuclear accent placement have two important implications for the interpretation of the stress judgments in our corpus. First, we must control for special emphasis in determining when stress will shift. By coincidence, the utterances with special emphasis on the first word of the target phrase all happened to be tokens where the following word is *antique*. In Table 4, the slightly higher percentage of unshifted judgments in *Chinese antique* relative to *Chinese dresser* could be due to the inclusion of tokens with special emphasis on *Chinese* only in the first set, rather than to the stress pattern in the following word. The numbers in parentheses in Table 4 are percentages for tokens excluding cases where there was an implicit or explicit contrastive emphasis on *Chinese*. When these cases are excluded, the effect of the following word on the distribution of shifted and unshifted judgments is not significant ($X^2[2]=2.40, P>0.3$).

Second, the intonation pattern may provide crucial insights into the percept of stress shift. As Bolinger [12] and others have noted, accents in English can only appear on "strong" syllables (with full vowels). Although a nuclear pitch accent is expected to fall on the syllable with strongest stress in a word, prenuclear pitch accents can go on any strong syllable. Thus, words such as *Chinese* and *fifteen*, which have two strong syllables, can take a prenuclear pitch accent on either without imparting special contrastive emphasis to that syllable. This suggests that, absent any difference in intonational prominence, the two syllables in these words have equal stress; any difference in perceived stress should reflect accent placement.

Table 5. Correlation of judgments for accent and stress

P0 or N0*	PP or 00	ON, PN, or OP
69%	2%	6%
10%	3%	1%
8%	0%	2%

*P=pre-nuclear accent, 0=no accent, N=nuclear accent

In order to test this hypothesis, we compared stress judgments with intonation transcriptions for the transcribers who provided both. The comparisons for *Chinese* are shown in Table 5. The hypothesis predicts that the majority of judgments will fall along the diagonal in this Table: the first syllable will be perceived as having greater stress if it is accented and the second is not; the two syllables will be perceived as equal if both have prenuclear accents or no accent; and the second syllable will be perceived as more stressed if it is accented and the first is not or if it has the nuclear accent and the first a mere prenuclear accent. The hypothesis is borne out in 74% of the cases. We conclude from this result that the stress percept in these words does depend on the pitch accent pattern, but that other factors also influenced the stress transcriptions.

What factors are these? One is that words in English must be marked for an "inherent" primary stress in order to account for strong native speaker intuitions about where to put the nuclear accent in isolation pronunciations. This "inherent" lexical stress no doubt influenced some of the stress judgments, and probably accounts for the 18% of cases where the transcriber did not mark the first syllable as more stressed even though it was perceived as having the stronger accentual prominence. The fact that nearly half of these were marked as having primary stress on the second syllable attests to the strong

psychological reality of the lexical marking. A second factor is the consequent inclination to think of stress as a relational property. That is, since only one syllable can be lexically stressed, native speakers might be biased toward always marking one syllable as stronger. (In fact, the two transcribers with least phonetic training together marked only 1 token of *Chinese* as having equal stress.) This bias probably accounts for the 2% of cases where the two syllables were marked as having equal accentual prominence but unequal stress.

Despite these two conflicting factors, stress and accent markings agreed in 74% of the cases. We conclude that relative stress in words such as *Chinese*, where both syllables are strong, is a matter of accent placement. In isolation or under special emphasis on the word as a whole, the nuclear accent in the intonation contour would normally be placed on the second syllable. In a phrase such as *Chinese antique*, on the other hand, the potential for prenuclear accent on the first syllable means a potential for shifting primary stress away from this "inherent" location. Thus, in understanding stress shift, as in demonstrating polysyllabic shortening, it is crucial to take into account the intonation pattern.

REFERENCES

- [1] M. Liberman & A. Prince, "On Stress and Linguistic Rhythm," *Linguistic Inquiry*, vol. 8, pp. 249-336, 1977.
- [2] B. Hayes, "The Phonology of Rhythm in English," *Linguistic Inquiry*, vol. 15, pp. 33-74, 1984.
- [3] E.O. Selkirk, *Phonology and Syntax*, 1984.
- [4] W.E. Cooper & S.J. Eady, "Metrical Phonology in Speech Production," *J. Memory and Language*, vol. 25, pp. 369-384, 1986.
- [5] I. Lehiste, "The Timing of Utterances and Linguistic Boundaries," *JASA*, vol. 51, pp. 2018- 2024, 1972.
- [6] A.W.F. Huggins, "On Isochrony and Syntax," In G. Fant & M.A.A. Tatham, eds., *Auditory Analysis and Perception of Speech*, pp. 255-464, 1975.
- [7] C.A. Fowler, *Timing Control in Speech Production*, 1977.
- [8] L.H. Nakatani & J.A. Schaffer, "Hearing 'Words' without Words: Prosodic Cues for Word Perception," *JASA*, vol. 63, pp. 234-245, 1978.
- [9] B. Rakerd, W. Sennet, & C.A. Fowler, "Domain-final Lengthening and Foot-level Shortening in Spoken English," *Phonetica*, vol. 44, pp. 147- 155, 1987.
- [10] M.E. Beckman & J. Edwards, "Lengthenings and Shortenings and the Nature of Prosodic Constituency," In J. Kingston & M.E. Beckman, eds., *Papers in Laboratory Phonology I*, pp. 152-178, 1989.
- [11] J.B. Pierrehumbert, *The Phonology and Phonetics of English Intonation*, 1980.
- [12] D. Bolinger, *Two Kinds of Vowels, Two Kinds of Rhythm*, 1981.

Work supported by the National Science Foundation.