Phrasal Prominence Location is Influenced by IP Boundary Location in the Absence of Stress Clash

Soundess Azzabou - Kacem,[1] and Alice Turk[2]

[1]University of Edinburgh, UK & University of Sousse, Tunisia, [2]University of Edinburgh, UK

1 Introduction

The location of prominence in double-stressed words like *thirteen*, *mundane*, *Maltese* is not fixed. Several factors can influence the early assignment of prominence in such words, such as the presence of stress clash with more prominent following words (as in, e.g., *thirteen MEN; Chinese WATCH;* cf. Fig. 1), and their adjacency to domain initial constituent boundaries (Halle & Idsardi, 1995). The goal of the experiment reported in this paper is to test whether the **early** assignment of phrasal prominence in double-stressed words, can occur when they are Intonational Phrase (IP) initial, in the absence of stress clash with a following word.
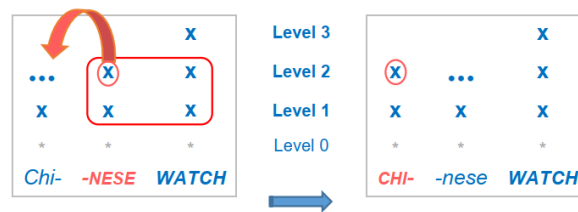

Figure 1. Metrical grid of the double-stressed word *Chinese* representing the possible shifting process from late prominence on syllable 2, *–nese* (left panel) to early prominence on syllable 1, *Chi-* (right panel). The boxed (x)-s are the four metrical beats involved in the movement, or a clash configuration as defined by Selkirk, (1984).

English post-lexical stress clash, as defined by Selkirk (1984), implies that a double-stressed word (e.g., *thirteen* in *thirteen MEN*), can be the site of leftward stress shift from a right-hand 'accentable' syllable (e.g., -*teen*) to another 'accentable' syllable located **before** it (e.g., *thir-*), when the double-stressed word is followed by a **stronger** prominence, (e.g., *MEN*). Contrary to the definition of clash in other languages like German, in English, the directionality of the movement is leftward only (e.g., no shift predicted rightwards from *con-* to -*test* in *SPORTS contest;* cf. Liberman & Prince, 1977, p. 311. Furthermore, only the weaker prominence is the site of potential stress shift; cf. Selkirk (1984).

2 Likely predictors of early phrasal prominence

The **early** location of prominence in double-stressed words like *thirteen* has been studied in pre-nuclear contexts, i.e. in contexts where *THIRteen* is followed by a clashing main phrasal prominence (e.g., *MEN* in *THIRteen MEN*); e.g., Grabe & Warren (1994). Stress clash was claimed to be one of the conditions favouring the realisation of early prominence. But, other studies showed that stress clash is not a requirement for the realisation of early prominence (e.g., Thompson, 1980; Hayes & Puppel, 1984; Hayes, 1995). The first empirical evidence of this kind, provided by Cooper & Eady (1986, Experiment 3), showed that stress clash may not be the only predictor of early prominence in, e.g., *MISsissippi RElatives* (clash) and *MISsissippi LEgisLAtion* (no clash). As both of these phrases were produced in isolation and preceded with silence, which is a correlate of the strongest prosodic boundary, this result raises the question whether

Proceedings of *AMP 2021*

the presence of this utterance-initial boundary before *Mississippi* affected the likelihood of early pre-nuclear phrasal prominence.

In fact, we know that early prominence could serve as a left-edge phrase onset marker, (Bolinger, e.g., 1958b; Halle & Vergnaud, 1987; Shattuck-Hufnagel, Ostendorf & Ross, 1994; Halle & Idsardi, 1995; and for French, Astésano, Bard & Turk, 2007, and Garnier, Baqué Dagnac & Astésano, 2016). Shattuck-Hufnagel et al. studied double-stressed words in a radio news style corpus (the Boston Radio Corpus) of American English. In their data, 12 out of the 14 double-stressed tokens realised with 'Early Accent' were reported *not* to occur in contexts where the main lexical prominences were adjacent. And furthermore, eight out of these 12 tokens were phrase-initial. Although the precise stress clash status of the materials in their corpus was not clear, their evidence is consistent with the view that the presence of a prosodic boundary at the left edge of a phrase is a good predictor of early pre-nuclear prominence.

Here, we provide a further test of the possibility of phrase-initial phrasal prominence, in the absence of stress clash with a following word. Our study consists of a comparison of phrasal prominence locations in double-stressed target words (e.g. *Maltese*) when produced in isolation (and thus surrounded by strong prosodic boundaries) vs. when produced in a (*Say* ___ again) carrier sentence, where the prosodic boundaries surrounding the target are expected to be weaker, and the following word is unaccented and therefore does not cause stress clash.

3   Paradigm and predictions

The main research question we deal with in this paper is whether word-initial prominence is encouraged in isolated double-stressed words, as expected if initial IP boundaries attract early prominence.

3.1   *Experimental Paradigm: manipulation of context*   The strength of the prosodic boundaries surrounding the target item (e.g., *Maltese*) is manipulated to elicit two renditions of the target.

In the **Isolated condition**, the target is produced in isolation and, therefore, surrounded by strong utterance onset and offset prosodic boundaries. The target must carry the phrasal nucleus because it constitutes the only segmental material in the Intonational Phrase.

In the **Embedded condition**, the target is elicited in the frame sentence *Say* [target] *again*, where the target is also expected to carry the phrasal prominence, but now is Intonational-Phrase-medial and surrounded by weaker prosodic boundaries relative to the Isolated condition. Note that the longer phrase creates no stress clash with the following word since the following word *again* does not carry nuclear prominence. Table 1 presents examples of the Embedding experiment paradigm.

Table 1. *Examples of stimuli for targets in the two conditions. In each case, the speakers saw the stimulus at the centre of the screen along with preceding and following fillers.*

Experimental conditions	
Isolated condition	Embedded condition
banana	*Say banana again.*
Maltese	*Say **Maltese** again.*
Rebecca	*Say Rebecca again.*

3.2   *Predictions*   In the Embedded condition (1.b), the phrasal prominence is predicted to fall on the second syllable of the target, (e.g., *-tese* in *Maltese*) as the boundary immediately preceding the target (e.g., *Maltese*) is weaker in this condition than in the Isolated condition, and there is no stress clash with a following word.

(1) **a. Isolated:** $_{IP}$ [$_\omega$ [***Maltese***] $_\omega$] $_{IP}$

     b. Embedded: $_{IP}$ [*Say* $_\omega$ [***Maltese***] $_\omega$ *again*] $_{IP}$

In the Isolated condition (1.a), early prominence (i.e. phrasal prominence placed on the initial syllable *Mal-* in *Maltese*) is predicted to be more likely, as the target word *Maltese* is utterance- and phrase-initial and the prosodic boundary is stronger. In other words, if this embedding manipulation elicits a higher incidence of early prominence in the Isolated condition, then this will constitute evidence that prominence can occur on

the first syllable in the target words, even in the absence of stress clash with a following word. And furthermore, such results would be consistent with proposals that the strength of the left-edge boundary influences the location of phrasal prominence.

Note that both conditions are clash-free. Whilst in the Isolated condition, the target is preceded and followed by silence, in the Embedded condition, the sequence of the two (pre-nuclear + nuclear) lexical items (e.g., the monosyllabic pre-nuclear word *Say* and the word carrying the nuclear prominence *Maltese*) would not present a clash configuration if the primary word-stress is on the 2nd syllable (e.g. *-tese* in *Maltese)*. In addition, on the assumption that stress shift directionality is leftward only, there is no 'accentable' syllable before *Say* as a potential docking site for stress shift.

4. Production materials and methods

4.1   *Speakers*   The present experiment used 12 native (Scottish, Southern British, and NE American) speakers of English (11 female and one male), all students or recent alumni of the University of Edinburgh, UK. One student was aware of the phenomenon under investigation and her data were discarded. Data from the remaining speakers were used in the analysis.

4.2   *Materials*   The experimental materials are the following 14 double-stressed target items (*antique, nineteen, Dundee, shampoo, courgette, campaign, canteen, Bombay, mundane, Mumbai, Chinese, Maltese, upstairs* and *Corfu*) and 26 tri-syllabic filler items carrying the main prominence on the middle syllable (e.g., *potato, fantastic, magenta*). Materials were presented in three lines on the screen. Target stimuli were presented at the centre of the screen with fillers on the top and bottom lines.

5. Experimental design and procedure

The target words elicited in this experiment were randomised within two experimental parts. The speakers read all the words in the Isolated condition first, then produced the material in the Embedded condition after a short break, as pragmatically it made more sense to produce all the *Say* [target] *again* items after producing the [target] first. They were asked to read the individual words (e.g., *banana, Chinese, Rebecca*) or sentences (e.g., *Say banana again*; *Say Chinese again*; *Say Rebecca again*) appearing on the three lines of the screen as if there were no relations between them. All the speakers read all the materials within each condition in a different random order. In the Embedded condition, the sentences were uttered with a pre-nuclear prominence on *Say* and a nuclear prominence on the target word. If the speakers put the main phrasal prominence on *Say* instead of the target, they were instructed to highlight the middle word in the sentence '*Say* [target] *again*', without pausing between the individual words. Sound recordings were made using a high-quality microphone, and digitised at a sampling rate of 48 KHz, at a bit depth of 16.

6. Perception method[1]

6.1   *Preparation*   The goal of the perception tests was to identify that location of the phrasal prominence location in the target utterances. The target utterances (i.e., the target word in the Isolated condition, and *Say* [target] *again* in the Embedded condition) were manually excerpted from the main recording with the preceding and following fillers in preparation for the perception tests. For the 11 speakers, one token for each of the 14 target items in each of the two conditions was extracted and carefully labelled. Because one token in the Isolated condition was missing from the recording, its mate in the Embedded condition was excluded from the perception test. A total of 306 tokens were therefore used to collect judgements about the location of phrasal prominence.

6.2   *Procedure*   A seven-step scale with options ranging from 'Early sure' to 'Late sure' was designed to collect the judgements of prominence location (cf. the top part of Table 2, below). The expectation was that the use of this fine-grained 7-step scale would produce a more precise categorisation of prominence

---

[1] Similar methods were used in the paper by the same authors to appear in the Proceedings of Speech Prosody 2022 (cf. Azzabou - Kacem & Turk, 2022).

location than, for example, a forced choice scale with only two options ('Early' or 'Late'). Three expert judges listened to and provided judgements for all the 306 tokens (3 judges * 306 tokens totalling 918 judgements).

6.3   *Categorisation of early prominence….*To determine whether 'Early' prominence was perceived, the next step was to identify the tokens paired by item and speaker across conditions (henceforth token-pairs), where at least two judges shared the same percepts of prominence location. Notwithstanding the precision of the fine-grained 7-step scale used for the collection of prominence location judgements, it may result in more chances for disagreement between judges, because of the large number of options (seven) for each token. To maximize the number of 'Early' vs 'Late' token-pairs used for the analysis, the results of the perceptual categories that were used in the original collection of data (i.e., the 7-step scale) were aggregated.

Table 0. *Seven-step judgement collection scale and methods for the categorisation of perceived early prominence*

Category	1	2	3	4	5	6	7
Original 7-step scale	Early Sure	Early not so sure	Both or early	Both sure	Both or late	Late not so sure	Late Sure
Relaxed 3-step scale	Early prominence			Both	Late prominence		
Dichotomous scale	Prominence pattern that is changed relative to the default 'Late'				Late prominence is the default		

Here, an important question came to the fore: Which categories of the 7-step scale should be considered as 'Early' categories? To answer this question, we need to understand what early prominence is. Would early prominence be any deviation of perceived prominence judgements from the 'default' or 'canonical' late pattern of the citation form referred to in dictionary entries? In this case, any judgement not including 'Late' would be considered as a changed pattern, and would by default be classified as 'Early', that is categories (1-4) in Table 2. Or, should early prominence be rather construed as a perceptual domain that includes any categories symmetrically equivalent to the 'Late' categorisation? We opted for the latter categorisation method using a **Relaxed 3-step scale.** The cut-off point was any judgement of 'Early' (categories 1-3), not including the 'Both Sure' category (category 4; cf. Table 2)**.**

6.4   *Variables….*Two types of perceptual results are presented. First, the overall *rates* of 'Early', 'Both' and 'Late' judgements are described for each judgment sample. The perceived Early prominence *rates* are the proportion of perceived early prominence in each experimental condition.

To test whether the overall rate of perceived 'Early' prominence is different across the Isolated vs. Embedded conditions, a prominence perception *score* consisting of the number of 'Early' prominence judgements (0-3) was calculated for each target token. The *scores* are also used as the dependent variable in the statistical tests.

7. Perception results

The perceptual results show excellent agreement between judges on the location of phrasal prominence. The judgements of prominence location were collected using the 7-step scale described in § 6.2 and categorised using the Relaxed 3-step scale method; cf. § 6.3). Of the 306 tokens used in the analysis, there were 294 tokens (96%) where at least two judges agree on prominence location, 215 tokens where all of the judges agree (70%), and 12 tokens with no agreement between judges on prominence location (4%). Across experimental conditions, of the 12 tokens where judges disagree, 11 were produced in the Isolated condition. For nine of the 'mates' of these 11 tokens produced in the other condition (i.e., the Embedded condition), all the judges agreed on the location of prominence.

In the following subsections, we present four sets of perceptual results. The first set shows the rates of perceived prominence for the tokens as paired across conditions by speaker and item (cf. § 7.1). The second set (shown in § 7.2) is about the perceived 'Early' prominence scores for every token, computed from the

responses of the three judges, and followed by the corresponding paired samples *t*-test. The last two sets are to test whether the preceding (two sets of) results are driven by particular judges (§ 7.3), speakers or items (§ 7.4). The Relaxed 3-step scale for the categorisation of judgements of early prominence (cf. § 6.3) was used in all sets of perceptual results.

7.1 *The paired results* This section presents the results of the perceptual judgements of the location of phrasal prominence for the tokens where at least 2 of the 3 judges agreed on the categorisation of prominence in both tokens of the target item (Isolated, Embedded) for each speaker. Figure 1, below, shows the distribution of judgements of prominence for tokens pairs e.g., (*Maltese*, *Say Maltese again*) within speaker. In the Embedded condition, 96% of the tokens (135 of 141) were reported to be perceived with 'Late' phrasal prominence (e.g., on –*tese* in *MalTESE*), while in the Isolated condition only 48% (68 of 141) were reported to be perceived as 'Late', and 5% (7 out 141) were judged to have both syllables of the target (*Mal-* and –*tese)* as equally prominent.
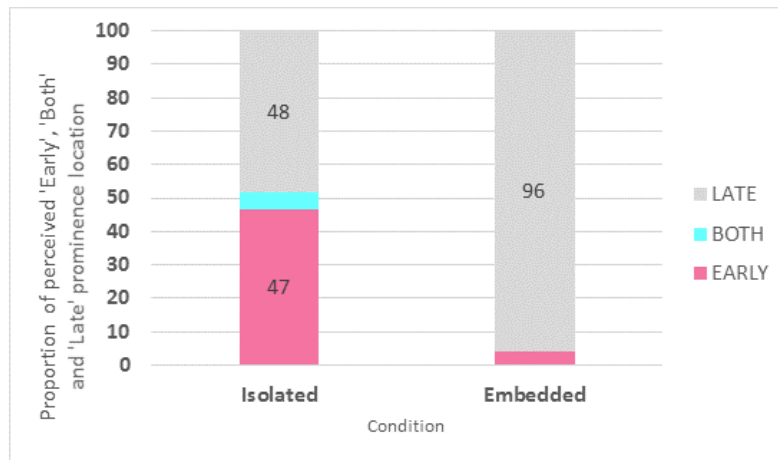

Figure 1. Distribution of the location of perceived prominence in the target, in the Isolated, e.g., *Maltese,* and Embedded condition, e.g., *Say Maltese again*, for the 282 token pairs where at least 2 judges agree.

7.2 *Perceived early prominence scores per token* The mean perceived prominence scores were the number of 'Early' judgements provided (for each target token) by each of the three judges (cf. § 4.3.6 in Chapter 4). The 'Early' scores were higher in the Isolated condition ($M = 0.54$; $SD = 1.05$) than those for the same speakers and items in the Embedded condition ($M = 0.17$; $SD = .055$); $t(152) = 4.237$; $p < .001$ (2-tailed).

To make sure that the above results are not restricted to a particular judge or a small subset of speakers or items, the following steps were taken: 1) the distributions of perceptual judgements were analysed, then the cross-tables for the distribution by speakers and items in Appendices 9.1 and 9.2 were examined.

7.3 *Distribution of prominence location per judge* Table 2, below, shows that all three judges perceived early prominence in the Isolated condition more often than in the Embedded condition, and all three judged that more Embedded items were late-prominent than early prominent.

Table 2. *Distribution of prominence location judgments by condition and judge using the Relaxed 3-step-scale prominence categorisation method (described in § 6.3).*

Condition	Location	Judge 1	Judge 2	Judge 3	At least 2 judges agree (tokens not paired across conditions)
Isolated	EARLY	52	86	63	67 (47%)
	BOTH	24	3	24	7 (5%)
	LATE	77	64	66	68 (48%)
Embedded	EARLY	5	11	10	7 (5%)
	BOTH	4	6	10	0 (0%)
	LATE	144	136	133	145 (95%)
	N	306	306	306	294

The question that is addressed in the next subsection is whether these results are driven by particular speakers or particular items.

7.4   *Distribution of prominence location by item and by speaker*   The distribution of prominence by speaker (cf. Appendix 1) and by item (cf. Appendix 2) used the tokens where at least 2 judges agree, not considering any pairing across conditions. Appendix 1 shows that the early prominence trend is well spread across speakers: 10 of the 11 speakers were judged to have produced early prominence in the Isolated condition. Speaker P02 did not show a pattern of early prominence that the judges agreed upon. Appendix 9.2 shows that all of the target items used in the experiment showed a proportion of early prominence that is at least two times higher in the Isolated condition compared to the Embedded condition. The results of these analyses suggest that the cases of early prominence in the Isolated condition are not specific to particular lexical items or speakers.

8   Discussion

The results of the perceptual judgements show a very clear tendency towards late prominence in the Embedded condition (e.g., on *–tese in MalTESE*), while in the Isolated condition less than half of the targets were judged late-prominent[1]. Thus, the experimental manipulation of the prosodic context of the target via embedding affected the location of perceived phrasal prominence as predicted. The stronger left-edge boundary and the absence of a preceding word increased the perceived incidence of early prominence in the Isolated condition, as compared to the Embedded condition. These results indicate that early phrasal prominence can be heard in the absence of stress clash with a following stronger prominence. In addition, they are consistent with the findings from previous literature showing that early prominence, in general, can function as a domain-initial marker (Shattuck-Hufnagel et al., 1994; Astésano et al., 2007; Garnier et al., 2016).

A potential alternative explanation could be found if we consider, contrary to what is assumed in the literature, the idea that stress shift could be bi-directional in English (both leftward and rightward movement allowed), and if we consider the possibility that the target words in our experiment might have underlying initial main stress (instead of final main stress as is widely assumed). On this view, *Say* and *Mal-* in *Maltese* may have clashed in the Embedded condition, and clash may have been alleviated by moving the prominence from *Mal-* to *-tese*. However, we note that roughly half of our materials showed early prominence in the Isolated condition, with half showing prominence on the 2nd syllable. If the alternative explanation is correct, we would have expected more cases of early prominence in the Isolated condition. We therefore prefer the early prominence-as-domain-initial-marker account, and speculate that the right-edge boundary may have affected the rates of perceived early prominence since our materials were both IP-initial and IP-final. .

The results also present evidence that the location of the nuclear prominence can be assigned to either syllable of the disyllabic target (e.g., *thirteen*) when it is produced in isolation. This result indicates that

---

[1] We note that our experiment does not distinguish between effects of prosodic boundaries on the speaker, the listener, or both.

English stress may actually not be strictly culminative, as any of the two foot-level prominences of the double-stressed word can have equal chances to carry the main lexical prominence.

   Overall, the results are consistent with the view that doubly stressed words can show stress shifting to demarcate the left IP edge. Because this study uses contexts completely free of stress clash with a following word (Selkirk, 1984), its results provide evidence that other factors, such as the marking of domain onsets (Halle & Vergnaud, 1987; Halle & Idsardi, 1995) and/or the absence of a *preceding* word, can influence the movement of prominence location.

9 Appendices

9.1 *Appendix 1* Distribution of prominence location by speaker in the paired items where any two judges agreed on prominence location

Speaker			EARLY	Both	LATE	Total
P01	Condition	Isolated	10	0	4	14
		Embedded	0	0	14	14
	Total		10	0	18	28
P02	Condition	Isolated	0	0	12	12
		Embedded	0	0	12	12
	Total		0	0	24	24
P03	Condition	Isolated	3	0	9	12
		Embedded	0	0	12	12
	Total		3	0	21	24
P04	Condition	Isolated	1	1	12	14
		Embedded	0	0	14	14
	Total		1	1	26	28
P05	Condition	Isolated	8	0	6	14
		Embedded	0	0	14	14
	Total		8	0	20	28
P06	Condition	Isolated	6	1	2	9
		Embedded	1	0	8	9
	Total		7	1	10	18
P07	Condition	Isolated	8	0	5	13
		Embedded	0	0	13	13
	Total		8	0	18	26
P08	Condition	Isolated	11	1	2	14
		Embedded	0	0	14	14
	Total		11	1	16	28
P09	Condition	Isolated	3	1	10	14
		Embedded	0	0	14	14
	Total		3	1	24	28
P10	Condition	Isolated	12	1	0	13
		Embedded	0	0	13	13
	Total		12	1	13	26
P11	Condition	Isolated	4	2	6	12
		Embedded	5	0	7	12
	Total		9	2	13	24
Total	Condition	Isolated	66	7	68	141
		Embedded	6	0	135	141
	Total		72	7	203	282

9.2 *Appendix 2*   Distribution of prominence location by item

Item			EARLY	Both	LATE	
antique	Condition	Isolated	4	0	7	11
		Embedded	0	0	11	11
	Total		4	0	18	22
Bombay	Condition	Isolated	4	1	4	9
		Embedded	0	0	9	9
	Total		4	1	13	18
campaign	Condition	Isolated	7	0	4	11
		Embedded	0	0	11	11
	Total		7	0	15	22
canteen	Condition	Isolated	6	1	3	10
		Embedded	1	0	9	10
	Total		7	1	12	20
Chinese	Condition	Isolated	7	1	3	11
		Embedded	0	0	11	11
	Total		7	1	14	22
Corfu	Condition	Isolated	6	0	5	11
		Embedded	1	0	10	11
	Total		7	0	15	22
courgette	Condition	Isolated	2	0	7	9
		Embedded	0	0	9	9
	Total		2	0	16	18
Dundee	Condition	Isolated	5	1	4	10
		Embedded	2	0	8	10
	Total		7	1	12	20
Maltese	Condition	Isolated	3	1	6	10
		Embedded	0	0	10	10
	Total		3	1	16	20
Mumbai	Condition	Isolated	2	0	7	9
		Embedded	1	0	8	9
	Total		3	0	15	18
mundane	Condition	Isolated	2	2	5	9
		Embedded	0	0	9	9
	Total		2	2	14	18
nineteen	Condition	Isolated	6	0	4	10
		Embedded	0	0	10	10
	Total		6	0	14	20
shampoo	Condition	Isolated	6	0	4	10
		Embedded	0	0	10	10
	Total		6	0	14	20

Upstairs	Condition	Isolated	6	0	5	11
		Embedded	1	0	10	11
	Total		7	0	15	22
Total	Condition	Isolated	66	7	68	141
		Embedded	6	0	135	141
	Total		72	7	203	282

References

Astésano, Corine Bard, Ellen Gurman & Turk, Alice (2007). Structural Influences on Initial Accent Placement in French. *Language and Speech, 50* (3), 423-446.

Azzabou - Kacem, Soundess & Turk, Alice (2022). Fine gradations of prosodic boundary strength can drive the assignment of prominence. In Frota, Sonia. & Vigário, Marina (eds.), *Proceedings of Speech Prosody 2022* (23-26 May 2022). Lisbon, Portugal.

Bolinger, Dwight (1958). A theory of pitch accents in English. *Word, 14*, 109-149.

Cooper, William, E. & Eady, Stephen J. (1986). Metrical phonology in speech production. *Journal of Memory and Language, 25*, 369-384.

Garnier, Laury, Baqué ; Lorraine, Dagnac, Anne & Astésano, Corine (2016). Perceptual investigation of prosodic phrasing in French. *Proceedings of Speech Prosody 2016*, 1153-7, Boston, MA, USA.

Grabe, Esther & Warren, Paul (1994). Stress Shift: Do speakers do it or listeners hear it? In B. Connell (Ed.), *Laboratory Phonology IV: Phonology and Phonetic Evidence* (pp. 95-110). Oxford: Oxford University Press.

Halle, Morris & Idsardi, William (1995). General Properties of Stress and Metrical Structure. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 403 - 443). Oxford: Blackwells.

Halle, Morris & Vergnaud, Jean-Roger (1987). Stress and the Cycle. *Linguistic* Inquiry, *18*(1), 45-84.

Hayes, Bruce (1995). *Metrical stress theory : Principles and case studies*. Chicago: University of Chicago Press.

Hayes, Bruce & Puppel, Stanislaw (1984). On the rhythm rule in Polish. In H. van der Hulst, & N. Smith (Eds.), *Advances in Nonlinear Phonology* (pp. 59-81). Dordrecht: Foris.

Liberman, Mark, & Prince, Alan (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8*, 249-336.

Selkirk, Elisabeth O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass: MIT Press.

Shattuck-Hufnagel, Stefanie, Ostendorf, Mari, & Ross, Kenneth N. (1994). Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics, 22*, 357-388.

Thompson, Henry Swift. (1980). *Stress and Salience in English: Theory and Practice*. Palo Alto, California: Xerox Palo Alto Research Center.