A Phonetic and Phonological Analysis of Dual and Multiple Focuses in Standard Chinese

Yuan Jia[1], Aijun Li[1], Ziyu Xiong[1]

[1] Phonetics Lab, Institute of Linguistics, Chinese Academy of Social Sciences

`jiayuan,liaj,xiongziy@cass.org.cn`

Abstract

The present study investigates the acoustic manifestation and phonological nature of prominences induced by dual and multiple focuses in Standard Chinese (hereinafter SC). Results show that double focus exerted two prominences in the target sentence and these two prominences performed similar acoustic manifestations. However, under a multiple focus condition, only the rightmost focused constituent realized prominence in the sentence. Based on this evidence, this study proposed that the prominence pattern in SC have hierarchical levels, and this phenomenon can be accounted for by the distinction of *nuclear accent* and *pre-nuclear accent* as in English. Specifically, in SC, the *nuclear prominence* is characterized by obligatory and unique nature, while the *pre-nuclear prominence* is optional and secondary. There is no restriction on the appearance of *nuclear prominence*, while the appearance of *pre-nuclear prominence* is constrained by focus condition.

Index Terms: dual focus, multiple focus, nuclear prominence, pre-nuclear prominence.

1. Introduction

There is almost as much variation in the notion of focus as there are different ways of approaching focus phenomena. This fact, which has been repeatedly pointed out in the literature (Szabolcsi [1], King [2], Kiss [3], and Winkler [4]), is investigated in the fields of semantics, syntax, phonology, phonetics, or pragmatics of focus, which makes it difficult to compare. In the prosodic literature, two major types of focus are usually distinguished: *broad* or *wide focus*, and *narrow focus*. Broad focus is often referred to as (new-)information focus or presentational focus, and narrow focus is often referred to as identificational or contrastive focus (e.g., King [2] and Kiss [3]). Based on this definition, previous studies have approached focus phenomena in two major ways: (i) the phonetic approach, and (ii) the phonological approach. In the former, instrumental and experimental means are adopted to look for acoustic differences between broad and narrow focuses. This difference is usually displayed in terms of fundamental frequency (F_0) and sometimes also duration and intensity. Thus, the phonetic approach concentrates on the detailed descriptions of physical continua, not on the phonological structure. In contrast, in the phonological approach, the emphasis is on the non-lexical phonological structure of the focused constituents, and issues like accent distribution and prosodic phrasing.

In the phonetic approach, cross-language studies have demonstrated that larger F_0 excursions and durational lengthening are important acoustic correlates for the focused constituents, specifically, English, Dutch, Danish and Spanish are illustrative examples, as shown in Xu [5], Nooteboom & Kruyt [6] and Thorsen [7], Toledo [8], respectively. These authors emphasize the effect of F_0 in focus signaling, according to Xu [5], a narrow focus is realized by expanding the pitch range of the on-focus stressed syllables and suppressing the pitch range of the post-focus syllables in English. Nooteboom and Kruyt [6] propose that the F_0 overrides the syntactic devices (e.g. word order) as a cue to the "given/new" distinction in Dutch. In Standard Danish, Thorsen [7] points out that although F_0 is the prime cue to "focus" and "emphasis for contrast" as in English or Dutch. The specific F_0 details are different in this language: it is the F_0 reduction of surrounding rise-falls or the lifting out of the focused item, and not the presence of a more elaborate F_0 movement on the focused item. Toledo [8] maintains that features of pitch, intensity, and duration can be taken to stand for focus in Spanish as in English. However, Toledo concludes that although pitch features like peak F_0 and pitch range may cue focus, they show less consistent performances than in English. On the whole, loudness peak seems to be the best correlate of focus in the speech data observed.

With regard to the phonological approach, it mainly discusses two types of effects induced by focus: (i) stress and accent effect; and (ii) phrasing effect[1]. The first type of effect is clearly showed in the discussion of the relation between focus and nucleus placement in English and Dutch by Ladd [9] and Gussenhoven [10]. To this relation, Ladd adds the role played by abstract prominence patterns: the nucleus signals the focus of the sentence, and the nucleus is assigned to the element that bears the sentence accent. In English, the final prominence pattern (w(eak)-s(trong)) signals broad focus, whereas early prominence (s-w) usually conveys narrow focus. Moreover, they propose that the statements of the location of focus and accents involve two aspects, focus distribution is determined by contextual factors and accent placement is restricted by the language structural rules, e.g., Sentence Accent Assignment Rule (SAAR) proposed by Gussenhoven [11].

Following the path suggested by the previous analysis, the present study mainly investigates the prosodic effects of dual and multiple focuses in SC. Its aim is threefold: (i) the acoustic manifestations of double and multiple focuses; (ii) the hierarchical levels of prominences conveyed by the focus in discussion; and (iii) the theoretical explanation of the mapping relation between *focus* and *prominence* in SC. It is expected that the present examination of focus may provide important evidences for the prosodic effects of focus in cross-language studies. In order to accomplish this aim, this study employs acoustic experiment and (a version of) the theory of *Intonational Phonology*, as developed by Ladd [9] and Gussenhoven [10]. The chosen approach here is simultaneously theoretically and empirically based, in which experimentally collected speech data is employed to investigate questions about the abstract categories of phonological structure.

---

[1] The present mainly discusses the relation between focus and prominence distribution and detailed introduction of phrasing effect of focus was not included.

2. Methodology

An acoustic experiment is adopted to investigate the acoustic manifestations of prominences and corresponding relation between (dual or multiple) focus placement and prominence distribution in SC. As is mentioned in the previous study that F_0 is the primary cue to signal focus, the present study also employed F_0 as the parameter to examine the location and realizing manner of prominences in SC.

2.1. Materials selection

With regard to the research aim of the acoustic experiment, the following factors were considered in the material design: tonal and segmental combinations of focused constituents and syntactic structure of the sample sentences. Previous studies on the focus realization in SC demonstrated that tonal combinations of the focused constituents exert no effect on the distribution of the prominence (refer to, Xu [12]), thus the present study only selected the tone2 components as the target constituents to examine the performances of the 'L-H' tonal permutation. In order to offset the effect of the segments on the F_0, the study adopted the voiced initials to compose the target words, specifically, they are: S={Liu2 Min2(Liumin)}[2]; Adv={Ling2 Chen2(early morning)}; V={Ti2 Ba2(promote)}; O={Mao2 Lan2(Maolan)}[3]. As for the formation of the target sentence, Xu [13] points out that the canonical form of the Chinese sentence is in the order of 'SVO', though 'SOV' is also found as an alternative. Based on this observation, the target sentence is composed as the following (i):

(i) Liu Min2 Ling2 Chen2 Ti2 Ba2 Mao2 Lan2 Le0.

 liu min early morning elevate mao lan le

 (Liumin elevated Maolan in the early morning).

To approach the intended focus conditions, the study followed Zubizarreta [14][4] to adopt double and multiple *wh*-opertors. The guide sentences and the focus conditions are displayed in the following examples (ii)-(vii):

(ii) Fa1 Sheng1 Le0 Shen2 Me0 Shi?

 happen le what

 (What happened?)

Liu Min2 Ling2 Chen2 Ti2 Ba2 Mao2 Lan2 Le0[+BF] .

(iii) Shei2 Ling2 Chen2 Ti2 Ba2 Shei2 Le0?

 who early morning elevate whom le

 (Who elevated whom in the early morning?)

Liu Min2[+NF][5] Ling2 Chen2 Ti2 Ba2 Mao2 Lan2 [+NF] Le0.

(iv) Shei2 Shen2 Me0 Shi2 Hou0 Ti2 Ba2 Shei2 Le0?

 who when elevate whom

 (Who and when elevated whom?)

Liu2Min2[+NF] Ling2Chen2[+NF] Ti2 Ba2 Mao2 Lan2 [+NF] Le0.

(v) Shei2 Ling2 Chen2 Zen3 Me0 Shei2 Le0

 who early morning do what whom

 (Who do what to whom in the early morning?)

Liu2Min2[+NF]Ling2Chen2 Ti2 Ba2[+NF] Mao2 Lan2[+NF] Le0.

(vi) Liu2 Min2 Shen2 Me0 Shi2 Hou0 Zen3 Me0 Shei2 Le0?

 liu min when   do what whom

 (When and where Liumin do what to whom?)

Liu2Min2 Ling2Chen2[+NF] Ti2 Ba2[+NF] Mao2 Lan2 [+NF] Le0.

(vii) Shei2 Shen2 Me0 Shi2 Hou0 Zen3 Me0 Mao2 Lan2 Le0?

 who when do what mao lan

 (Who and when do what to Maolan?)

Liu2Min2[+NF]Ling2Chen2[+NF] Ti2 Ba2[+NF] Mao2 Lan2 Le0.

---

[2] Liumin and Maolan are Chinese names.
[3] The symbols of 'S', 'Adv', 'V' and 'O' stand for the Subject, Adverb, Verb and Object, respectively.
[4] Zubizarreta [14] adopted double and multiple *wh*-operators to investigate the relationship between focus and prominence in English and German.
[5] The features of [+BF] and [+NF] are adopted to stand for the *broad focus* and *narrow focus*.

2.2. Recording

All the above mentioned sentences were contained in the recording schema with two times repetitions for each sentence. Eight Standard Chinese speakers, four females and four males, aged within 20-45, were recruited as the subjects. These subjects were divided into four groups, each contains two women or two men. The recording was conducted in the sound-proof booth in the Phonetics Lab (CASS)[6]. During the recording procedure, each *wh*-question and target sentence pair appeared on the screen in random order. Within each group, one subject was asked to read the *wh*-questions and the other read the target sentences as the answer to the question in normal speed without any irregular pause. The speakers were instructed to read the sentences as naturally as possible according to the given context, and were free to repeat them in case they considered their reading not fluent or unnatural. When the schema was finished recording, the subjects were asked to change the *asking-answe*r role. We got a total of 32 samples for each target sentence.

2.3. Data labeling and extraction

The syllabic boundaries of all the 'wav' files were annotated in the utterance and they were double checked to ensure the accuracy of the data. The F_0 data extraction was conducted based on the 'PitchTier' files with each syllable being selected ten points. SPSS 10 was employed to get the F_0 means, and One-Way ANOVA was conducted to test the significance of the differences of the highest and lowest points of the F_0 values for the words in various focus conditions.

3. Phonetic and phonological analysis of dual and multiple focuses

This part employs F_0 as the parameter and examines the acoustic manifestations and phonological nature of the double and multiple focuses in SC. Its aim lies in: (i) to explore the corresponding relation between *focus* and *prominence*; (ii) to investigate the hierarchical levels of prominences in SC.

3.1. Dual focuses

This part mainly discusses the distribution and the realizing manner of prominences induced by double focuses. The (i), (ii) and (iii) are adopted in this part.

The following Figure1 is the F_0 means of the sentence "LiuMin2 Ling2Chen2 Ti2Ba2 Mao2Lan2 Le0(Liumin elevated Maolan in the early morning)". Within this figure, the bottom part of the X-Axis illustrates the focus environment of the sentence, specifically, 'BF' stands for 'Broad Focus' and "DN-S-O" indicates that the sentence contains double narrow focuses which distribute on subject and object items. Further, the top part of the X-Axis displays the content of each syllable in the utterance. And the Y-Axis states the pitch range of the sentence which is fixed as 110Hz-260Hz in accordance with the pitch range of all the speakers in this study.
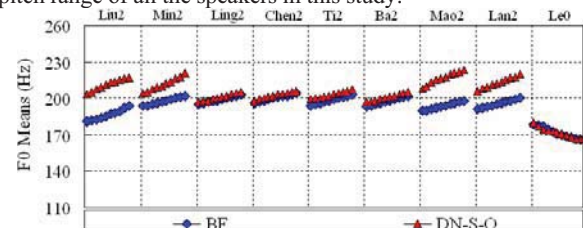

Figure 1: *F_0 Means of broad and double focuses sentences*

---

[6] Chinese Academy of Social Sciences.

As for the F_0 prominences of dual focuses in English, Eady et al [15] propose that regardless of whether the sentence contains one or two focuses, the F_0 peaks and word durations can always signal focus. The main difference between the single focus and dual focus is that the F_0 between the two focuses exhibit no lowering. The present figure of the double focus in SC displays a similar F_0 prominence distribution with English. Specifically, the prominence distribution corresponds with the focused constituents and also no F_0 compression of the constituents between the two focuses. A closer examination of the two prominences reveals that they perform similar realization, i.e., both the *L* and *H* tone of the two syllables are raised to some extent in comparison with the broad focus sentence. One-Way ANOVA was used to compare the significance of differences of the F_0 of the words in the positions of subject, adverb, verb and object in the contours of "BF" and "DN-S-O". Bonferroni post hoc test was used to examine the differences of the minimum and maximum pitch values of the syntactic component in different focus conditions. Results have demonstrated that the lowest points of the subject and object constituents in double focus sentences are significantly different from the broad focus: under the subject position, [By Narrow Focus: $F_{min}(1, 32)=17.24$, P=0.00 and F_{max}[7]$(1, 32)=37.82$, P=0.00]; under object position, [By Narrow Focus: $F_{min}(1, 32)=18.74$, P=0.00 and $F_{max}(1, 32)=38.29$, P=0.00]. However, the subject and object constituents are not obviously different from each other in pitch values with $P_{min}>0.05$ and $P_{max}>0.05$. Further, the components in the positions of adverb and verb under narrow focus exhibit that they are also similar with the units in broad focus condition. As it is known, F_0 goes slightly downward in general condition; however, the second prominence in Figure 1 does not show any lowering in comparison with the first prominence on the subject position in the sentence. And, this kind of phenomenon displays cross-language feature, i.e., Jaeggli [16] investigates the prominence patterns within one phrase in Romance and he proposes that although the first focused constituent begins with high pitch, the main prominence falls at the end of the phrase. Thereafter, we propose that the right prominence in the sentence is the main prominence and the left is the secondary one in the sentence.

3.2. Multiple focuses

Part 3.1 discusses the co-existence of the prominences induced by dual focuses. Due to the existence of the declination phenomena and the non-lowering of the second prominence, it points out that the rightmost one is the main prominence. In this part, evidences are provided from the multiple focuses to further examine the phonetic details of prominences and (i), (iv), (v), (vi), (vii) sentences are selected.

In Figure 2 the focuses are dwelled on three non-adjacent syntactic components, i.e., subject, adverb and object constituents illustrated by the symbol "NF-S-Ad-O"; subject, verb and object components as described by the sign "NF-S-V-O". The content of the X-Axis and Y-Axis are identical with Figure 1.
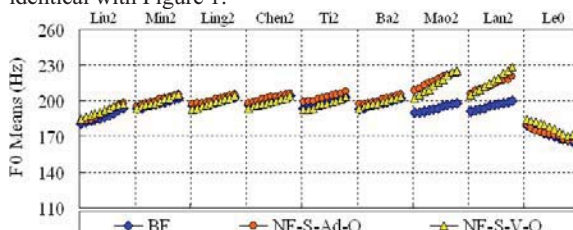

Figure 2: *F_0 Means of broad focus and multiple focuses*

In comparison with contour 'BF', it is obvious in the contour 'NF-S-Ad-O' that the F_0 prominence locates on the object constituent which exhibits higher pitch register. Both the *L* tone and *H* tone behave higher pitch than the same item in broad focus condition. Further, the pitch register of the object is also higher than the other syntactic components, i.e., subject, adverb and verb within the same sentence. Results of Bonferroni post hoc test show that the minimum and maximum pitch values of the object constituent in 'NF-S-Ad-O' significantly differ from the one in 'BF': [By Narrow Focus: $F_{min}(1, 32)=16.24$, P=0.01 and $F_{max}(1, 32)=39.01$, P=0.00]. Pitch values of other constituents in the sentences are not different from the ones in the 'BF' sentence. From these results, it can be seen that although there are three focuses in one sentence, i.e., subject, adverb and object, only the object (the rightmost unit) exerts F_0 prominence. As for the contour 'NF-S-V-O', the subject and the adverb and verb constituents perform resemble pitch registers, and the object item exhibits higher pitch register, with both of the *L* and *H* tones of the word "mao2lan2" displaying higher than the other syntactic component. Results of Bonferroni post hoc test show that the pitch values of the object constituents are different from those in broad focus, e.g. [By Narrow Focus: $F_{min}(1, 32)=14.32$, P=0.012 and $F_{max}(1, 32)=39.35$, P=0.00]. Moreover, the prominence in 'NF-S-Ad-O' resembles with the one in the contour 'NF-S-V-O', although these two contours deserve different distribution of focus, they display almost identical performance of prominence. Therefore, when multiple constituents in the sentence are served as the focuses, they can not realize prominence simultaneously. The rightmost component is the only anchor for the F_0 prominence.

In Figure 3, the focus is designed to distribute on three adjacent constituents, e.g., adverb, verb and object constituents as marked by the symbol 'NF-Ad-V-O', and subject, adverb and verb constituents as described in 'NF-S-Ad-V'. In this manner, we can examine the specific phonetic realization of the prominence induced by three adjacent focused constituents.
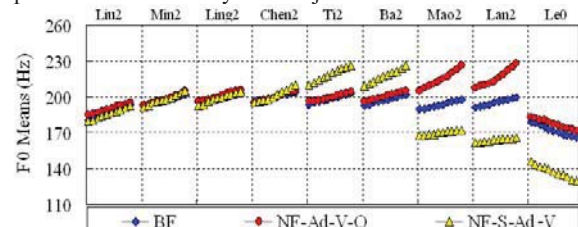

Figure 3: *F_0 Means of broad focus and multiple focuses*

Within the contour 'NF-Ad-V-O', although three focuses are put onto the adverb, verb and object constituents, the prominence only locates on the object item, specifically, it distributes on the rightmost focus bearing unit. The specific phonetic realization of the prominence is due to the raising of the *L* tone and *H* tone of the two syllables. In comparison with the 'BF' contour, the lowest and highest points of the object in 'NF-Ad-V-O' is significantly different from the one in 'BF' contour. Results of Bonferroni post hoc test support this observation, i.e., [By Narrow Focus: $F_{min}(1, 32)=16.32$, P=0.00 and $F_{max}(1, 32)=39.78$, P=0.00]. These numbers further suggest that there is intonational prominence on the rightmost\focus bearing unit. As for contour 'NF-S-Ad-V', it illustrates in the above figure that the verb constituent "ti2ba2" shows a higher pitch register than the subject, adverb and the object constituents in the same sentence. More specifically, both the *L* tone and the *H* tone of the verb constituents are raised by the narrow focus. However, the focus bearing units of subject and the adverb constituents show no obvious effect upon the pitch raising. In contrast to the former constituents, the object item gets pitch register lowering. Results

---

[7] 'min' and 'max' are taken to mean the minimum and maximum values.

of the Bonferroni post hoc test have shown that the pitch values of the verb constituent is significantly different from the one in 'BF' condition, e.g., [By Narrow Focus: $F_{min}(1, 32)=14.19$, P=0.01 and $F_{max}(1, 32)=36.44$, P=0.00]. All of these results demonstrate that the verb constituent is the prominence bearing unit with the post-constituent being compressed significantly. The realization of the prominence in the surface form is triggered by the rightmost focus. Hereinafter, the evidences of the adjacent focused constituents further suggest that the rightmost unit is the prominence bearing unit in the sentence.

4. Discussion and conclusion

The present study systematically investigates the distribution of the prominences triggered by dual and multiple focuses in SC. Taking the F_0 as the parameter, results have shown that under a dual focus condition, the two focuses can realize two prominences simultaneously; however, there exists level differences between these two prominences, i.e., the right one bears the primary nature and the left one the secondary. Under a multiple focus condition, the F_0 evidences suggest that there is only one prominence in the target sentence and only the rightmost focused constituent can serve as the anchor for the prominence.

The dual and multiple focuses induced F_0 prominence in SC obtained in this study have two implications: (i) the rightmost focused constituent can always realize prominence in the surface form; and (ii) although two prominences can co-exist with each other, it has hierarchical distinctions. Therefore, this study suggests that level difference in SC can be related to the classification of *nuclear accent* and *pre-nuclear accent* in the intonation language as English and Dutch, which states that the nuclear accent is both obligatory and potentially unique and the pre-nuclear accent is of optional and secondary nature (refer to, Ladd [9], Gussenhoven [17]).

Evidence of acoustic manifestations of dual and multiple focuses show that SC obtains the phonological entities of nuclear accent and pre-nuclear accent as in English. However, SC is a tonal language, the acoustic reflection of the focus is the F_0 prominence, and it is realized above the original tonal target *L* or *H*. What constitutes an interesting phenomenon in the above analysis is that identical level of focus can not lead to the same level of prominence, only the double focuses can trigger double prominences and the multiple focuses can only realize one prominence on the rightmost constituent. This evidence suggests that the *nuclear prominence* bears the unique and obligatory nature and the *pre-nuclear prominence* has the optional and secondary nature in SC. The evidence of the multiple focuses further demonstrates that the appearance of nuclear prominence is unmarked. As for the underlying causes for restricting the distribution of the prominence, it is the metrical structure of SC. In SC, the rightmost position is the metrically strongest position. Under identical focus condition, the metrically strongest positions deserve the nuclear prominence.

On the whole, *focus distribution* and *prominence distribution* exhibits non-symmetrical relation in SC. Similar cases can also be observed in English, it is pointed out that there remain a number of disagreements about how *focus* is conveyed by *accent*. In many ways these disagreements represent the continuation of a decades-old debate about the relation between focus and accent. In the literature, there are two major ways of dealing with the relationship between *focus* and *accent*, i.e., the structure-based approach and highlighting-based approach. The first approach was sketched by Ladd [9] and Gussenhoven [11], this view deals with the relation between focus and accent by distinguishing the *distribution of focus* and the *distribution of accent*. That is, the linguistic description of accent patterns involves two complementary but essentially separate aspects: a statement about which parts of an utterance are focused, and a statement about how a given pattern of focus is conveyed by the location of the accent. However, within the highlighting-based approach (refer to, Bolinger [18]), it rejects the distribution between focus distribution and accent distribution and maintains the validity of the bi-directional relation between focus and accent, thus, in this view, if a word is focused, it is accented. The evidence of the asymmetric relation between focused parts and prominent parts in SC have supported the observations made by the structure-based approach that the speaker's decision about what to focus is subject to all contextual influences. However, once the focused part of the utterance is specified, the accent pattern follows more or less automatically by language-specific rules.

5. Acknowledgements

6. References

[1] Szabolcsi, Anna. "The semantics of topic-focus articulation. Formal Methods in the Study of Language", 513-540. Amsterdam: Mathematische Centrum, 1981.

[2] King, Patricia H., "Configuring Topic and Focus in Russian", Stanford: Center for the Study of Language and Information Publications, 1995.

[3] Kiss, Katalin., "The focus operator and information focus", Working Papers in the Theory of Grammar, Vol. 3, N2, Budapest, 1996.

[4] Winkler, Suanne., "Focus and Secondary Predication", Berlin: Mouton, 1997.

[5] Xu Yi and Ching X. Xu., "Phonetic realization of focus in English declarative intonation", Journal of Phonetics, 33: 159-197, 2005.

[6] Nooteboom, Sieb and J. Kruyt., "Accents, focus distribution, and perceived distribution of given and new information: An experiment", Journal of the Acoustical Society of America 74: 1512-1524, 1987.

[7] Thorsen Nina., "Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish", Proceedings of the IXth International Congress of Phonetic Sciences. Volum II The Relation between Sentence Prosody and Word Prosody, 417-423, 1975.

[8] Toledo, Guillermo., "Señales Prosodicas del Foco", Revista Argentina de Lingüistica 5(1-2): 205-230, 1989.

[9] Ladd, D. Robert., "Intonational Phonology", Cambridge: Cambridge University Press, 1996.

[10] Gussenhoven Carlos., "Focus and Sentence Accents in English. Focus and Natural Language Processing", Working Papers of the Institute for Logic and Linguistics. Heidelberg, 83-92, 1994.

[11] Gussenhoven Carlos., "Focus, mode and nucleus", Journal of Linguistics 19: 377-417, 1983.

[12] Xu Yi., "Effects of tone and focus on the formation and alignment of F_0 contours", Journal of Phonetics, 27, 55-105, 1999.

[13] Xu, Liejiong., "Manifestation of informational focus", Lingua 114: 277-299, 2004.

[14] Zubizarreta, Maria Luisa., "Prosody, Focus, and Word Order"., London: Massachusetts Institute of Technology Press,1998.

[15] Eady Stephen, et al., "Acoustical characteristics of sentential focus: narrow vs. broad and single vs. dual focus environments", Language and Speech 29: 233-251, 1986.

[16] Jaeggli, O., "Topics in Romance syntax. Dordrecht: Foris", 1982.

[17] Gussenhoven Carlos., "The Phonology of Tone and Intonation"., Cambridge: Cambridge University Press, 2004.

[18] Bolinger, Dwight., "Accent is predictable (if you're a mind-reader)", Language 48: 633-644, 1972.

[This paper was published in Speech Prosody, 2010]