

Disambiguate or not? – The role of prosody in unambiguous and potentially ambiguous anaphora production in strictly Mandarin parallel structures

Luying Hou¹, Bert Le Bruyn¹, René Kager¹

¹Utrecht University

L.Hou@uu.nl, B.S.W.LeBruyn@uu.nl, R.W.J.Kager@uu.nl

Abstract

It has been observed that the interpretation of pronouns can depend on their accentuation patterns in parallel sentences as "John hit Bill and then George hit him", in which 'him' refers to Bill when unaccented but shifts to John when accented. While accentuation is widely regarded as a means of disambiguation, some studies have noticed that it also extends to unambiguous anaphors [7-10]. From the perspective of production, however, no strong experimental confirmation was found for the 'shift' function of accented pronouns, which is due to the fact that production research has mainly focused on corpora [5, 6]. Hence, the nature of the accent on anaphors still remains obscure. By manipulating referential shift and ambiguity, this study explores the role of prosody in anaphora production in strictly Mandarin parallel structures. The results reveal a significantly higher F₀ and longer duration for anaphors in referentially shifted conditions, suggesting that anaphoric accentuation signals a referential change in strictly parallel structures in Mandarin. No evidence was found that ambiguity plays a role in anaphoric accentuation. This finding challenges the general view on accented pronouns and will deepen our understanding on semantics-prosody relationship.

Index Terms: pronoun, anaphora, production, ambiguity, accent, prominence, \mathbf{F}_0

1. Introduction

1.1. Background

It has been noted that the interpretation of pronouns may depend crucially on prosody in spoken language [1]. As illustrated in (1) – a coordinated sentence with a parallel structure, the potentially ambiguous pronoun 'him' has different antecedents under different accentual patterns (capitalization indicates accentuation): it refers to 'Bill' when unaccented but to 'John' when accented.

(1) John hit Bill and then George hit him (Bill). John hit Bill and then George hit HIM (John).

Studies on pronoun resolution propose that listeners employ a *parallel function strategy*: an unaccented pronoun co-refers with a prior constituent in the same syntactic position, while an accented pronoun shifts this preference to the competitor in a different position [2-4]. Since listeners are unable to resolve the correct antecedent of a pronoun in sentences like (1) via cues other than prosody, pronominal accentuation is pervasively assumed to be a disambiguation strategy [5-12]. Different from this ambiguity-oriented view, some studies point out from a generalized perspective that accentuation also extends to unambiguous anaphors [13-15],

as illustrated in Lakoff's example (2) [16] with unambiguous pronouns and (3) with full NP. Moreover, it is argued that disambiguating information cannot surpass prosodic cue in resolution [2, 14], as shown in (4), in which the unaccented 'he' cannot be identified with 'Bill' and listeners are forced to associate 'he' with 'Mary'. These examples show that the connection between prosody and ambiguous pronouns may be an instance of a broader relationship between prosody and referential shift, which is defined as the anaphor and its antecedent being in different syntactic positions.

- (2) John insulted Mary and then SHE insulted HIM.
- (3) John hit Bill and then George hit JOHN.
- (4) Mary hit Bill, and then he hit HARRY.

Though *referential shift* is a more plausible trigger of accentuation on pronouns than *ambiguity*, the role and the exact nature of anaphoric accentuation in parallel structures still remain obscure.

On the one hand, the possibility has not been ruled out that ambiguity jointly plays a role with referential shift. Moreover, the disambiguation effect of prosody seems so self-evident that accented unambiguous anaphors has been much ignored, and many studies have modified Lakoff's example by only using ambiguous pronouns, which runs the risk of establishing and reinforcing a causal relationship between ambiguity and accentuation. Hence, experimental work is needed to assess the respective contributions of referential shift and ambiguity. To our knowledge, ambiguity has never been investigated as a potential factor in anaphora production.

On the other hand, though intuitively convincing, experimental evidence is still lacking relating accentuation and referential shift. Most experimental studies on production have failed to find a distinct 'shift' function of accented pronouns [9, 17]. A major reason is that almost all those studies used corpus materials, in which there may be limited cases of referential shift and there exist many uncontrolled confounding factors. Studies on perception have demonstrated various strategies and factors in pronoun resolution. For instance, some evidence indicates that a subject assignment preference is the dominant strategy other than parallel function strategy [18, 19]. Smyth's result [2] shows that lack of control over parallelism accounts for the conflicting results. Besides sentence structure, there are other factors involved such as discourse coherence [8, 20, 21] and causality [22] introduced by the verb or conjunctions. For example, in 'John hit Bill and then he fell', contrary to either the parallelism or the subject assignment account, 'he' refers to 'Bill' due to coherence; in 'John impressed Bill because he ...' and 'John admired Bill because he ...', 'he' is more expected to refer to 'John' in the first sentence but to 'Bill' in the second sentence due to the direction of causality implied within the verbs. Factors affecting resolution may also

influence production. Hence, in order to check for the role of *referential shift*, it is necessary to use strictly controlled parallel structures.

1.2. The Current Study

The current study aims to answer the much-neglected questions about anaphoric accentuation in production in strictly parallel structures: first, does ambiguity play a role in the prosodic patterns? Specifically, does the degree of accentuation differ when other strong disambiguating cues are present? Second, is there a systematic correlation between accentuation and *referential shift*?

Previous studies are based on intonational languages such as English. It is noted that the interaction between prosody and anaphors may be universal [14]. To add cross-linguistic data, this study tackles the above questions via a production experiment in a tonal language - Mandarin. Results on the interaction of tone and information structure in Mandarin show that Mandarin indicates prosodic prominence by use of duration, intensity and F₀ range. Especially, F₀ is raised for high tone at varied focal positions as in English [23]. Hence, this paper restricts all concerned anaphors to high tone, which, combined with the same canonical word order as English, makes it directly comparable with previous observations on English. For this reason, this paper will continue to use 'accent' to denote prosodic prominence at phrase level in Mandarin to avoid terminological confusion, though it remains an issue of debate. Meanwhile, the interaction between tone and accentuation is eliminated, making it also directly comparable among different anaphors and conditions.

To minimize confounding factors, experimental materials are strictly controlled. Elicited sentences are always triplets of short sentences. The first sentence sets the scene, and the other two sentences depict certain actions that happened in temporal order in strictly parallel structures, in which the verbs are identical, as shown in (5) (S-subject; V-verb; O-object). In previous studies, negative verbs such as 'hit' and 'insult' were most frequently used. To achieve a semantic balance, three verbs with no perceived causal bias are selected: one positive verb '帮' (help), one neutral verb '背' (carry) and one negative verb '排' (push). '然后' (and then) is used as the conjunction. All the verbs have a high tone so that different verbs will not influence the prosodic pattern of the sentence frame (5).

(5) 大家 在……。 开始 S_1VO_1 ,然后 S_2VO_2 。 everyone(present article)first S_1VO_1 then S_2VO_2 Everyone was doing... First S_1VO_1 , and then S_2VO_2 .

Three types of anaphors are under investigation, including the third person singular pronouns '他 /tha/' (he/him) and '她 /tha/' (she/her), the plural pronoun '他们 /thaman/' (they/them) and proper names. In Mandarin, the third person singular pronouns share exactly the same pronunciation with a high tone, thus this paper treats them as the same pronoun. There are five characters involved – '小刚 /kaŋ/' (Xiao Gang, male), '小当 /taŋ/' (Xiao Dang, male), '小汤 /thaŋ/' (Xiao Tang, male), '小芳 /faŋ/' (Xiao Fang, female) and '小康 /khaŋ/' (Xiao Kang, female). 'Xiao' indicates hypocorism, while 'Gang' 'Dang' 'Tang' 'Fang' and 'Kang' are either first names or surnames. The selection of these names meets the following criteria: 1) they are all typical Chinese names; 2) their last syllables share the same rhyme; 3) the last syllables share the same tone with the third person singular pronouns'.

To approach the research goal, *Referential Status* and *Ambiguity* are manipulated in both subject and object positions for three types of *Anaphoric Form* (the third person singular pronouns, plural pronoun and proper names):

Referential Status is defined as the relative syntactic position of an anaphor and its antecedent. It has three values: Unchanged means that an anaphor remains in the same position as its antecedent; Shifted means that an anaphor and its antecedent are in opposite positions; Switched means that the syntactic positions of anaphors in the last sentence and their antecedents are switched. The reason for making a distinction between Shifted and Switched is that information structure may influence the prosodic patterns. In the Shifted condition there is one anaphor and a new character in the last sentence, while there are two anaphors and no new characters in the last sentence in the Switched condition. Referential shift is involved in both Shifted and Switched conditions.

Ambiguity has two values: Ambiguous and Unambiguous. Ambiguity is manipulated by anaphoric form and number agreement of a pronoun: in the Ambiguous condition the antecedent of a pronoun and its competitor $(S_1 \text{ and } O_1)$ are of the same number, while in the Unambiguous condition the antecedent of a pronoun and its competitor are of different number, or S_1 and O_1 are both proper names.

Table 1 illustrates the examples for the last two sentences under different conditions within each factor (the bolded elements are co-referential):

Table 1. Examples under different conditions.

Proper name: Subject; Unchanged; Unambiguous					
开始	小剛	帮	了	小芳,	
first	Xiao Gang	help	PERF	Xiao Fang	
然后	小剛	帮	了	小当	
then	Xiao Gang	help	PERF	Xiao Dang	ζ.
First Xiao Gang helped Xiao Fang, and then Xiao Gang helped Xiao Dang.					
Singular pronoun: Object; Switched; Ambiguous					
7170	小当 Xiao Dang		了 d PERF	小 刚 Xiao Gang	
then		d PER	F Xia	o Dang	Xiao Dang
First Xiao Dang carried Xiao Gang , and then he carried Xiao Dang. Plural pronoun: <i>Object</i> ; <i>Shifted</i> ; <i>Unambiguous</i>					
开始	小当 7	印 小	芳	推了	小刚
first	Xiao Dang d	and Xi	ao Fang	oushed PERF	Xiao Gang
然后	小汤	惟 了	他们		
then	Xiao Tang p	oush PE	RF them		
First Xiao Dang and Xiao Fang pushed Xiao Gang, and then Xiao Tang pushed them .					

Fundamental frequency F_0 and duration are taken as parameters indicating prosodic prominence. In accordance with previous observations on accentuation of unambiguous anaphors [7-10], it is hypothesized that accentuation occurs whenever there is a referential shift regardless of ambiguity; hence higher F_0 and longer duration are expected for the *Shifted* and *Switched* conditions. For the potential influence of ambiguity, if ambiguity does play a role, it is anticipated that anaphors in the *Ambiguous* condition have a higher F_0 and longer duration than anaphors in the *Unambiguous* condition.

2. Methodology

2.1. Experimental Design

To elicit utterances as naturally as possible, a cooperative story-telling-and-understanding game is designed. The speaker first looks at three pictures on the screen and then describes what happened to a listener using the fixed sentence frame (5). The three pictures correspond to the three sentences in (5): the first picture sets the scene – climbing a mountain, playing basketball or playing a game; the other pictures depict certain actions with different agents and patients. The occurrence of the five characters is balanced. The presentation of stimuli and auto-recording are achieved using ZEP [24], a system for implementing and running (psycholinguistic) experiments.

All anaphors are balanced in both subject and object positions. For pronouns, there is only one concerned pronoun in the last sentence in conditions *Unchanged* and *Shifted*, but it is possible to have two pronouns in the *Switched* condition. This study focuses on cases with maximally one pronoun in the last sentence. There are 53 stimuli for the purpose of this study. These stimuli (together with 10 other stimuli for a related study) are divided into four tasks, in which the speakers are instructed to follow a specific constraint on using a certain type of anaphors in a fixed syntactic position.

2.2. Data Collection

2.2.1. Participants

36 Mandarin Chinese speakers (18 male and 18 female; ages 18-26, mean age 20.4) at Beijing International Studies University and Communication University of China participated in the experiment. All the speakers began to speak Mandarin Chinese before or when they entered primary school (7 native speakers of Mandarin, 22 native speakers of Beijing dialect, 7 native speakers of other dialects in north China). An experiment assistant acted as the listener throughout the 36 experimental sessions. Participants received a financial compensation.

2.2.2. Procedure

The participants were tested in a quiet room. In the experimental setting, the speaker sat in front of the screen and the listener sat back to back with the speaker so that they could not see each other, which prevented other non-verbal communications such as eye contact and gestures. There was a short oral introduction before the experiment. The listener was asked to fill in an answer sheet based on the descriptions and the participants were told that their cooperation would be graded based on both the speaker's fluency and the listener's correct rate. In this way, the speakers would be more involved. No communication was allowed between the speaker and the listener throughout the experiment.

The experiment began with a familiarization practice. The images of the five characters were presented followed by two rounds of a name-figure mapping test. Then the tasks began. In the first task, the speaker was asked to use full names to refer to all the characters; in the last two tasks, the speaker was asked to use a pronoun to refer to either the agent or the patient in the last picture, respectively (the second task is irrelevant to this study). In each trial, three pictures were firstly presented. When the speaker clicked the start button, the pictures disappeared and the audio recording began. There

were four practice trials in each task. The whole experiment was self-paced.

2.3. Data extraction and analysis

The obtained audio recordings (in wav format) were annotated and labelled via Praat [25]. Since the names share the same rhyme, only the rhymes in the second syllables were annotated to reduce the duration difference introduced by different onsets. The same method also applied to pronouns. Mean F_0 and duration of annotated syllables were extracted after manual correction of F_0 . In order to eliminate individual difference, the mean F_0 was normalized into Z-score. SPSS was employed for statistical analysis and the Generalized linear mixed model (GLMM) was adopted, which suits best to the categorical data in current study (random effects factors are speaker ID and sentence ID; fixed effects factors are Referential Status, Ambiguity and Anaphoric Form, which is not a concerned factor in this study but needed for GLMM).

3. Results

3.1. F₀ analysis

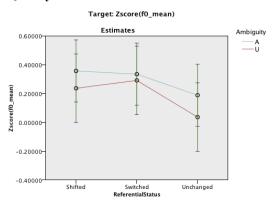


Figure 1: Interaction between Referential Status and Ambiguity for normalized F₀ in Subject.

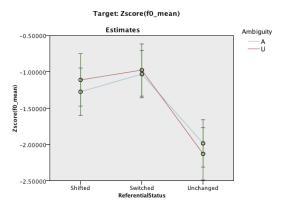


Figure 2: Interaction between Referential Status and Ambiguity for normalized F_0 in Object.

GLMM analysis of the normalized F_0 mean of subject anaphors shows: 1) F_0 for *Shifted* and *Switched* is significantly higher than *Unchanged* (p < 0.001) and there is no significant difference between *Shifted* and *Switched* (p = 0.331). 2) No significant difference is found between the *Ambiguous* (A) and *Unambiguous* (U) conditions (p = 0.431). The results for

object anaphors show: 1) F_0 for *Shifted* and *Switched* is significantly higher than for *Unchanged* (p < 0.001) and F_0 for *Switched* is significantly higher than for *Shifted* (p = 0.012). 2) No significant difference is found between the *Ambiguous* and *Unambiguous* conditions (p = 0.361). No interaction (p > 0.05) between *Referential Status* and *Ambiguity* is found for either subject or object, as illustrated in Figures 1 and 2.

3.2. Duration analysis

Target: duration Estimates Ambiguity Ambiguity A — U Shifted Switched | Unchanged | ReferentialStatus

Figure 3: Interaction between Referential Status and Ambiguity for Duration in Subject.

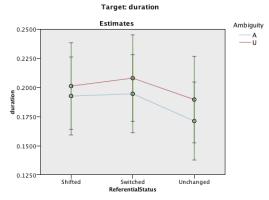


Figure 4: Interaction between Referential Status and Ambiguity for Duration in Object.

The results for duration demonstrate a similar pattern. For subject anaphors: 1) Duration in the *Shifted* and *Switched* conditions is significantly higher than *Unchanged* (p = 0.005) and there is no significant difference between *Shifted* and *Switched* (p = 0.911). 2) There is no significant difference between the *Ambiguous* and *Unambiguous* conditions (p = 0.322). For object anaphors: 1) There is no significant difference the *Ambiguous* and *Unambiguous* conditions (p = 0.619). 2) Duration in the *Shifted* and *Switched* conditions is significantly higher than in the *Unchanged* (p < 0.001) condition; and there is no significant different between *Shifted* and *Switched* (p = 0.141). No interaction (p > 0.05) between *Referential Status* and *Ambiguity* is found for both subject and object, as illustrated in Figures 3 and 4.

4. Discussion

In F₀ analysis, the main effects and interactions for both subject and object demonstrate that anaphors receive an accent in the referentially shifted (*Shifted* and *Switched*) conditions,

and no evidence shows that ambiguity plays a role in the accent assignment. The results for duration exhibit the same pattern: both subject and object anaphors are significantly longer in referentially shifted conditions, and ambiguity does not influence duration. These results provide evidence from Mandarin for the observations (1) - (3): anaphors undergoing a referential shift acquire an accent. Since no significant difference is found between the Ambiguous and Unambiguous conditions for both F₀ and duration, the 'disambiguation effect' of accent on ambiguous pronouns in strictly Mandarin parallel structures comes from a generalized relationship between prosody and referential status. There are several accounts of this correlation between prosody and referential shift, such as repetition-based account [26], centering-based account [27] and stress-shift account [28], but no consensus has been made as to why referential shift triggers anaphoric accentuation. Hence, more theoretical work is needed in future studies.

Despite the consistent patterns in both syntactic positions, both F₀ and duration exhibit a distinct asymmetry in different positions, as illustrated in Figures 1-4. Subject anaphors have an overall higher Fo and a shorter duration than object anaphors, which we attribute to the declination effect and boundary tone. In addition, F₀ differences in subject position are much smaller than in object position, which indicates that subject anaphors have a narrower range in signalling accentuation while object anaphors have a wider range. In the Unchanged condition, the anaphor becomes given information and thus is de-accented. Because a subject anaphor is at the beginning of a sentence, F₀ cannot be lowered to a large extent, whereas for object anaphors, the subject in the utterance is contrastive focus, and object anaphors become attached to the verb and F₀ is much lowered due to post-focus F₀ drop [29]. This implies that referential shift of object anaphors may be more easily indicated by prosody, which is supported by the significant F₀ difference between Switched and Shifted in object position. Whether this asymmetry also influences perception calls for further study.

5. Conclusions

Through a production experiment, this study explores the role of prosody on anaphors in Mandarin. On the basis of the findings, it can be concluded that in strictly Mandarin parallel structures, accent on anaphors is introduced by a referential shift. No evidence was found that ambiguity plays a role in anaphoric accentuation influence accentuation, hence the role of prosody in anaphora production is not to disambiguate but to signal different referential statuses. Whether this is a language universal conclusion needs further investigation in other languages. This finding challenges the general view on accented pronouns and will deepen our understanding on semantics-prosody relationship. Considering the possible wide misunderstanding on accented ambiguous pronouns, it is time to reconsider the nature of accented anaphors, and future studies should treat the interaction between prosody and pronouns with more caution.

6. Acknowledgements

This study is supported by the China Scholarship Council – Utrecht University (CSC-UU) PhD-Programme. We express our thanks to Chris van Run, Kirsten Schutter and Wenyong Hu for their technical and statistical support, and Xiaoli Dong for her practical suggestions.

7. References

- [1] Akmajian, Adrian, and Ray Jackendoff. "Coreferentiality and stress." *Linguistic Inquiry* 1, no. 1 (1970): 124-126.
- [2] Smyth, Ron. "Grammatical determinants of ambiguous pronoun resolution." *Journal of Psycholinguistic Research* 23, no. 3 (1994): 197-229.
- [3] Theune, Mariët. "Parallelism, coherence, and contrastive accent." In Proc. of the 6th European Conference on Speech Communication and Technology. 1999.
- [4] Poirier, Josée, Matthew Walenski, and Lewis P. Shapiro. "The role of parallelism in the real-time processing of anaphora." *Language and Cognitive Processes* 27, no. 6 (2012): 868-886.
- [5] Grober, Ellen H., William Beardsley, and Alfonso Caramazza. "Parallel function strategy in pronoun assignment." *Cognition* 6, no. 2 (1978): 117-133.
- [6] Taylor, Ryan C., Laurie A. Stowe, Gisela Redeker, and John CJ Hoeks. "Comprehension of marked pronouns in Spanish and English: Object anaphors cross-linguistically." *The Quarterly Journal of Experimental Psychology* 66, no. 10 (2013): 2039-2059.
- [7] Dogil, Grzegorz, Jonas Kuhn, Jörg Mayer, Gregor Möhler, and Stefan Rapp. "Prosody and discourse structure: Issues and experiments." In Proc. of the ESCA Workshop on Intonation: Theory, Models and Applications, Athens, Greece. 1997.
- [8] Chambers, Craig G., and Ron Smyth. "Structural parallelism and discourse coherence: A test of centering theory." *Journal of Memory and Language* 39, no. 4 (1998): 593-608.
- [9] Wolters, Maria, and David Beaver. "What does he mean." In Proc. of the Annual Meeting of the Cognitive Science Society. 2001.
- [10] Venditti, Jennifer J., Matthew Stone, Preetham Nanda, and Paul Tepper. "Toward an account of accented pronoun interpretation in discourse context: Evidence from eye-tracking." Rutgers Center for Cognitive Science: Technical Report RuCCSTR-68 (2001).
- [11] Kothari, Anubha. "Accented pronouns and unusual antecedents: A corpus study." In Proc. of the 8th SIGdial Workshop on Discourse and Dialogue. 2007.
- [12] Rello, Luz, and Joaquim Llisterri. "Temporal Prosodic Features in Pronominal Anaphora Resolution in Spanish."
- [13] Wasow, Thomas. *Anaphora in Generative Grammar*. Vol. 2. John Benjamins Publishing, 1979.
- [14] Oehrle, Richard T. "Common problems in the theory of anaphora and the theory of discourse." In H. Parret, M. Sbisà and J. Verschueren (eds.): Possibilities and Limitations of Pragmatics: Proceedings of the Conference on Pragmatics, Urbino, July 8– 14, 1979 (1981): 509-30.
- [15] De Hoop, Helen. "On the interpretation of stressed pronouns." In R. Blutner & H. Zeevat (eds.): Optimality Theory and Pragmatics, pp. 25-41. Palgrave Macmillan UK, 2004.
- [16] Lakoff, George. "Presupposition and relative well-formedness." In D.S. Steinberg and L.A. Jakobovits (eds.): Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology (1971): 329-340.
- [17] Wolters, Maria, and Donna K. Byron. "Prosody and the resolution of pronominal anaphora." In *Proceedings of the 18th* conference on Computational linguistics-Volume 2, pp. 919-925. Association for Computational Linguistics, 2000.
- [18] Crawley, Rosalind A., Rosemary J. Stevenson, and David Kleinman. "The use of heuristic strategies in the interpretation of pronouns." *Journal of Psycholinguistic Research* 19, no. 4 (1990): 245-264.
- [19] Zuckerman, S., N. Vasić, E. Ruigendijk, and S. Avrutin. "Experimental evidence for the subject rule." In *Proceedings of LATL*, vol. 18, 2002.
- [20] Wolf, Florian, Edward Gibson, and Timothy Desmet. "Discourse coherence and pronoun resolution." *Language and Cognitive Processes* 19, no. 6 (2004): 665-675.
- [21] Kertz, Laura, Andrew Kehler, and Jeff Elman. "Grammatical and coherence-based factors in pronoun interpretation."

- In Proceedings of the 28th Annual Conference of the Cognitive Science Society, pp. 1605-1610. 2006.
- [22] Garvey, Catherine, Alfonso Caramazza, and Jack Yates. "Factors influencing assignment of pronoun antecedents." *Cognition* 3, no. 3 (1975): 227-243.
- [23] Xu, Yi. (1999). "Effects of tone and focus on the formation and alignment of F 0 contours". *Journal of Phonetics* 27, no. 1 (1999): 55-105.
- [24] Veenker, Theo. J. G. (2014). "The Zep Experiment Control Application (Version 1.10) " [Computer software]. Beexy Behavioral Experiment Software. Available from http://www.beexy.org/zep/
- [25] Boersma, Paulus P. G. "Praat, a system for doing phonetics by computer." Glot international 5. 2002.
- [26] Gleitman, Lila R. "Coordinating conjunctions in English". Language 41, no. 2 (1965): 260-293.
- [27] Kameyama, M. "Stressed and unstressed pronouns: complimentary preferences". In: P. Bosch and R. van der Sandt (eds.): Focus: Linguistic, Cognitive and Computational Perspectives. Cambridge University Press, 1999.
- [28] Reinhart, Tanya. Interface Strategies: Optimal and Costly Computation. Linguistic Inquiry Monograph 45. Cambridge, US: MIT Press, 2006.
- [29] Cooper, William E., Stephen J. Eady, and Pamela R. Mueller. "Acoustical aspects of contrastive stress in question-answer contexts." *The Journal of the Acoustical Society of America* 77, no.6 (1985): 2142-2156.