

Early and Late Nuclei in Yes-No Questions: Tails or High Rises?

Nancy Hedberg, Juan M. Sosa and Emrah Görgülü

Department of Linguistics
Simon Fraser University, Burnaby, Canada
{hedberg; sosa; ega5}@sfu.ca

Abstract

In this paper we analyze the placement of the nucleus in yes-no questions and its significance in American English. We show that the vast majority of positive yes-no questions are expressed through a low rise, often with at least one word as a tail. High rise is the second most common yes-no question nucleus; and so we are interested in the question of why a high-rise instead of a low rise contour is sometimes selected by the speaker. It can be difficult to tell whether a question should end in a late high rise, or whether an early low-rise should be postulated with a tail that is part of the rise. We bring phonetic criteria to bear on this question, and also show that post-nuclear tails tend to consist of function words or else of information that is in some sense given in the discourse. Finally, we present evidence that the discourse function of high rises overlaps with the function of tails to such an extent that it is economical to consider the high pitch accent of a high rise nuclear tune as simply an accented part of what otherwise could be analyzed as a tail.

1. Introduction

While it is universally acknowledged that English yes-no questions differ prosodically from declarative statements by being pronounced with rising instead of falling intonation, many aspects of question intonation are still little understood, especially in natural speech. We claimed in [1] that there is a template for unmarked yes-no questions in American English. That is, a typical yes-no question starts with a falling head and then ends with a low rise (L*H-H%). In that paper and the present one we combine the broad constituent analysis of [2], which distinguishes Head + Nucleus + Tail as subparts of an Intonation Phrase, with the fine-grained ToBI notation of [3], which annotates each pitch accent, phrase accent and boundary tone. Other end shapes to questions are also possible, e.g. questions ending in a high-fall (H*L-L%) and questions ending in a high-rise (H*H-H%). In the current study, we aim at describing the end shapes of yes-no questions with more extended data and refined analysis than we did in [1].

The issue of nucleus placement is generally assumed to have semantic or pragmatic conditioning because the nucleus marks the end of the focus—or new information—of an intonation unit, and material after the nucleus has to be given in some sense in the discourse [4]. Tails are defined by [2] as “all syllables following the nucleus” (p. 15). Tails are widely assumed to consist of either function words or material that is either given information or relatively less informative than final accented material [4,5,6]. In the autosegmental ToBI system, the phrase accent and boundary tone in a nuclear tune are extrapolated over the tail and determine its phonetic shape, which is rising in most yes-no questions.

We are particularly interested in analyzing tails and high-rises because in a more fine-tuned second stage of coding, we found early nuclei in the questions more often than we did in our first phase of coding that was reported on in [1]. This perceived discrepancy caused us to go back and check our earlier work. As a result of this re-examination, we reanalyzed some questions that we earlier analyzed as late high rises as early low rises with rising tails. Intonationally, a tail differs from a high rise in that it goes up at a steady inclination from the early L*, whereas the H* pitch accent of the high rise has a descent or plateau just before it (see figures 1 and 2).

We find this result encouraging because we had in general been successful at finding semantic or pragmatic justification for classifying tails, while we were unable to come up with a very consistent semantic or pragmatic explanation for the high rises that we reported in [1]. We did come up with some sub-generalizations, however: e.g. that high rises mark an anaphoric or deictic element that would normally be unaccented, or that the information encoded is related to material in the discourse context instead of being ‘out of the blue’. It is noteworthy that these pragmatic characteristics are precisely those characteristics that have been identified in the literature and in our data as characteristic of tails. This fact leads us to hypothesize that the lexical/pragmatic characteristics of high rises are a subset of those of tails.

2. Method

Our data are taken from the CallHome Corpus of American English [7, 8] and the Fisher English Corpus [9]. 104 examples of positive yes-no questions from CallHome were analyzed in [1], and we have added to that corpus 241 additional questions from CallHome and 74 questions from Fisher. In total we have 419 questions. Only positive yes-no questions with the syntactic form of interrogative sentences are considered.

The first 104 questions were annotated according to the ToBI system and were coded for constituent analysis by the three authors of [1]. The remaining 241 questions were subsequently annotated for ToBI categories by the first two authors of the current paper, and the final 74 questions were subsequently annotated by the first and third author of this paper. The CallHome corpus consists of telephone calls between people who know each other and the Fisher corpus consists of telephone calls between people who do not know each other. We used Praat (v. 4.4.04) and Pitchworks (v. 8.9.5.5) for phonetic analysis of the speech files.

Our phonological analysis follows the ToBI guidelines quite closely, but we have supplemented ToBI categories with a category of “upstep” (annotated as $\dot{\uparrow}$) and “increased range” (annotated as \uparrow) when such annotation seemed warranted. A three-way coding reliability study based on this system was reported on in [1], with the resulting transcriber-pair-word agreement of 75.7% on presence and type of pitch accent

concluded to be typical for reliability results reported on for ToBI coding in the literature. Now that the first two authors of this paper have coded 241 more questions together, we feel even more confident that we have arrived at a reliable system of tonal coding.

After performing the ToBI annotations, we classified the questions into groups exhibiting tails and high rises, and carefully reviewed the phonetic basis for annotating the resulting examples, while also examining the transcripts to ascertain possible semantic and pragmatic conditioning of these patterns. We did the phonetic analysis before we did the pragmatic analysis, thus avoiding circular reasoning.

3. Results

3.1 Classification of Nuclei

The classification of the final nuclear contour in each yes-no question is shown in Table 1. It is noteworthy that the vast majority of questions ended in some type of low rise (333/419 or 79.5%). The second most common category of contour was the high rise, which occurred 9.3% of the time. In our analysis in [1], we had postulated 18 high rises out of 104, or 17.3%. As noted above, we were especially interested in exploring reasons for this difference in annotation results, and this caused us to reanalyze some of the earlier high-rises before arriving at the distribution in Table 1.

Table 1: *Distribution of Nuclei*

Nucleus	ToBI category	Number
Low rise	L*HH%	311
	L*H _i H%	6
	L*H _i H%	5
	L*+HHH%	5
	L*LH%	6
High rise	H*HH%	34
	!H*HH%	5
High fall	H*LL%	15
	!H*LL%	4
	L+H*LL%	4
Level	H*HL%	12
	!H*HL%	2
	¡H*HL%	1
	L*HL%	2
Low fall	L*LL%	7
Total		419

3.2 Classification of Tails

We classified all the tails in our data semantically and pragmatically, finding that they fall into the categories detailed in Table 2. Examples of each of these are shown below.

- (1) **Does it move around** yet?
L*+H H* L*HH%
- (2) Do you **stay** at like **hostels** and stuff?
H* L*HH%
- (3) Did you **speak** to her?
L*HH%

Table 2: *Lexical/Pragmatic Distribution of Tails*

Tails	Number
Function words: adverbs, prepositions, copula (1)	42
Semantically light content words (2)	9
Personal and demonstrative pronouns (3)	52
Words following normal early stress (4), (5)	17
Activated open proposition (6), (7), (8)	24
Unactivated mutual knowledge (9), (10), (13)	4
Words following contrastive stress (11), (12)	14
Total:	162

Examples (4) and (5) are instances of normal focal stress on a pre-final word. (4) is an example of compound stress and (5) is an example of an intransitive verb that normally takes stress on its subject [10]:

- (4) Are you **thinking of doing public** interest law?
H* !H* L*HH%
- (5) Oh **did** your **sister** die?
L* ¡H*HL%

Examples (6) and (7) show instances of narrow focal stress on an activated open proposition [11]. In (6) it has been asserted that the addressee has been to Holland before, hence the open proposition (OP), "x has been here before", is activated. In (7) it has been asserted that the addressee had been at the beach but not in the water with the baby, and the speaker is concerned about the baby being left in the sun; hence the OP, "A did/didn't do something in relation to the beach", is activated.

- (6) Has **Kim** been here before?
L*HH%
- (7) But did you **go on** the beach?
L* L*HH%

Note that these tails were identified phonetically as such before we examined the transcript to ascertain whether positing a tail was justified pragmatically.

Information that is inferable but not directly activated is sometimes considered 'given' enough to warrant marking as a tail. Thus, the question in (8) was uttered after the addressee had mentioned that the referent under discussion drives her child to school, so the OP, "She does/does not pick him up" is inferable.

- (8) **Does** she pick him up?
L*HH%

Unactivated but mutually familiar [12] or mutually known information is also sufficient to trigger prosodic treatment as a tail, as shown in (9) and (10) In (9), it is mutually known, even inferable, that her name is or isn't Arlene. That is, the speaker is implying that she should know the name but can't

remember. Thus, the answer will fulfill a reminder function and the question sounds like an echo question in spite of the low rise.

- (9) Um **Arlene**, was **that** her name?
L*HH% L*HH%

(10) is similar in that the speaker implies that she should know whether they had ever met Rhea since people normally remember whether they have met someone. Again the speaker is asking for a reminder of mutually known but unactivated information. That both examples involve left dislocation perhaps indicates a characteristic feature of this type of question.

- (10) Rose's sister Rhea, **did** we ever meet her?
L*HH%

It must be acknowledged, of course, that all questions presuppose that the addressee knows the answer to the question. The situation is different in (9) and (10). Here, the information is marked as also expected to be known by the speaker.

Finally, in (11) it has been activated that a third person has received the addressee's letter, so the determiner in "my letters" is explicitly contrastive. In (12), becoming a "dual citizen" is implicitly contrasted with just becoming a "citizen."

- (11) Did you **ever get** any of **my** letters?
H* !H* L*HH%

- (12) Are you **gonna become** a **dual** citizen?
L* L* L*HH%

3.3. High-Rises Reclassified as Tails

In (13) we show an example of a question that we analyzed in 2005 as a late high-rise (a), but reanalyzed in 2007 as an early low-rise (b):

- (13) Is **Liat** still going with that guy?
a. L* L* H*HH%
b. L*HH%

In the present research, we came up with a principled phonetic motivation for classifying speech chunks as part of the tail as opposed to receiving a nuclear accent. As Figure 1 shows, in this case, there seems to be a steady inclination between 'Liat' and 'guy' except for the little hook up at the end of 'guy'.

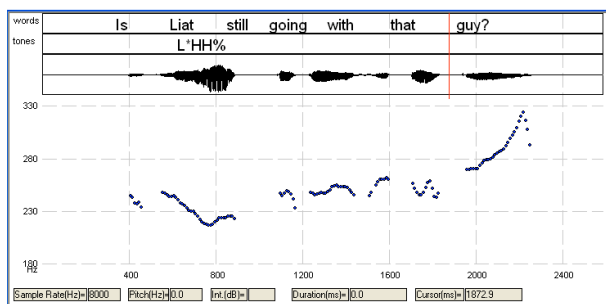


Figure 1: *Reanalyzed low-rise + tail*

We suggest that there is perhaps also a good pragmatic case for saying that 'Liat' is actually the nucleus with a long rising tail following. It is common ground that Liat had been going with that guy. (Note that 'still' indicates a presupposition, plus there is reminder 'that' [12] on 'that guy'). It thus appears again that mutually known but not activated information is enough to trigger deaccenting and consequent marking as a tail.

3.4 Remaining High-Rises

Example (14), on the other hand, is an example where we did not reclassify the high-rise as a tail. Here, 'I sound stucky to x' is activated, and 'you' is contrastive. ("I sound stucky to myself—do I sound stucky to you?"). This contrastiveness apparently justifies a pitch accent on the pronoun 'you', which thus becomes the nucleus. The generalization from [1] that high-rise occurs on items that are normally unaccented but are nevertheless accented holds here.

- (14) Do **I** sound stucky to **you**?
L*+H H*HH%

We also didn't reclassify example (15).

- (15) Is **anybody coming** before **that**?
L* L* H*HH%

The accent here appears to be one that signals that the referent is activated but not in the focus of attention and thus needs to be stressed in order to signal focus shift [12]. As shown in Figure 2, the pitch remains low until it reaches the nucleus on 'that' where the substantial jump into the high level tone takes place entirely during that one word.

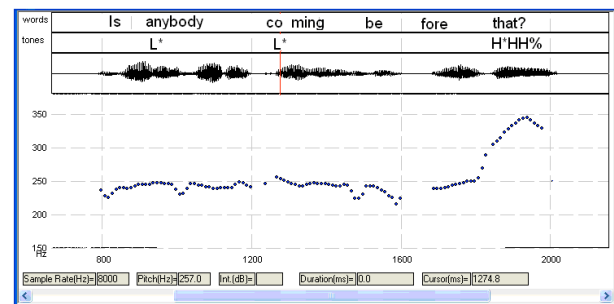


Figure 2: *Non-reanalyzed final high rise*

3.5 High Rises in the Overall Data

There were 39 high rises in our overall data. We classified them as shown in Table 3, finding that all of them mark normally unaccented but nevertheless accented material. That is, it would have been appropriate, at least in another context, to deaccent these items completely and have them constitute a tail or part of a tail.

For example, the normally unaccented final word in a nominal compound gets a H* accent in (16).

- (16) Did you **get** my **post card**?
H* L* H*HH%

Three questions contained a single H*HH% accent, which of course cannot be considered an accented functional tail as

there is no previous accented item to serve as a functional nucleus. However, these accents do mark given information. In (17), the proposition that the addressee has explained it is activated, and in (18), the high-rise falls on a pronoun.

- (17) Can you **explain** that again?
H*HH%
- (18) But are **they** gonna still come?
H*HH%

Table 3: Lexical/Pragmatic Distribution of High Rises

Tails	Number
Function words: adverbs, prepositions, adjunct PPs	12
Semantically light content words	2
Personal and demonstrative pronouns (14), (15), (18)	5
Words following normal early stress (16)	3
Other activated information (17)	14
Unactivated mutual knowledge	3
Total:	39

Here 'they' is a contrastive narrow focus filling an activated open proposition 'x is/is not going to still come'. The pitch track in Figure 3 shows the H* accent on 'they', followed by a steadily rising tail. Notice that the nucleus is higher than the previous unstressed syllables.

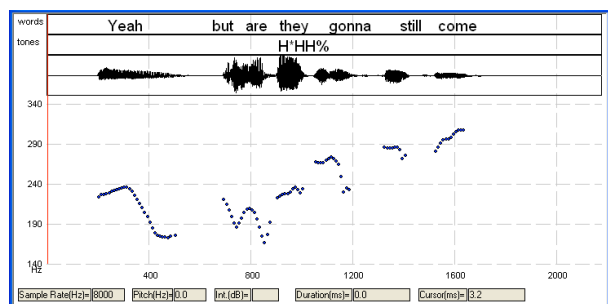


Figure 3: High rise with tail

4. Conclusion

We can confirm that the low rise (L*HH%) is the normal yes-no question nuclear tune in American English, and that tails are used to mark information that is in some sense 'given' in the discourse. When material within such information for some reason needs an accent, it receives a H* accent, with the result that the question receives a high-rise (H*HH%) nucleus.

It now doesn't matter (very much) for annotation if we give a H* to a 'given' item or not (in a functional tail). That is, we don't have to deliberate over whether the nucleus should be early or late in a question that sounds like it might have a late H* since we have a common lexical/pragmatic explanation for the two accent patterns. This could help in providing coding guidelines for relatively naïve annotators.

Our analysis can also contribute to the 'beat' versus pitch accent issue. Stressable words in a 'deaccented' tail may get a beat but are still 'deaccented'. An interesting question is whether these unaccented beats are imaginary, i.e. derived

from the language system, just not currently being put to use. We can say that it is possible to put them to use, in which case they receive a H* accent in a yes-no question.

With regard to both conclusions, it is interesting to note that Cruttenden suggested in [13] that low-rise and high-rise yes-no questions sound 'light' compared to 'serious' falling tones, with high-rise tones sounding more 'casual' than low-rise tones. He also said that high-rising tones are "much more frequent in American English than in British English", which is why Americans sound casual to the British while the British sound formal to Americans. (p. 59).

The overwhelming prevalence of low-rise nuclei in our data suggests that Cruttenden is wrong about high-rise occurring frequently in American English. Our conclusion is supported by Pierrehumbert & Hirschberg's assertion in [14] that yes-no questions tend to be marked with L*HH% in American English. Finally, our explanation for the high-rise tones, i.e. that they mark information that is given in the discourse, also suggests that Cruttenden's impressionistic assertion is wrong about interpreting high-rises as 'casual'.

5. References

- [1] Hedberg, Nancy; Sosa, Juan M.; Fadden, Lorna, 2006. Tonal Constituents and Meanings of Yes-No Questions in American English. *Proceedings of Speech Prosody 2006*, Dresden, Germany.
- [2] O'Connor, J. D.; Arnold, G. F. 1973. *Intonation of Colloquial English* (2nd edition). Longman.
- [3] Beckman, M.; Ayers-Elam, G. Guidelines for ToBI Labelling. Ohio State University.
- [4] Halliday, M. A. K. 1967. Notes on transitivity and theme in English: Part. 2. *Journal of Linguistics* 3. 199-244.
- [5] Zacharski, R.; Monaghan, A.I.C.; Ladd, D.R.; Delin, J. 1995. *BRIDGE: Basic Research on Intonation for Dialogue Generation*. University of Edinburgh.
- [6] Zacharski, R. 1993. *A Discourse Pragmatics Model of Pitch Accent in English*. Unpublished Ph.D. dissertation, University of Minnesota.
- [7] Carnavan, A.; Graff, D.; Zippener, G. 1997. CALLHOME American English Speech. Linguistic Data Consortium, Philadelphia.
- [8] Kingsbury, P.; Strassel, S.; McLemore, C.; McIntyre, R., 1997. CALLHOME American English Transcripts. Linguistic Data Consortium, Philadelphia.
- [9] Cieri, C.; Graff, D.; Kimball, O.; Miller, D.; Walker, K. 2004. Fisher English Training Speech Part 1 Speech; Transcripts. Linguistic Data Consortium, Philadelphia
- [10] Schmerling, S. 1976. *Aspects of English Sentence Stress*. University of Texas Press.
- [11] Prince, E. 1986. On the syntactic marking of presupposed open propositions. In Farley, A.; Farley, P.; McCullough, K.-E., eds., *Papers from the Parasession on Pragmatics and Grammatical Theory*, Chicago Linguistic Society: 208-22.
- [12] Gundel, J. K., Hedberg, N; Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69.274-307.
- [13] Cruttenden, A. 1986. *Intonation*. Cambridge: Cambridge University Press.
- [14] Pierrehumbert, J.; Hirschberg, J. 1990. The meaning of intonation contours in the interpretation of discourse. In Cohen, P.R.; Morgan, J.; Pollack, M. E. (eds.) *Intentions in Communication*. MIT Press, 271-311.