The Prosody and Meaning of Wh-Questions in American English

Hedberg, Nancy, Juan M. Sosa, Emrah Görgülü, Morgan Mameni

Department of Linguistics, Simon Fraser University, Burnaby, BC, V5A 1S6 Canada

We report on a corpus study of the intonation and meaning of 200 spontaneous wh-questions in American English. The most frequent final nuclear contour is falling, and this category correlates with the most frequent pragmatic functions of wh-questions in general, such as requesting elaborative detail, opening a subtopic and directing information flow. The pragmatic function of rising wh-questions is shown to be a generalization of the echo-question pattern, with the interrogator intending to signal with the rising intonational contour that he or she is not attempting to take the floor from the ongoing speaker, but is rather attempting to support the ongoing speaker's discourse topic by requesting background information or asking for clarification of inaudible information. We conclude that distinct nuclear contours in wh-questions correlate with differences in their pragmatic function.

Index Terms—**wh-questions, intonation, pragmatic meaning, ToBI, American English, corpus study.**

I. INTRODUCTION

THIS PAPER reports on a corpus study of wh-questions in American English. It is widely assumed that wh-questions in English tend to be falling in intonation, but this prediction has not been tested in a corpus study of spontaneous speech. Our goal was to confirm this prediction, and also to study variations in intonational shape with a view to establishing a correlation between the intonational shape of wh-questions and the meanings they convey as they occur in actual discourse.

Halliday [1] suggests that wh-questions tend to be falling in intonation because the polarity is "known" in wh-questions just as it is in declarative statements and unlike in interrogative yes-no questions, which tend to be rising [2, 3]. Gussenhoven [4] says that rising contours indicate 'testing' while falling contours indicate 'addition', but he doesn't explain why wh-questions, which presumably involve 'testing' as do yes-no questions, should be falling in intonation.

Pierrehumbert & Hirschberg [2] argue for a compositional semantic analysis of ToBI categories. They claim that the low phrase accent (L-) and low boundary tone (L%) (typical of wh-questions) both indicate completion, i.e. lack of connection to subsequent discourse; but this doesn't explain how the wh-question is linked to its subsequent answer just as a yes-no question is.

Bartels [5] accounts for the intonational difference between the two types of question by claiming that a wh-question evokes an existential presupposition in the body of the question, which is asserted by the speaker and thus is marked by an L- phrase accent. Steedman [6] claims that the most typically falling contour indicates an "uncontroversial rheme" (H*) with speaker commitment (L-L%). This analysis, like Bartels', is perhaps justified by the observation that the body of a wh-question is presupposed and thus is uncontroversial and speaker-committed.

Bolinger [7] claims that the intonational contour of a wh-question involves an early rise with a subsequent fall. He suggests that wh-questions tend to fall rather than rise as do yes-no questions because they have a "more demanding nature". Moreover, he suggests that reclamatory wh-questions tend to be rising, i.e. those questions with which speakers ask for a repetition because they failed to understand something (echo questions).

We will show that our data in general support the claims of the previous literature. The dominant pattern of wh-questions is that they end in a falling nuclear contour (81%), and the primary alternative pattern is that of ending in a rise (18%). Furthermore we find that rising questions occur when the speaker wishes to obtain information in order to follow or support the conversation but does not want to take the floor, similar to what Bolinger concluded. Our primary goal, however, is to show that a fine-grained classification of the functions that wh-questions play in dialogue can elucidate motivations behind speakers' choices of different nuclear contour patterns.

II. METHOD

The data were taken from the CallHome Corpus of American English [8], a corpus of 30-minute recorded telephone calls between people who know each other, with 10-minute segments of each of the 120 conversations transcribed; and the Fisher English Corpus [9], a corpus of transcribed ten-minute conversations on assigned topics between people who do not know each other. 200 wh-questions were extracted from the two corpora, with 87 questions coming from the CallHome corpus and 113 questions coming from the Fisher Corpus. Utterances consisting only of a wh-word were not included.

Our phonological analysis follows the ToBI guidelines [10] quite closely, but we have supplemented ToBI categories with a category of "upstep" (annotated as ¡) when such annotation seemed warranted. We used Praat (v. 4.4.04) and Pitchworks (v. 8.9.5.5) for phonetic analysis of the speech files. The search for questions in the transcripts was performed partially automatically, and we extracted the wav files using GoldWave. The last three authors annotated the sound files together. Our ToBI coding system was tested for intercoder

reliability in an earlier study [11], 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 annotation in the literature.

After performing the ToBI annotations, the last two authors classified the wh-questions into groups exhibiting different final nuclear contours, listened to the examples again and examined the transcripts to ascertain possible semantic and pragmatic conditioning of the intonational patterns. We did the phonetic analysis before we did the semantic/pragmatic analysis, thus avoiding semantic bias in the prosodic annotation.

The wh-questions were classified pragmatically according to the function that the wh-question played in the conversation. Then, the pragmatic classification was compared to the classification of nuclear contours to see if any pragmatic functions of the different nuclear contours could be identified. The categories of the pragmatic classification are discussed in section C below.

III. RESULTS AND DISCUSSION

A. The nuclear contour.

The classification of the final nuclear contour in each wh-question is shown in Table 1.

TABLE 1
NUCLEAR CONTOUR FOR WH-QUESTIONS

Nucleus	ToBI Category	Number
High Fall	H*LL%	64
	!H*LL%	34
Rise Fall	L+H*LL%	42
	L+!H*LL%	6
	L+¡H*LL%	1
	¡L+H*LL%	1
Low Fall	L*LL%	14
Low Rise	L*HH%	25
	L*H¡H%	1
	L*LH%	1
High Rise	H*HH%	3
	!H*HH%	1
Fall Rise	H*LH%	2
Rise-Fall-Rise	L+H*LH%	2
	L+!H*LH%	1
Level	H*HL%	1
	H*!HL%	1
Total		200

It can be seen from the table that High-Falls (H*LL% and variants) are the most frequent contour, occurring 98 times, or 49% of the time, followed by its closely related alternative, the Rise-Fall (L+H*LL% and variants), which occurs 50 times, or 25% of the time. 14 instances of Low-Falls (L*LL%) occurred. The third most frequent contour was the Low-Rise (L*HH% and variants), which occurred 27 times, or 13.5% of the time. Other contours occurred more rarely.

B. The Pitch Accent on the Wh-word

Whether or not the wh-word itself is pitch-accented, and if so how, is also worthy of study. Table 2 shows the distribution of wh-words in the data and shows the pitch accent occurring on the wh-words. Cases where the wh-word received the nuclear accent are indicated in the table. It can be seen that most wh-words exhibit the high that Bolinger noted as the start of a typical wh-question. In the case of L*+H, the initial high typically occurs on the auxiliary word, which is the syllable after the monosyllabic wh-word that realizes the L* component of the rising pitch accent.

Steedman [12] posits the wh-word as the 'theme' of the wh-question because it evokes but does not select from an alternative set. He thus predicts that it should be marked with an L+H* or L*+H pitch accent. Hedberg & Sosa [13] found that 60.7% of the 34 positive wh-questions that they examined exhibited an L+H* pitch accent. However almost all of the questions were asked by the moderator of the discussion program examined, who spoke very emphatically. The present study shows that that pattern of results is not typical of American English. Table 2 illustrates that only 13/200 or 6.5% of wh-question words were marked L+H*. .

TABLE 2
PITCH ACCENT ON THE WH-WORD

	H*	L*+H	L+H*	L*	L+¡H*	No Accent	Total
What	27 H*LL%: 1	33	5 L+H* LL%: 1	1 L*HH%: 2		25	95
How	21	13	3 L+H* LH%: 1		1	7	46
Where	7 H*LL%: 1	5	2	L*HH%: 2	1	2	20
When	8	8		L*HH%: 1			17
Why	2 H*LL%: 1	10	1			1	15
Who	4	1				1	6
Which	1						1
Total	73	70	13	6	2	36	200

C. Pragmatic classification.

We started our analysis by first examining the immediate environment of the questions in the transcript and discovered five binary dimensions which appeared to influence the intonational contour of the questions. Based on these five dimensions, we arrived at eleven pragmatic categories which describe the conversational function of wh-questions in discourse. Before introducing the five dimensions, we need to introduce the term *Interrogator*, which we need to distinguish from the *Speaker*. We designate the Speaker as the person who has the floor in the conversational turn and the *Hearer* as the other conversation participant whom the Speaker addresses. The *Interrogator* designates the person who asks the question, which, at question time, could either be the Speaker or the Hearer. See examples below.

The five dimensions are as follows.

(i) Information seeking: For every question we decided whether that question was information seeking or not. Non-information seeking questions include rhetorical questions, back-channeling and questions to self. In other words, we

classified a question as information seeking when the interrogator uses the question as a strategy to remedy an informational gap.

(ii) Floor orientation: This dimension determines whether the interrogator is the (S)peaker or the (H)earer. That is, S might ask a question to pass the floor to H, or else H might ask a question without claiming the floor.

(iii) Topic sensitivity: This dimension checks whether the question retains the topic of immediate discourse or changes it.

(iv) Interruption: This dimension only applies to cases where H asks the question and checks whether the question interrupts S.

(v) Content Givenness: This dimension determines whether the question asks for new information, or whether it clarifies information already provided.

Based on these dimensions, we classified each question into one of the following eleven categories. Examples are given of the first five categories.

1. Elaborative Detail (The curious interrogator): S is talking about X and H asks a question regarding X to get more detailed information. Questions in this category inquire about an at-issue topic. That is H's question builds on the information S intends to provide. The pitch track of (1), a typical falling example, is shown in Figure 1.

(1) [S has been speaking regarding a pending settlement]
 S: You know on the uh you're sitting on the steps waiting
   for the judge and that's when they settle you know.
 H: yeah. wh- when **when** are you **going** to **court**?
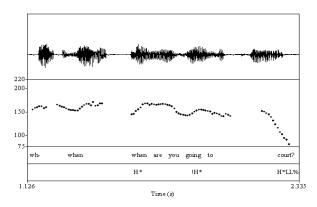   H*   !H*   H*LL%

Figure 1. High-Fall

2. Sub-Topic Initiator (The deferring interrogator): S asks the question to pass the floor to H. Topic can be the same or different.

(2) [S has been talking about a particular TV show]
 S: You know that probably men versus women could
   work […] What **else** is out there?
               H*LL%

3. Directing Information Flow (The controlling interrogator): H asks the question to influence the content of S's topic. H has the remote control, so to speak.

(3) [S has been talking about several issues in her personal
   life as a response to various questions that H directs at S]
 S: I work ten at night to six in the morning. Bo Bo's
   happy because uh he don't have to go back to day
   care.
 H: **When** do you **sleep** then?
   H*   L+H*LL%

4. Rhetorical (The un-inquisitive interrogator): This is not a homogeneous group: comprises statements couched in question form, questions to self, as well as back-channeling. Essentially, this is a class of questions that are not information seeking.

(4) [Following a pause in the conversation, S recollects her
   thoughts]
 S: **What** was I gonna **say**?
   L*+H   H*HH%

5. Supplementary Information (The interruptive interrogator): H interrupts S without claiming the floor, but merely requests background information relevant to the topic that is not at-issue or critical to S's content. The pitch track of (5), a typical rising example, is shown in Figure 2.

(5) [S has been worrying about *where* she could stay when
   she visits]
 S: But if not I mean I'm just coming anyway but I have
   no idea like where I'll go-
 H: **When** are you **planning** on **coming**?
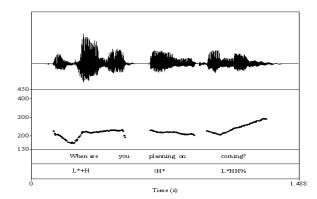   L*+H   !H*   L*HH%

Figure 2. Low-Rise

6. Reciprocal Question (The symmetrical interrogator): S asks the same question of H which S has been providing an answer to, i.e. exchanging mutual questions.

7. Topic Initiator (The entrepreneurial interrogator): This is any question that starts a new topic (generally at the beginning of the discourse), and it is floor-neutral.

8. Clarification (The uncertain interrogator): H is clarifying information that is audibly not clear.

9. Concedes to New Topic (The agreeable interrogator): S has initiated a new topic, and H asks a question on the new topic as a strategy to encourage S to continue.

10. Returns to Old Topic (The nostalgic interrogator): S asks a question to return to an old topic, either from a short digression or from an elaborate change in topic.

11. Presupposition failure (The perplexed interrogator): H raises the issue of missing certain information which S seems to presuppose H shares.

The frequency of each category is summarized in Table 3.

TABLE 3
PRAGMATIC FUNCTIONS

ED	Elaborative detail (at issue)	50
STI	Sub-topic initiator (passing the floor)	26
DIF	Directing information flow	26
R	Rhetorical (not information seeking)	26
SI	Supplementary Information (not at issue info.)	15
RQ	Reciprocal question	13
TI	Topic initiator (floor neutral)	14
CL	Clarifying	8
CNT	Concedes to new topic	8
ROT	Returning to old topic (from digression)	6
PF	Presupposition failure	4
UNC	Unclassified	4
	TOTAL	**200**

D. Nuclear contour and pragmatic function.

Table 4 presents the pragmatic function of wh-questions and their distribution with respect to their nuclear contour. The table illustrates the central finding that captures the contrast between rises and falls. Importantly, the most frequent categories that occur with a falling contour are ED, STI and DIF. It is interesting to note that High-Falls are not only the most frequent nuclear contour, but also that they exhibit the most frequent functions in exact order of frequency, i.e. ED, STI, DIF, R. Within Downstepped-High-Falls, the table shows a small dispreference for STI compared to High-Fall and a small preference for RQ. Rise-Falls show a small preference for DIF. Such small tendencies are worthy of further study.

TABLE 4
PRAGMATIC FUNCTIONS AND NUCLEAR CONTOURS

Function	Rise	Low Fall	High Fall	!High Fall	Rise Fall	Fall Rise	Level	Total
ED	2	1	24	11	11	1		**50**
STI	2		12	3	8		1	**26**
DIF			8	6	11	1		**26**
R	5	3	7	3	8			**26**
SI	13	1		1				**15**
TI	2	2	2	3	4	1		**14**
RQ		1	3	6	2	1		**13**
CL	5		2				1	**8**
CNT	1	3	2		1	1		**8**
ROT		2	2		2			**6**
PF	1		1		2			**4**
UNC		1	1	1	1			**4**
Total	**31**	**14**	**64**	**34**	**50**	**5**	**2**	**200**

While none of these categories are frequent with the rising contour, the function of rising wh-questions is predominantly SI and CL. Notably, these two latter categories rarely occur with a falling contour.

IV. CONCLUSION

We reported on a corpus study in which wh-questions occurred with a falling nuclear contour 81% of the time, and with a rising contour 18% of the time. Our pragmatic analysis shows that the difference between falling and rising wh-questions is correlated with differences in their discourse function. Falling questions are most often used to get more detailed information about an ongoing topic, to open up a new subtopic or to influence the development of an ongoing topic. Rising questions are most often used to ask for background information, and also to clarify information that is not audible. This pattern of results demonstrates through careful study of the prosody and pragmatics of actual corpus examples that the widespread perception that wh-questions are generally falling, and also that reclamatory wh-questions (e.g. echo questions) are typically rising is justified.

V. ACKNOWLEDGMENTS

REFERENCES

[1] Halliday, M. A. K. 1994. An Introduction to Functional Grammar. 2nd edition. London: Edward Arnold.
[2] Pierrehumbert, Janet; Julia Hirschberg. 1990. The meaning of intonation contours in the interpretation of discourse. In Cohen, P.R.; Morgan, J.; Pollack, M.E. (eds.) Intentions in Communications. MIT Press. 271-311.
[3] Hedberg, Nancy; Juan M. Sosa; Emrah Görgülü. 2008. Early and Late Nuclei in Yes-No Questions: Tails or High Rises? Proceedings of Speech Prosody 2008, Campinas, Brazil, May 2008.
[4] Gussenhoven, Carlos. 1984. On the Grammar and Semantics of Sentence Accents. Dordrecht: Foris.
[5] Bartels, Christine. 1999. The Intonation of English Statements and Questions. A Compositional Interpretation. New York & London, Garland Publishing.
[6] Steedman, Mark. 2007. Information-Structural Semantics for English Intonation. In Chungmin Lee, Matthew Gordon and Daniel Büring (eds.), Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation. Springer: Dordrecht. 245-264.
[7] Bolinger, Dwight. 1989. Intonation and Its Uses: Melody in Grammar and Discourse. Stanford, Stanford University Press.
[8] Canavan, Alexandra; David Graff, and George Zipperlen. 1997. CALLHOME American English Speech. Linguistic Data Consortium. Philadelphia.
[9] Cieri, Christopher; David Graff; Owen Kimball; Dave Miller; and Kevin Walker. 2004. Fisher English Training Speech Part 1 Speech; Transcripts. Linguistic Data Consortium. Philadelphia.
[10] Beckman, M.; Ayers Elam, G. 1997. Guidelines for ToBI labeling, Version 3, Ohio State University.
[11] Hedberg, Nancy; Juan M. Sosa; Lorna Fadden. 2006. Tonal Constituents and Meaning of Yes/No Questions in American English. Proceedings of Speech Prosody 2006, Dresden, Germany.
[12] Steedman, Mark. 2000. Information Structure and the Syntax-Phonology Interface. Linguistic Inquiry 31. 649-689.
[13] Hedberg, Nancy; Juan M. Sosa. 2002. The Prosody of Questions in Natural Discourse. Proceedings of Speech Prosody, Aix-en-Provence, France. 375-378.