

A Prosodic Analysis of Wh-words in Standard Chinese

Fang Hu (PhD Student)

Department of Chinese, Translation and Linguistics, Faculty of Humanities and Social Sciences

Abstract

This paper is a phonetic study of prosody of wh-words in wh-questions, yes/no-questions and echo questions in Standard Chinese (SC). The production data from 4 speakers show that there is a prosodic difference of wh-words and the corresponding VPs between different question types. Wh-words in wh-questions are the focus of sentence, whereas in yes/no-questions, VPs are the focus. The focused constituent is pitch accented so that its lexical tonal melody is retained and sometimes reinforced, while the lexical tonal melody of the corresponding unfocused constituent is compressed and sometimes reduced to a level tone. Speakers usually produce echo questions with a raised F0 register and an expanded F0 range of the wh-words to convey a meaning of surprise. However, no consistent data suggest that either duration or amplitude correlates with the focus vs. non-focus distinction.

Procedure

Subjects and material. Four SC speakers in their late twenties provided the data: two male (M1 and M2) and two female (F1 and F2). M2, F1 and F2 were all born and raised up in Beijing, and M1 was born in Shandong, a Mandarin speaking area as well, and moved to Beijing with his family when he was 15.

Two wh-words, *shui* 'who/anyone' and *shen-me* 'what/anything', are chosen as the target wh-words, because they have the same lexical tonal melody MH, although *shui* is a monosyllabic and *shen-me* is a disyllabic word. Sentences are designed so that the target wh-word occurs either at the beginning, in the middle, or at the end of a sentence. See (1):

- (1) a. *shui lai-le* (MH M-H)
Who/anyone come-ASP
- b. *ni kan-jian shui lai-le* (MH HL-L MH M-H)
You watch-see who/anyone come-ASP
- c. *zhang-san mai-le shen-me* (H-H L-H M-H)
Name buy-ASP what/anything

The above sentences are in three-way ambiguous, namely a wh-question, a yes/no-question, or an echo question. And we distinguish two types of echo question: one is used when the listener is not sure of what the speaker said, the other is used when the listener knows what has been uttered but is surprised to hear it. The sentence can be disambiguated by adding a proper question marker, namely a wh-question marker 'ne' or a yes/no-question marker 'ma', which both have a high level tone. In this way, six different type sentences are designed for each sentence in (1). And each type of question was put into a corresponding dialogue: wh-questions or yes/no-questions were triggered by a proper answer, and echo questions were designed as a respond to a statement. In order to distinguish the two types of echo-question, additional statement was written in parenthesis next to the target question to provide the subjects with enough information.

Recording and analysis. Each designed dialogue was printed on a half-A4 size paper. A total of eighteen dialogues were pseudo-randomized so that one similar dialogue did not appear right after the other in sequence. For each dialogue, one speaker read the first sentence (a question or statement), and the second speaker read the second sentence (an answer corresponding to the question or an echo question corresponding to the statement). The eighteen dialogues were repeated five times by each subject. The recording was made in a soundproof booth with a Sony PCM-R700 Digital Audio Recorder and a Shure SM-58 Microphone. The speech was analyzed using Kay's CSL4400 speech analysis software. In each target utterance, F0 values of the lowest and highest points in the wh-word, the lowest and highest points in the target VP, and the question particle (if there is any) were measured; durations of the wh-word, VP, and question particle (if any) were measured; peak amplitude values of the wh-word, VP, and question particle (if any) were measured.

Results

Because no consistent results were detected on the duration and amplitude data, only the F0 data were discussed in this poster presentation.

WHQs and YNQs. Results show that there is an intonational difference of wh-words and the corresponding VPs between the wh-question (WHQ) and yes/no-question (YNQ). In a WHQ, the wh-pronoun is focused and consequently accented, so that the lexical tonal melody MH is retained and sometimes reinforced while the lexical tonal melodies of the corresponding VP, MH or LH are somewhat weakened and sometimes leveled to a M (or H) and L tone respectively. Typically, in a WHQ, the highest F0 point on wh-words is much higher than that on the corresponding VPs, and the pitch range of wh-words is considerably greater than that of VPs, irrespective of the position of wh-words in a sentence; in a YNQ, the VP has a much higher highest F0 point and a much greater pitch range. Figure 1-3 is an example of F0 contours of the female speaker 2 when producing WHQs and YNQs with or without particles respectively (wh-words in a sentence-initial, sentence-middle, or sentence-final position).

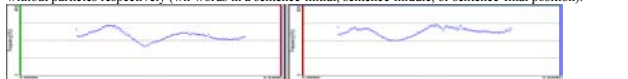


Figure 1: F0 contours of sentences 'shui lai-le ne?' (top left), 'shui lai-le ma?' (top right), 'shui lai-le?' (bottom left), and 'shui lai-le?' (bottom right) by the female speaker 2.

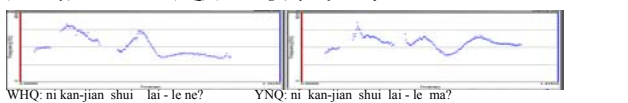


Figure 2: F0 contours of sentences 'ni kan-jian shui lai-le ne?' (top left), 'ni kan-jian shui lai-le ma?' (top right), 'ni kan-jian shui lai-le?' (bottom left), and 'ni kan-jian shui lai-le?' (bottom right) by the female speaker 2.

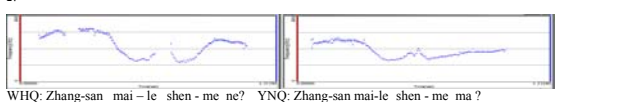


Figure 3: F0 contours of sentences 'zhang-san mai-le shen-me ne?' (top left), 'zhang-san mai-le shen-me ma?' (top right), 'zhang-san mai-le shen-me?' (bottom left), and 'zhang-san mai-le shen-me?' (bottom right) by the female speaker 2.

The detailed F0 results of WHQs and YNQs of four speakers are summarized in table 1 through table 6:

Table 1: Mean F0 values (in Hz, n=5), with SDs in parentheses right after the means, of wh-words and the following VPs in WHQs and YNQs with particles (wh-words in a sentence-initial position)

Sent. Type & subjects	Wh-word		VP		Particle	
	L	H	L	H		
WHQ	M1	146(4)	178(4)	151(8)	174(12)	172(12)
	M2	135(7)	150(11)	127(9)	136(11)	132(10)
	F1	216(16)	289(21)	227(16)	229(19)	220(19)
	F2	216(11)	305(67)	200(22)	266(41)	238(30)
YNQ	M1	137(6)	151(8)	136(14)	179(12)	171(14)
	M2	140(8)	154(8)	127(4)	149(6)	132(3)
	F1	246(25)	285(13)	223(19)	266(13)	239(34)
	F2	256(15)	288(16)	211(12)	295(14)	260(13)

Table 2: Mean F0 values (in Hz, n=5), with SDs in parentheses right after the means, of wh-words and the following VPs in the ambiguous WHQs and YNQs (wh-words in a sentence-initial position)

Sent. Type & subjects	Wh-word		VP		
	L	H	L	H	
WHQ	M1	146(10)	182(6)	156(13)	163(16)
	M2	131(7)	156(10)	128(10)	131(9)
	F1	205(20)	268(29)	180(14)	190(18)
	F2	210(120)	285(29)	178(16)	193(14)
YNQ	M1	138(7)	151(6)	136(11)	182(16)
	M2	132(2)	146(2)	124(5)	149(4)
	F1	240(19)	278(26)	228(22)	289(29)
	F2	237(16)	268(17)	212(8)	307(9)

Table 3: Mean F0 values (in Hz, n=5), with SDs in parentheses right after the means, of wh-words and the following VPs in WHQs and YNQs with particles (wh-words in a sentence-middle position)

Sent. Type & subjects	Wh-word		VP		Particle	
	L	H	L	H		
WHQ	M1	152(9)	184(9)	147(3)	168(5)	160(10)
	M2	133(7)	146(9)	125(7)	131(10)	127(8)
	F1	210(27)	261(16)	196(13)	204(22)	207(13)
	F2	201(16)	271(21)	169(27)	225(35)	203(37)
YNQ	M1	143(8)	158(8)	144(10)	182(11)	177(8)
	M2	131(3)	147(6)	123(7)	147(8)	132(9)
	F1	232(3)	275(7)	216(10)	253(14)	236(23)
	F2	229(25)	264(35)	186(21)	273(35)	242(35)

Table 4: Mean F0 values (in Hz, n=5), with SDs in parentheses right after the means, of wh-words and the following VPs in the ambiguous WHQs and YNQs (wh-words in a sentence-middle position)

Sent. Type & subjects	Wh-word		VP		
	L	H	L	H	
WHQ	M1	138(6)	171(3)	135(10)	135(10)
	M2	126(5)	137(10)	121(6)	129(4)
	F1	190(8)	261(13)	170(5)	176(9)
	F2	198(18)	262(21)	187(30)	209(33)
YNQ	M1	138(11)	161(7)	142(7)	188(6)
	M2	130(5)	146(5)	124(2)	150(5)
	F1	240(22)	285(25)	221(35)	302(39)
	F2	255(10)	272(10)	222(11)	294(6)

Table 5: Mean F0 values (in Hz, n=5), with SDs in parentheses right after the means, of wh-words and the preceding VPs in WHQs and YNQs with particles (wh-words in a sentence-final position)

Sent. Type & subjects	VP		Wh-word		Particle	
	L	H	L	H		
WHQ	M1	115(11)	134(12)	113(6)	152(6)	177(14)
	M2	114(5)	122(5)	116(3)	133(6)	136(6)
	F1	Creaky	202(10)	178(9)	224(7)	253(27)
	F2	133(17)	180(9)	138(13)	231(27)	277(34)
YNQ	M1	90(4)	136(11)	106(2)	106(2)	118(2)
	M2	Creaky	128(4)	118(2)	119(3)	117(3)
	F1	Creaky	215(24)	192(34)	213(23)	238(15)
	F2	138(14)	180(11)	152(7)	192(25)	212(23)

Table 6: Mean F0 values (in Hz, n=5), with SDs in parentheses right after the means, of wh-words and the preceding VPs in the ambiguous WHQs and YNQs (wh-words in a sentence-final position)

Sent. Type & subjects	VP		Wh-word		
	L	H	L	H	
WHQ	M1	107(7)	136(22)	107(4)	157(22)
	M2	115(5)	124(6)	115(8)	142(7)
	F1	Creaky	193(9)	165(12)	206(9)
	F2	144(12)	184(13)	155(10)	264(26)
YNQ	M1	Creaky	134(8)	98(1)	102(8)
	M2	Creaky	128(7)	117(3)	124(3)
	F1	Creaky	232(12)	172(12)	172(13)
	F2	Creaky	237(7)	Creaky	Creaky

Echo questions 1. F0 results of wh-words and the corresponding VPs in echo questions containing the emotion of surprise are summarized below in table 7:

Table 7: Mean F0 values (in Hz, n=5), with SDs in parentheses right after the means, of wh-words and the corresponding VPs in echo questions

Wh-words sent-initial	Wh-word		VP		
	L	H	L	H	
Echo	M1	153(21)	206(26)	193(24)	236(41)
	M2	146(5)	163(8)	145(9)	153(6)
	F1	211(28)	300(33)	232(29)	245(32)
	F2	217(28)	397(54)	308(36)	406(54)
Wh-words sent-mid.	Wh-word		VP		
	L	H	L	H	
Echo	M1	173(8)	217(15)	197(13)	234(25)
	M2	142(5)	168(5)	149(4)	152(5)
	F1	213(28)	327(30)	242(33)	260(38)
	F2	254(33)	392(70)	307(48)	396(110)
Wh-words sent-final	VP		Wh-word		
	L	H	L	H	
Echo	M1	132(11)	142(12)	138(15)	257(23)
	M2	121(2)	124(2)	127(3)	190(11)
	F1	142(27)	194(6)	185(16)	356(24)
	F2	193(17)	206(11)	186(9)	434(57)

Compared with the data of normal WHQs in table 2, 4 and 6, wh-words in echo questions with the emotion of surprise have a much greater pitch range, except of one case where the F0 range is smaller when the speaker M2 produce the wh-word in the sentence-initial position. In addition to the pitch range difference, wh-words in echo questions with the emotion of surprise tend to have a higher F0 register, namely both the highest and the lowest F0 point are usually (although not always) much higher than those in normal WHQs, and this is also true for the exceptional case in terms of F0 range. Moreover, the unfocused VPs also tend to have a much higher F0 register, especially when the wh-word is not in a sentence-final position, and in the speakers of M1 and F2, the highest F0 point of VPs is even much higher than that of the preceding wh-words. As for the case where wh-words are in the sentence-final position, although the highest F0 point of the unfocused VP is sometimes even lower than that in normal WHQs, the creaky voice at the lowest F0 point of VPs, which is common in normal WHQs, only occur twice in the speaker F1, which also indicates the raise of the F0 register. All these suggest that in an echo question, speakers generally tend to raise the pitch register of the whole sentence and expand the F0 range of wh-words to express the emotion of surprise. And the data also suggest that although wh-words in WHQs are the focus and consequently pitch accented, wh-words in echo questions need extra pitch accent to convey the meaning of surprise.

Echo questions 2. For the echo questions used when the listener is not sure of what the speaker said, no consistent F0 data is found in this study. Roughly speaking, speakers sometimes use an F0 pattern more similar to that of a normal WHQ, but sometimes use an F0 pattern more similar to that of an echo question conveying the emotion of surprise.

Conclusion

(1) Wh-words in WHQs are the focus of sentence, whereas in YNQs, wh-words are not the focus, instead the corresponding VPs are the focus, irrespective of the position of wh-words in the sentence.

(2) Data from the all four speakers show that focus is manifested with pitch accent, an effect of intonation. The focused constituent, either wh-words or VPs, retains and sometimes expands its lexical tonal melody, while the corresponding unfocused constituent compresses and sometimes reduces its lexical tonal melody to a level tone.

(3) The presented production data suggest that the possible readings of ambiguous question sentences as WHQs, YNQs and one type of echo questions can be disambiguated by the prosodic differences summarized in (1) and (2).

(4) No consistent data of duration and amplitude are detected which suggest a correlation with the focus vs. non-focus distinction, although speakers sometimes do produce the focused wh-words in WHQs with considerably longer duration than produce the unfocused wh-words in YNQs.

References

- Jun, S.; Oh, M., 1996. A prosodic analysis of three types of wh-phrase in Korean. *Language and Speech* 39(1), 37-61.
- Pierrehumbert, J., 1980. *The phonology and phonetics of English intonation*. Ph.D. Dissertation, MIT.
- Pierrehumbert, J.; Beckman, M., 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.
- Selkirk, E., 1995. Sentence prosody: Intonation, stress, and phrasing. In *Handbook of phonological theory*, J. Goldsmith (ed.), Oxford: Blackwell, 550-569.
- Ladd, D., 1996. *Intonational phonology*. Cambridge, U.K.: Cambridge University Press.
- Gårding, E., 1987. Speech act and tonal pattern in Standard Chinese: Constancy and Variation. *Phonetica* 44: 13-29.