

SIMILAR, BUT NOT THE SAME: NEUTRAL AND CONFIRMATION-SEEKING QUESTIONS IN MANDARIN

Li-Fang Lai

Pennsylvania State University, USA

lk15302@psu.edu

ABSTRACT

In some Germanic and Romance languages, speakers use multiple syntactic and/or prosodic cues to distinguish neutral questions (NQs) from confirmation-seeking questions (CSQs). In Mandarin, this query-check contrast between NQ and CSQ is less clear because both share similar syntactic frames. To provide insights into the issue, this study tested whether the phonetic realizations of NQ and CSQ are different in Orchid Island Mandarin (Taiwan), and if so, which prosodic correlates contribute to the differentiation. Acoustic parameters – F_0 slope, mean pitch height, and speech rate – were measured at the sentence and word levels. At the sentence level, CSQs have a steeper F_0 slope and are higher in pitch than NQs. Plus, words embedded in CSQs were spoken slower than those in NQs. Taking global and local cues together, I argue that in Orchid Island Mandarin, NQ and CSQ each “own” their melodic pattern and are not always interchangeable.

Keywords: prosodic cue, neutral question, confirmation-seeking question, Mandarin, speech rate

1. INTRODUCTION

The speech act of questioning is common to all human languages [5]. Two types of questions are universally found: (a) open questions, often marked by WH-words and have virtually unlimited possible answers, and (b) yes/no questions which only prompt positive or negative responses [26]. Merely treating yes/no questions as a single category, however, is unsatisfactory as it may conceal the intricate interaction between pragmatics, syntax, and prosody in subtypes of yes/no questions, such as neutral questions, confirmation-seeking questions, declarative questions, etc. This necessitates a more fine-grained analysis of different types of yes/no questions to help advance our understanding of language production and the encoding of information [13].

In this study, I restrict attention to **neutral yes/no questions** (questioners are clueless about the answer, henceforth NQs)¹ and **confirmation-seeking yes/no questions** (questioners produce propositions to

confirm their initial guess, henceforth CSQs) [4, 8, 20] in Orchid Island Mandarin (Taiwan). In what follows, I give a review of how the two types of questions are constructed and contrasted in different languages.

1.1. NQ and CSQ: Formation and contrast

In some Germanic and Romance languages, there is a well-established NQ-CSQ distinction. In English, NQs are characterized by subject-verb inversion, whereas canonical CSQs are formed by using a declarative construction with a rising intonation. In other languages, speakers signal the NQ-CSQ contrast through prosodic means: in Leipzig German, NQs present a high boundary tone but CSQs display a low boundary tone [12]; in Bari Italian, NQs are marked by a L+H* accent, while CSQs are signaled through a H*+L accent [10]; in Puerto Rican Spanish, a high peak is associated with the nuclear syllable in NQs but is located in the syllable preceding the nuclear one in CSQs [1]; in Catalan, NQs are signaled by a \uparrow H+L* pitch accent, while CSQs are featured by a H+L* pitch accent [27]; in European Portuguese, a H+L* accent is attested in NQs, while H* and L+H* accents mainly occur in CSQs [21].

To sum up, speakers may use syntactic differences to mark the NQ-CSQ distinction. When syntactic cues are not available, speakers resort to prosodic means to fulfill their communicative intentions.

However, in some languages, the NQ-CSQ contrast is not so clear-cut. For instance, in Mandarin, NQ and CSQ often share similar surface syntactic structure and intonation patterns. Whether an utterance is an NQ or a CSQ depends largely on the discourse-pragmatic context. Thus, whether and how Mandarin speakers employ nuanced prosodic differences to encode question type calls for exploration.

1.2. Yes/No questions in Standard Mandarin

In Mandarin, declaratives are constructed using a Subject-Verb-Object order and have a falling intonation [6, 7, 13, 18, 24]. Regarding yes/no questions, the most widely used types include:

(a) Particle questions – formed by attaching a question particle *-ma/-ba/-ne* to the end of the

declarative [19, 20]. Both NQ and CSQ belong to this category.

- (b) Declarative questions – formed by using a declarative syntactic frame with a rising intonation to express the speaker’s incredulity/surprise.

Researchers reported interaction between pragmatics, syntax, and prosody in utterances: [24] showed that yes/no questions are realized with higher tones than the corresponding declaratives. [6, 7, 13, 18] further added that within yes/no questions, declarative questions are overall higher in pitch and have wider final pitch expansion than particle questions.

As mentioned, in Mandarin, the NQ-CSQ contrast is less clear because both are syntactically marked by *-ma/-ba* and the pragmatic nuances between the two are not always easily detectable (see Table 1). A recent experimental study [13], however, shows that NQs and CSQs end with different final boundary tones (NQ: M% vs. CSQ: L%) and notes that CSQs are impressionistically slower than NQs. This study aims to provide acoustic-phonetic evidence for this.

1.3. Aims of the study

[13] has shown that NQs and CSQs end with different final boundary tones. This paper will investigate whether there are more fine-grained phonetic differences in the realization of F_0 values and speech rates between the two question types.

2. METHODS

The present study is based on the data collected for a larger project investigating Yami-Mandarin bilingual intonation patterns [13, 16, 17]. Particular emphasis was placed on the melodic patterns of Mandarin NQs and CSQs.

2.1. Participants

33 participants (10 males and 23 females, aged between 21 and 60) on Orchid Island, Taiwan, were recruited for the study. The majority of participants (28) reported Mandarin as their primary language in social interactions. Only five elder participants reported Yami (an indigenous language) as their primary language in daily conversations. Even so, the elder participants received six-years of Mandarin schooling and were proficient enough for effective communication.

2.2. Stimuli and corpus collection

Spontaneous Mandarin speech was elicited using The Interactive Card Game [13, 14]. Four sentence types

were elicited: NQ, Declarative, WH-question, and CSQ.

To compare sentence types across pragmatic conditions, the sentence-final lexical contents were kept constant. Six disyllabic target items with identical adjacent lexical tones were chosen: Tone11: [tʰjɛn¹.kʰɔŋ¹] ‘sky’; Tone22: [ʂɣ².tʰoɔ²] ‘tongue’, [xɣ².ljoɔ²] ‘river’; Tone33: [xai³.ʂweɪ³] ‘sea water’, [taɔ³.y³] ‘island’; Tone44: [qœ⁴.ljan⁴] ‘moon’. Ten fillers were also included.

Two interactive games were conducted. **Card-matching task**: participant 1 initiated the conversation by requesting a target card from participant 2. Upon hearing the request, participant 2 checked the randomly shuffled deck of cards in front of him/her to see if (s)he had the intended target participant 1 needed. If yes, (s)he gave the card to participant 1 to facilitate card matching. Participant 1 then specified which target was matched and put the pair aside. After that, participant 2 repeated the same procedure for card matching. This elicited six NQ-Declarative pairs from each participant. **Memory card game**: each participant randomly arranged the six target cards into a six-pocket holder and presented the layout to their partner to review for five seconds. Then, they took turns to ask each other if they could recall the order of the cards from memory. Upon hearing the answer (each item could only be mentioned once), the participants only replied “correct” or “incorrect” to their partner without revealing any further information. This elicited six WH-question-CSQ pairs from each participant.

In total, each participant provided 24 responses. 18 of them contained the six target items (denoted by * in Table 1, the blanks represent the target positions). In this paper, I restrict attention to NQs and CSQs. As seen in Table 1, both NQ and CSQ have five syllables, contain the same target word, and are marked by the question particle *-ma*.

Table 1: Interactive card game dialogues.

Task	Pragmatic Condition	Carrier sentence (tones omitted)
Card-matching	NQ	[ni joɔ ___ ma?]* ‘Do you have ___?’
	Declarative	[zɣ ʂl ___]* ‘This is ___.’
Memory card game	WH-question	1-6 [xao ʂl ʂɣ mɣ?] ‘What is #1-6?’
	CSQ	[na ʂl ___ ma?]* ‘Is that ___?’

2.3. Acoustic analysis

Three parameters – F_0 slope, mean pitch height, and speech rate – were measured at both the global (sentence) and local (target word) levels:

- (a) F_0 slope – defined as the difference between phrase-final and initial F_0 values [22] was calculated for each sentence and target word, to

reveal both direction and steepness of pitch change.

(b) Mean pitch height – the heights of F_0 over the whole sentence and the target word were measured to see if pitch height helps differentiate question type [11].

(c) Speech rate – computed in syllables per second (syll/sec) for each target sentence and word. The temporal dimension was included here because studies have shown that speech rate can be an important cue to sentence/question type [15, 26].

All pitch and duration measurements were done in Praat [3]. F_0 measurements were time-normalized and converted to semitone (st) by implementing the ProsodyPro script [28] to facilitate comparison across speakers.

2.4. Statistical analysis

A linear mixed-effects (LME) model was run, with six dependent variables (F_0 slope, mean pitch height, and speech rate, at the global and local levels). In the model, question type (NQ vs. CSQ) was included as a fixed effect and participant and word as random effects. All statistical modeling was performed in RStudio [23] using the lme4 package [2].

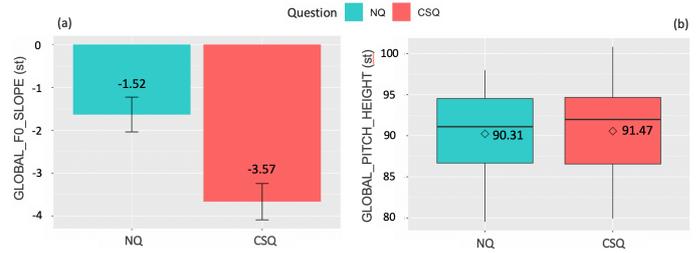
3. RESULTS

Overall, participants performed well on the tasks and had no difficulty distinguishing between NQs and CSQs. In total, 468 target sentences were elicited. Sentences containing overlapping, laughing, clear disfluency/hesitation, or background noise that would obscure pitch and contour information were eliminated. This yielded a smaller dataset containing 344 eligible sentences for acoustic analysis. For clarity, parameters measured at the global level are represented in uppercase; those at the local level, in lower case.

3.1. Global level

At the sentence level, the models found a main effect of question type on F_0 SLOPE and PITCH HEIGHT. Comparing the two questions, F_0 SLOPE is on average 2.05 ± 0.27 (st) steeper in CSQs than in NQs ($t = 7.65, p < .001$) (Figure 1a), and PITCH HEIGHT is higher in CSQs by 1.16 ± 0.19 (st) ($t = 6.19, p < .001$) (Figure 1b). No main effect of question type was found on SPEECH RATE ($t = 0.43, p = .67$).

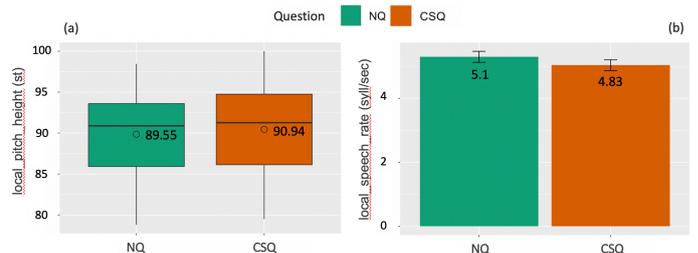
Figure 1: GLOBAL F_0 SLOPE (st) and PITCH HEIGHT (st) by question type.



3.2. Local level

At the word level, the models found a main effect of question type on pitch height and speech rate. Comparing the two questions, pitch height is higher in CSQs by 1.39 ± 0.19 (st) ($t = 7.33, p < .001$) (Figure 2a), and speech rate is slower in CSQs by 0.27 ± 0.08 (syll/sec) ($t = 3.3, p < .01$) (Figure 2b). The effect of question type on f_0 slope was just significant ($t = 1.93, p = .05$).

Figure 2: Local pitch height (st) and speech rate (syll/sec) by question type.



In sum, different cues are used at different levels – GLOBALLY, CSQs have a steeper F_0 SLOPE (st) and are higher in PITCH than NQs. Locally, CSQ words are higher in pitch and were spoken slower than NQ words, as summarized in Table 2.

Table 2: Summary of statistical analyses.

	F_0 SLOPE	MEAN PITCH HEIGHT	SPEECH RATE
GLOBAL	CSQ > NQ ***	CSQ > NQ ***	CSQ \approx NQ n.s.
local	f_0 slope	mean pitch height	speech rate
	CSQ \approx NQ n.s.	CSQ > NQ ***	NQ > CSQ **

4. DISCUSSION

A thorough review of literature on Mandarin yes/no question prosody shows that, compared to NQs and declarative questions, CSQs have received far less scholarly attention, presumably because CSQs share similar syntactic frame with NQs (marked by *-mal-ba*) [19, 20] and both are marked by non-rising terminal tones [13]. A detailed analysis of their F_0 slope, pitch height, and speech rate would help provide some clarity.

The results show that the phonetic realizations of NQ and CSQ are indeed different in Orchid Island Mandarin. Interestingly, speakers employ different prosodic cues, at different levels, to encode the question types. For **F₀ slope**, a significant difference was found at the sentence level as CSQs have a steeper negative F₀ slope than NQs. This is compatible with the findings that Mandarin CSQs end with a low boundary tone (L%) and NQ, a mid-level tone (M%) [13]. For **mean pitch height**, CSQs are higher in pitch than NQs at both the sentence and word level. For **speech rate**, CSQ words were spoken *slower* than NQ words. The speech rate effect, however, was absent at the sentence level as both NQs and CSQs were spoken at a similar rate (see Section 3.1). This necessitates an analysis of the temporal distributions of the two to see how they are locally different.

In the temporal analysis, all sentences were segmented into three fragments: (a) “pre-target” that contains the two syllables preceding the target word, (b) “target word”, and (c) the final particle *-ma* (Table 3). The relative proportions of each fragment against the whole utterance were calculated.

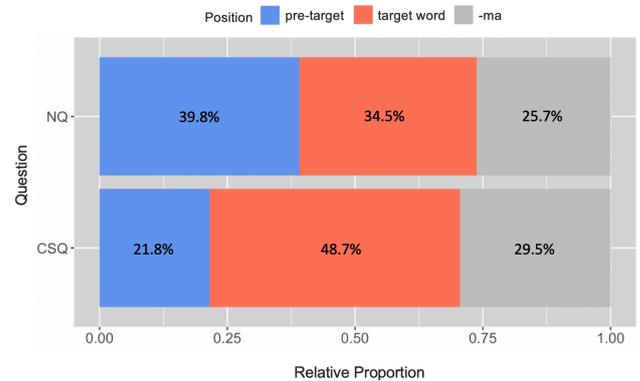
Table 3: Sentence segmentation.

Question	pre-target	target word	-ma	Translation
NQ:	[ni joʊ]	[_____]	[-ma]?	‘Do you have?’
CSQ:	[na ʂl]	[_____]	[-ma]?	‘Is that ___?’

Three dependent variables (the relative proportions of pre-target, target word, and *-ma* segments) were examined using the LME model (see Section 2.4). The models found a main effect of question type on all positions. Comparing the two questions, the pre-target portion is shorter in CSQs by 17.92 ± 1 (%) ($t = 18.92, p < .001$); the target word is longer in CSQs by 14.18 ± 1 (%) ($t = 17.68, p < .001$), and the *-ma* particle is longer in CSQs by 3.89 ± 1 (%) ($t = 5.11, p < .001$).

It is clear that CSQ and NQ have unique melodic patterns, and discourse pragmatics is important to the distinction. In CSQs, the target words are lengthened, and the slower rate may reflect the uncertainty of the pragmatic intent of the speaker [9, 15, 25]. The longer target word (viz. slower speech rate) in CSQs may also in part account for the auditory impression that CSQs are slower than NQs [13]. NQs show a different melodic pattern: the pre-target and target word portions are similar in length and the speech rate is relatively evenly spread over the entire utterance. Schematic presentations of CSQ and NQ are illustrated in Figure 3.

Figure 3: Schematic representations of the melodic patterns of Orchid Island Mandarin: NQ (upper) and CSQ (lower).



In addition to phonetic evidence, [20] mentioned that CSQs and NQs are propositionally and pragmatically different and thus prompt different responses. To illustrate, in Mandarin, the presupposition of CSQs permit “correct/incorrect” responses (as in the Memory card game). In contrast, the neutrality of NQs can only be answered with “yes/no” (as in the Card-matching task). CSQs and NQs, therefore, are not always interchangeable.

5. CONCLUSIONS

This paper investigates whether there are fine-grained phonetic differences in the realization of F₀ values and speech rates between NQs and CSQs in Orchid Island Mandarin. The results show that while being syntactically-similar, the speakers produced pitch and temporal variations to encode sentence type. Specifically, CSQs are characterized by a steeper GLOBAL F₀ SLOPE, are higher in pitch at both the GLOBAL and local levels, and have lengthened target words. NQs are featured with relatively level pitch contour, and the speech rate seems evenly spread over the entire utterance. The two questions are also propositionally and pragmatically different and thus permit slightly different responses. This study, ideally coupled with future perception studies, have the potential to facilitate a better understanding of language production and the encoding of information in communication and interpersonal interactions.

6. ACKNOWLEDGEMENT

I would like to thank my participants for providing me with invaluable research data. I would also like to extend my appreciation to Dr. Shelome Gooden and the anonymous reviewers who provided insightful comments from which this study has greatly benefited. All errors in this paper are my own.

7. REFERENCES

- [1] Armstrong, M. 2010. Puerto Rican Spanish intonation. In P. Prieto & P. Roseano (eds.), *Transcription of intonation of the Spanish language*. München: LINCOM Studies in Phonetics, 155-189.
- [2] Bates, D., Mäeçhler, M., Bolker, B., & Walker, S. 2015. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-8, <http://CRAN.R-project.org/package=lme4>.
- [3] Boersma, P., Weenink, D. 2018. Praat: Doing phonetics by computer (version 6.0.43) [Computer software]. Available from <http://www.praat.org/>.
- [4] Bolinger, D. W. 1989. *Intonation and its uses: Melody in Grammar and Discourse*. Stanford: Stanford University Press.
- [5] Chisholm, W. S., Milic, L. T. 1984. *Interrogativity: A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages*. Amsterdam: Benjamins.
- [6] Chuang, Y.-Y., Fon, J. 2016. Production and perception of incredulity in yes-no question intonation in Taiwan Mandarin. *Proc. 8th Speech Prosody* Boston, paper number 308:1-5.
- [7] Chuang, Y.-Y., Huang, Y.-H., & Fon, J. 2007. The effect of incredulity and particle on the intonation of yes/no questions in Taiwan Mandarin. *Proc. 16th ICPHS* Saarbrücken, 1261-1264.
- [8] Fiengo, R. 2007. *Asking questions: Using meaningful structures to imply ignorance*. New York: Oxford University Press.
- [9] Fujishin, R. 2016. *Natural bridges: A guide to interpersonal communication*. Boston: Pearson Education, Inc.
- [10] Grice, M., Savino, M. 1997. Can pitch accent type convey information-status in yes-no questions? In Alter, K., Pirker, H., Finkler, W. (eds.), *Proc. ACL 97 Workshop on Concept-to-Speech Generation Systems* Madrid, 29-38.
- [11] Huron, D., Shanahan, D. 2013. Eyebrow movements and vocal pitch height: Evidence consistent with an ethological signal. *The Journal of the Acoustical Society of America* 133(5), 2947-2952.
- [12] Kügler, F. 2003. Do we know the answer? Variation in yes-no question intonation. In Fischer, S. van de Vijver, R., Vogel, R. (eds.), *Linguistics in Potsdam 21*. Potsdam: Universitätsbibliothek Publikationsstelle, 9-29.
- [13] Lai, L.-F. 2018a. *Intonation in Contact: Prosodic Transfer and Innovation among Yami-Mandarin Bilinguals*. Dissertation: University of Pittsburgh.
- [14] Lai, L.-F. 2018b. The Interactive Card Game: A new data collection method for intonation study. Poster presented at *Experimental and Theoretical Advances in Prosody 4: Sociolectal and dialectal variability in prosody*. Amherst, USA.
- [15] Lai, L.-F., Gooden, S. 2015. What does the question sound like? Exploring wh- and yes/no interrogative prosody in Yami. *Proc. 18th ICPHS* Glasgow, paper number 790:1-5.
- [16] Lai, L.-F., Gooden, S. 2018a. Intonation in contact: Mandarin influence in Yami. *Proc. 9th Speech Prosody* Poznań, paper number 206:1-5.
- [17] Lai, L.-F., Gooden, S. 2018b. Tonal hybridization in Yami-Mandarin contact. *Proc. 6th Tonal Aspects of Languages* Berlin, paper number 41:1-5.
- [18] Lee, O. J. 2005. *The prosody of questions in Beijing Mandarin*. Dissertation: The Ohio State University.
- [19] Li, C. N., Thompson, S. A. 1982. Mandarin. In Chisholm, W., Milic, L. T., Greppin, J. A. C. (eds.), *Interrogativity: A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages*. Amsterdam: John Benjamins, 47-61.
- [20] Liing, W.-J. 2014. *How to ask questions in Mandarin Chinese*. Dissertation: City University of New York.
- [21] Mata, A. I., Santos, A. L. 2010. On the intonation of confirmation-seeking requests in child-directed speech. *Proc. 5th Speech Prosody* Chicago, paper number 118:1-4.
- [22] O'Rourke, E. 2009. Phonetics and phonology of Cuzco Quechua declarative intonation: An instrumental analysis. *Journal of the International Phonetic Association* 39(3), 291-312.
- [23] RStudio. 2018. RStudio: Integrated development environment for R (version 3.5.1) [Computer software]. Available from <http://www.rstudio.org/>.
- [24] Shen, X.-n. S. 1990. *The Prosody of Mandarin Chinese*. California: University of California Press.
- [25] Starkweather, C. W. 1980. Speech fluency and its development in normal children. In Lass, N. (ed.), *Speech and Language: Advances in Basic Research and Practice, Volume 4*, 143-200.
- [26] van Heuven, V. J., van Zanten, E. 2005. Speech rate as a secondary prosodic characteristic of polarity questions in three languages. *Speech Communication* 47, 87-99.
- [27] Vanrell, M. M., Mascaró, I., Torres-Tamarit, F., & Prieto, P. 2010. When intonation plays the main character: Information-seeking versus confirmation-seeking questions in Majorcan Catalan. *Proc. 5th Speech Prosody* Chicago, paper number 168: 1-4.
- [28] Xu, Y. 2013. ProsodyPro – A tool for large-scale systematic prosody analysis. *Proc. Tools and Resources for the Analysis of Speech Prosody* Aix-en-Provence, 7-10.

¹ In this paper, neutral questions are equivalent to information-seeking questions.