When is a *wh*-in-situ question identified in standard Persian?

Zohreh Shiamizadeh[1,2], Johanneke Caspers[1,2], Niels O. Schiller[1,2]
Leiden University Center for Linguistics[1], Leiden Institute for Brain and Cognition[2]

ABSTRACT

Wh-in-situ questions are typical of sentences which contain temporary syntactic ambiguity. The current research adopted the gating paradigm [7] to investigate when distinctive prosodic cues of the pre-wh part enable identification of wh-in-situ questions in standard Persian. A forced-choice sentence identification task was designed in which 40 pairs of gated wh-in-situ questions and declaratives were played to 20 Persian native speakers. As we hypothesized, correct identification responses were given from the first gate (75%). The statistical analysis demonstrated that response accuracy to declaratives is higher and reaction time to declaratives is shorter than to questions. It was also revealed that response accuracy, and confidence rating increases and reaction time decreases as the gate number increases.

The results corroborate proposals on the role of prosody in processing syntactic ambiguity [22,8], and suggest that the role of prosody in the identification of Persian wh-in-situ questions is earlier and richer in comparison to other languages.

Keywords: Persian, wh-in-situ questions, prosody, gating

1. INTRODUCTION

Prediction in speech comprehension is of great importance because it can indicate the sentence type before the end of the sentence and thus accelerate sentence processing and response preparation [8,9]. Previous literature demonstrated the influential role of prediction in processing speech [4,8,9,22], and of prosody in predicting the eventual syntactic structure of ambiguous sentences (e.g. [1,2,5,14,16,22,27]. The role of prosody in processing becomes prominent when other sources of information, such as syntactic information regarding the clause type, are absent from the utterance [8,9]. Wh-in-situ question is a type of interrogative which typically has local syntactic ambiguity. One language characterized by in-situ wh-questions is Persian (e.g. [13]) (see 1b). This study adopted the gating paradigm to investigate the role of prosody in processing and prediction of sentence type in Persian wh-in-situ questions vs. declaratives. The gating technique enables to limit the amount of information input by controlling for the temporal presentation of the acoustic signal. This property helps to determine when in the signal the discriminant acoustic information is accessible to feed the process of comparing competitors[1] and possibly lead to the correct prediction of the target [1]. The gating technique also helps to assess whether prediction improves as the listener progresses through the signal [8,10].

(1) a.	mærjæm	ʔæsr	bɑzi- kærd.
	Maryam	afternoon	play-do.PAST.3SG.
	"Maryam played in the afternoon."		
b.	mærjæm	kej	bɑzi- kærd ?
	Maryam	when	play-do.PAST.3SG.
	"When did Maryam play?"		

2. BACKGROUND

In a production experiment, [21] found that a greater pitch excursion of the pitch accents realized on the pre-wh words, a higher level of pitch mean, a higher F0 onset, a shorter duration, and a steeper inclination of the pitch contour of the pre-wh part contribute to the prosodic distinction of the pre-wh part in Persian wh-questions vs. declaratives. In a later perception study by [20], it turned out that the prosodic correlates of the pre-wh part of a sentence can cue correct identification of Persian wh-in-situ questions as opposed to declaratives in the absence of the wh-phrase at the sentence-initial position (response accuracy is 90.3%). This result raises a new question: where in the pre-wh part does the relevant distinctive prosodic information become available to feed the process of sentence type prediction? The current study aims to answer this question.

The previous gating studies mainly concentrate on yes-no questions and declarative questions as globally ambiguous sentences in languages other than Persian (e.g. [6,17,18,24,25]). As to our knowledge, the gating studies on prosody-driven perception of wh-in-situ questions vs. declaratives are limited to Mandarin Chinese [11,28]. This suggests that wh-in-situ is an understudied question type with respect to the role of prosody in decoding it in comparison to its matching declarative. Furthermore, Persian is an understudied language with respect to the role of prosody in sentence type identification and there is no gating study on the perception of Persian wh-in-situ questions vs. declaratives. This study fills this gap by adopting the gating paradigm to investigate the role of prosody in processing Persian wh-in-situ questions.

3. RESEARCH QUESTIONS, APPROACH AND HYPOTHESES

At what point can Persian native speakers use prosodic correlates to predict wh-in-situ questions before the wh-phrase is made audible? To answer this question, a forced-choice sentence identification task is designed, in which the gating method of stimuli presentation is applied. Twenty Persian native speakers listened to the gated pre-wh part of 20 wh-in-situ questions and 20 declaratives. After hearing each gate, participants had to decide as quickly as possible which sentence type the stimulus in the gate was extracted from, i.e. a declarative or a wh-question, and to indicate how confident they were about their response on a five-point scale.

Prosodic correlates differentiate wh-in-situ questions from declaratives from the beginning of the sentence (cf. §1). Thus, based on the assumption that listeners have the implicit knowledge of the correspondence between sentence type and prosody and are able to use it to process spoken utterances [22], we hypothesize that Persian native speakers can start sentence type prediction from the beginning of the sentence. Along the same lines, we predict that identification improves and confidence rating increases as the amount of discriminating prosodic information increases.

4. METHOD

4.1. Participants

Ten male and ten female Persian native speakers participated in this study. None of them reported any hearing impairment.

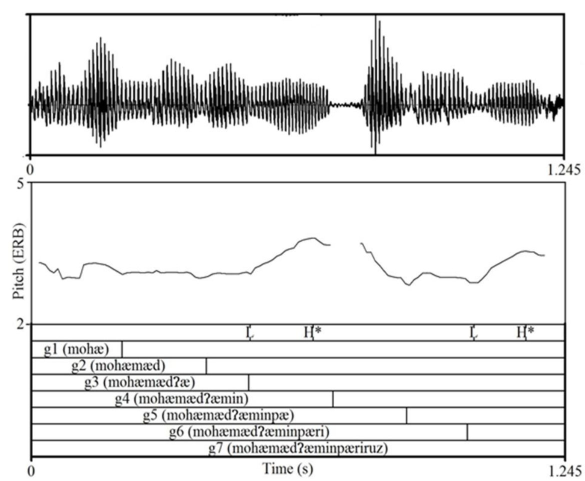
4.2. Materials

The pre-wh part of forty pairs of sentences elicited from a male and a female speaker in the production experiment by [21] comprises the stimuli of this experiment. To make the results of the current experiment more generalizable, we selected a male and a female speaker whose mean value of the acoustic measurements that characterize Persian wh-questions in the study by [21] (cf. §2) were closest to the mean value of the acoustic measurements in the production of all speakers. The structure of the wh-question and declarative stimuli of the production experiment is illustrated in (2) and (3), respectively.

(2) Subj[2]	Adv	Wh-phrase	Verb
(3) Subj	Adv	DWCs	Verb

The total number of the stimuli of this experiment equals 320 (1 subject x 1 adverb x 2 DWCs x 5 wh-phrases and the matching verbs x 2 sentence types x 2 speakers x 8 gates). Although only the pre-wh part of the sentences forms the stimuli of the current experiment, variation in the DWCs, the wh-phrases and their matching verbs are included in the formula to clarify how we arrived at 320 stimuli. The number of wh-questions and their matching declaratives was the same across wh-phrases. The pre-wh part of the sentences was separated from the remaining part of the sentence in Praat version 6.0.04 [3], and truncated into seven gates based on the number of the syllables it contained. The first gate contained the first two syllables of the pre-wh part (see 4 and Figure 1). One syllable was added at the following gates such that each gate contained the previous gate(s) plus one more syllable, e.g. gate 2 includes gate 1 plus the third syllable.

Figure 1. The seven gates of a declarative stimulus. The "L" and "H*" represent the valleys and the peaks of the realized pitch accents. The other tiers represent the gate boundaries. The letter g represents the word gate and the number designates the gate number.



(4) a. mohæmædʔæmin				pæriruz			
"Mohammadamin				two days ago"			
b. mohæ	|mæd	|ʔæ	|min	|pæ	|ri	|ruz	
gate1	|gate2	|gate3	|gate4	|gate5	|gate6	|gate7	

At gate 7 the pre-wh part which is ambiguous with regard to sentence type is completely presented (see 4 and Figure 1), at. The complete unambiguous version (CUV) of each item (see 5) was played at gate 8. However, it was not presented immediately after gate 7 (i.e. the pre-wh part) of the corresponding item. All of the CUVs of the items were presented at the end of the experiment after the first seven gates of all stimuli were played to the participants. The reason for doing this is that hearing the CUV of an item immediately after hearing the pre-wh part of the same item can provide participants with the opportunity to make an association between the prosody of the pre-wh part and the sentence type; thereby it can be practice for the participants in identifying the sentence type.

(5) a. mohæmædʔæmin pæriruz ʔæsr ʃenɑ- kærd.
Mohamadamin two days ago afternoon swim- do.PAST.3SG
'Mohammadamin swam in the afternoon two days ago.'
b. mohæmædʔæmin pæriruz kej ʃenɑ- kærd?
Mohamadamin two days ago when swim- do.PAST.3SG
'When did Mohamadamin swim two days ago?'

4.3. Procedure

A forced-choice sentence categorization task was designed in E-prime 2.0.10 [19]. Participants were seated in front of a computer in a quiet room. The stimuli were played to the participants through Sennheiser PC 141 Headset headphones. They were instructed to decide whether what they hear is going to be a wh-question or a declarative by pressing either V or M on the keyboard within four seconds after hearing the stimulus. Then they were asked to indicate how confident they were about their response on a five-point confidence scale, where one means "not sure at all" and five "completely sure" within four seconds. If participants did not give a response within four seconds, the experiment proceeded to the next stimulus automatically after two seconds. They were familiarized with the task by means of a practice session, including two non-experimental items. For each sentence, the first seven gates were played right one after the other in increasing order. When all seven gates of an item were played, the first gate of the next item was presented. Having accomplished the practice session, participants embarked on the main part of the experiment when they felt ready. The main session of 320 items was divided into five blocks. The final block contained the CUV of the items presented in the previous four blocks. Participants were instructed to take at least a three-minute break after each block. Every block started with a warm-up which consisted of two non-experimental items. The order in which the first four blocks were presented was randomized per participant. However, the fifth block was always presented at the end of the experiment to avoid a practice effect on sentence modality identification (see §4.2). The presentation order of the items within all blocks were randomized per participant. The procedure of the main session was identical to that of the practice session. The experiment took about 40 minutes.

4.4. Data analysis

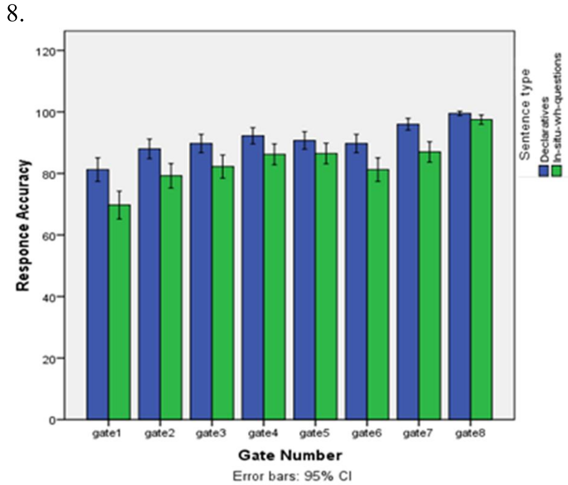
The response accuracy to declaratives and wh-questions was computed in terms of percentage correct and A′ [23]. Reaction times (RT) (in seconds) were calculated in terms of the time lapse between the stimulus offset and the response. Three separate two-way repeated measures ANOVAs (RM-ANOVA) were run on the accuracy, RT and confidence rating data to investigate the effect of sentence type, gate, and their interaction. The assumptions of the RM-ANOVA were met.

5. RESULTS

5.1. Response accuracy

Figure 2 represents the accuracy of sentence type perception for each sentence type across gates, indicating that response accuracy to declaratives is higher than response accuracy to questions. Mean response accuracy to questions and declaratives at gate one (75.5%) is above chance level (t (19) = 29.417, $p < 0.01$). Responses are transformed to A′ to correct for a possible response bias [23]. A two-way RM-ANOVA demonstrated that the main effect of gate (F (7,13) = 12.249, $p < .001$; Wilks' Lambda = .135, η_p^2 = .865) and sentence type (F (1,19) = 7.577, $p < .02$; Wilks' Lambda = .715, η_p^2 = .285) is significant.

Figure 2. Mean response accuracy across gates and sentence type. As mentioned in §4.3 the complete unambiguous version of the stimuli were presented at gate 8.



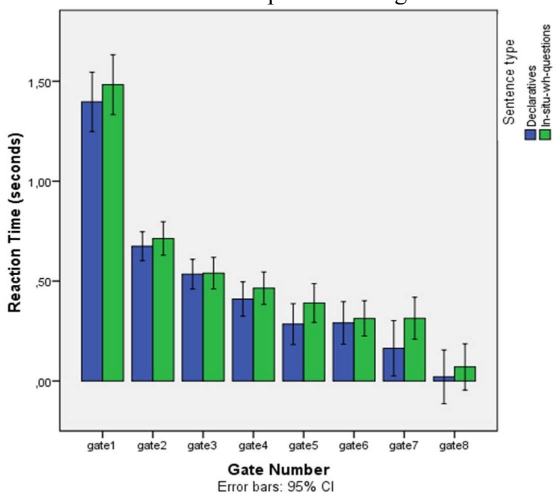
5.2. Reaction time analysis

The RT to declaratives was shorter than the RT to wh-questions within each gate (see Figure 3). The RT to stimuli decreases as the gate number increases, likely reflecting the increased availability of prosodic information as the gate number increases. A two-way RM-ANOVA showed that sentence type (F (1,19) = 11.583, $p < .01$; Wilks' Lambda = .621, η_p^2 = .379), gate (F (7,13) = 38.080, $p < .001$; Wilks' Lambda = .047, η_p^2 = .953) and the interaction of sentence type and gate (F (7,13) = 4.512, $p < .01$; Wilks' Lambda = .292, η_p^2 = .708) significantly affected RT. Pairwise comparison tests revealed that the difference between RT to all gates is significant ($p < .05$) except for the difference between gate 5 and 6 ($p > 0.5$).

5.3. Confidence rating

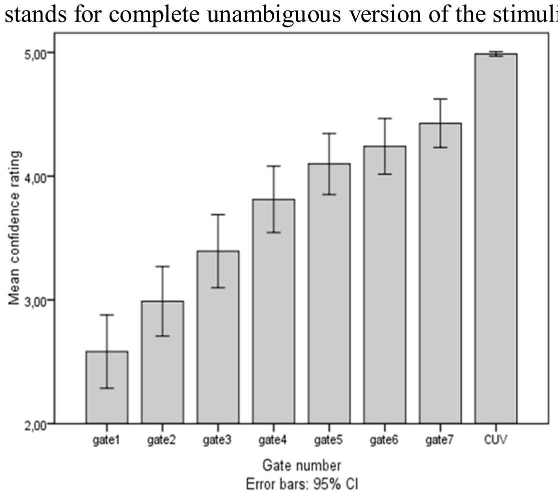
As observable in Figure 4, participants' confidence in their responses increased as the gate number also increased.

Figure 3. Mean reaction time across gates and sentence type. As mentioned in §4.3 the complete unambiguous version of the stimuli were presented at gate 8.



According to a two-way RM-ANOVA, the main effect of gate (F (7,13) = 20.872, p < .001; Wilks' Lambda = .082, η_p^2 = .918) was revealed to be significant. Pairwise comparison tests indicated that the difference between all gates with respect to confidence rating is significant p < .01.

Figure 4. Mean confidence rating across gates. CUV stands for complete unambiguous version of the stimuli.



6. DISCUSSION

The results confirm our hypotheses that listeners may start sentence type prediction from the first gate (75.5%) and identification improves as the amount of discriminating prosodic information increases. The highest confidence rating in the pre-wh part (the first seven gates which present syntactically ambiguous part of the sentence) is achieved at gate 7 (4.45). This implies that prediction can be possibly reset as the listener progresses through the acoustic signal which is in line with sentence processing theories (e.g. [12,15]). In line with earlier studies [20,25], response accuracy to declaratives is higher than response accuracy to questions. A possible reason for the decreased RT and the higher response accuracy to declaratives could be the higher frequency of declaratives in comparison to questions in daily conversation (as suggested by [24,25]).

The general result of this research corroborates several proposals suggested in the literature. First, prosody plays a prominent role in processing syntactically ambiguous sentences (e.g. [1,2, 5,14,16,22,27]) and models of spoken sentence processing may need to integrate the (online) use of prosody in interpreting these constructions [1]. Second, interlocutors may share the implicit knowledge that there is a syntax-prosody correspondence and draw on this knowledge to resolve the ambiguity of syntactically ambiguous sentences [22]. Third, prediction can be reset as more prosodic information is provided to the listener [8].

The first distinctive prosodic feature in Persian wh-in-situ questions is the F0 onset [21] while similar studies on the role of prosody in characterizing interrogatives in other languages do not report significant differences between the F0 onset of questions and statements (e.g., [6,28]). Upon the presentation of the first distinctive prosodic feature, the accuracy of sentence type identification in this study is higher than the accuracy of identification of interrogatives vs. declaratives in other languages (e.g. [6,24,28]). Therefore, it can be proposed that the effect of prosody on the identification of Persian wh-in-situ questions vs. declaratives is earlier and richer in comparison to its effect on the identification of interrogatives vs. declaratives in some of the languages reported in the literature.

7. CONCLUSION

The current experiment was purported to tackle the question at what point Persian native speakers can use prosodic correlates to predict wh-in-situ questions before the wh-phrase is made audible. The general conclusion is that the distinctive prosodic correlates to sentence modality contrast enable participants to predict the sentence type early in the utterance.

The current experiment does not provide direct evidence for the role that prosody plays during online language processing. Neurolinguistic research techniques, such as electroencephalography (EEG), are needed to examine the online use of prosody in the identification and processing of wh-in-situ questions. Due to its fine-grained time resolution, EEG experiments can give additional insights into the time course of prosody processing.

Acknowledgement. The experiment reported in this paper is related to the project "Understanding Questions" funded by the Netherlands Organization for Scientific Research (NWO). For discussion, the authors would like to thank the PIs of that project, Lisa Lai-Shen Cheng and Jenny Doetjes.

8. REFERENCES

[1] Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language, 30*, 644-663.

[2] Beach, C. M., Katz, W. F., & Skowronski, A. (1996). Children's processing of prosodic cues for phrasal interpretation. *Journal of the Acoustical Society of America, 99*, 1148–1160.

[3] Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer (Version 6.0.04) [Computer program]. Available from http://www.praat.org/.

[4] Brazil, D. (1981). The place of intonation in a discourse model. In C. Malcolm & M. Montgomery (Eds.), *Studies in discourse analysis* (pp. 146-157). London: Routledge & Kegan Paul.

[5] Carlson, K., Clifton, C., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language, 45*, 58–81.

[6] Face, T. L. (2005). F0 peak height and the perception of sentence type in Castilian Spanish. *Revista Internacional de Lingüística Iberoamericana, 2*, 49-65.

[7] Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics, 28*, 267-283.

[8] Grosjean, F. (1983). How long is the sentence? Prediction and prosody in the on-line processing of language. *Linguistics, 21*, 501-529.

[9] Grosjean, F. (1996a). Using prosody to predict the end of sentences in English and French: Normal and brain damaged subjects. *Language and Cognitive Processes, 11*, 107-134.

[10] Grosjean, F. (1996b). Gating. *Language and Cognitive Processes, 11*, 597-604.

[11] Gryllia, S., Yang, Y., Pablos, L., Doetjes, J., & Cheng, L. L. S. (2016, September). *Prosody as a means to identify clause type: a view from Mandarin*. Poster presented at Architecture and Mechanisms for Language Processing (AMLap), Bilbao.

[12] Hale, J. (2001). A probabilistic early parser as a psycholinguistic model. *Proceedings of the second meetings of the North American Chapter of the Association for Computational Linguistics (NAACL)* (pp. 1-8). Pittsburgh: Association for Computational Linguistics.

[13] Karimi, S. (2005). *A minimalist approach to scrambling: Evidence from Persian*. Berlin: Mouton de Gruyter.

[14] Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language, 40*, 153–194.

[15] Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126–1177.

[16] Nagel, H. N., Shapiro, L. P., & Nawy, R. (1994). Prosody and the processing of filler-gap sentences. *Journal of Psycholinguistic Research, 23*, 473–485.

[17] Petrone, C., & D'Imperio, M. (2011). From tones to tunes: Effects of the f0 prenuclear region in the perception of Neapolitan statements and questions. In S. Frota., G. Elordieta & P. Prieto (Eds.), *Prosodic categories: Production, perception and comprehension* (pp. 207-230). Dordrecht: Springer.

[18] Petrone, C., & Niebuhr, O. (2014). On the intonation of German intonation questions: The role of the prenuclear region. *Language and Speech, 57*, 108-146.

[19] Psychology Software Tools. (2012). E-Prime. Pittsburg, PA.

[20] Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2017). The role of prosody in the identification of Persian sentence types: declarative or wh-question? *Linguistics Vanguard*. doi: 10.1515/lingvan-2016-0085.

[21] Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2018). Do Persian native speakers prosodically mark wh-in-situ questions? *Language and Speech*. doi: https://doi.org/10.1177/0023830917753237.

[22] Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language, 48*, 103-130.

[23] Stanislaw H., & Todorov, N. (1999). Calculation of signal detection theory measures. Behavior Research Methods, Instruments, and Computers, 31, 137-149.

[24] Van Heuven, V., & Haan, J. (2000). Phonetic correlates of statement versus questions intonation in Dutch. In A. Botinis (Ed.), *Intonation, analysis, modeling and technology* (pp. 119–143). Dordrecht: Kluwer.

[25] Vion, M., & Colas, A. (2006). Pitch cues for the recognition of yes-no questions in French. *Journal of Psycholinguistics Research, 35*, 427-445.

[26] Wales R., & Taylor, S. (1987). Intonation cues to questions and statements: How are they perceived? *Language and Speech, 30*, 199-211.

[27] Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes, 10*, 457-486.

[28] Yang, Y. (2018). *The two sides of wh-indeterminates in Mandarin: A prosodic and processing account* (Doctoral dissertation, Leiden University). LOT: Utrecht.

---

[1] In this paper, the competitors are statements and wh-questions in Persian.

[2] Subject is abbreviated as Subj, adverb as Adv, and the words which replace the wh-phrase in the declarative sentences are named declarative wh-counterparts and abbreviated as DWCs. DWCs are animate direct object, inanimate direct object, adjunct of time, adjunct of manner, and adjunct of place.