

Zohreh Shiamizadeh¹ / Johanneke Caspers¹ / Niels O. Schiller²

The Role of Prosody in the Identification of Persian Sentence Types: Declarative or Wh-question?

¹ Leiden University Centre for Linguistics, Leiden University, 2311 BX, Leiden, Netherlands, E-mail: zshiamizadeh@gmail.com² Leiden University Centre for Linguistics, Leiden University, 2311 BX, Leiden, Netherlands**Abstract:**

It has been reported that prosody contributes to the identification of utterances which lack lexico-syntactic indicators of interrogativity but do have characteristic prosodic correlates (e.g. Vion and Colas 2006. Pitch cues for the recognition of yes-no questions in French. *Journal of Psycholinguistics Research* 35. 427–445). In Persian wh-in-situ questions, the interrogativity device (the wh-phrase) does not move to the sentence-initial position, and the pre-wh part is characterized by specific prosodic correlates (Shiamizadeh et al. 2016. *Do Persian native speakers prosodically mark wh-in-situ questions?* Manuscript submitted for publication). The current experiment investigates the role of prosody in the perception of Persian wh-in-situ questions as opposed to declaratives. To this end, an experiment was designed in which Persian native speakers were asked to choose the correct sentence type after hearing only the pre-wh part of a sentence. We hypothesized that prosody guides perception of wh-in-situ questions independent of wh-phrase type. The results of the experiment corroborate our hypothesis. The outcome is discussed in terms of Ohala's frequency code, and Bolinger's claim about the universal dichotomous association between relaxation and declarativity on the one hand and tension and interrogativity on the other hand.

Keywords: Persian, wh-in-situ questions, declaratives, prosody, sentence type identification**DOI:** 10.1515/lingvan-2016-0085**Received:** October 10, 2016; **Accepted:** March 17, 2017

1 Introduction

Wh-questions are expressions that use a wh-phrase to enquire about desired information. They can be divided into two groups: fronted and wh-in-situ questions. In fronted wh-questions the wh-phrase moves to the beginning of the sentence to form a wh-question (see example 1) whereas in wh-in-situ questions the wh-phrase does not move to the sentence-initial position (Carnie 2007; Chomsky 1977). One of the languages which is characterized by wh-in-situ is Persian (Abedi et al. 2012; Adli 2007; Gorjian et al. 2012; Kahnemuyipour 2009; Karimi 2005; Karimi and Taleghani 2007; Lotfi 2003; Megerdooonian and Ganjavi 2000; Mirsaedi 2006; Toosarvandani 2008). In Persian, wh-questions are in-situ by default; the wh-phrase does not need to move to the beginning of the sentence, rather it occurs at the same site where its declarative counterpart is expected to occur (see example 2).

- (1)
- a. Mary carries a book.
 - b. What does Mary carry?

(2)

a.

<i>Mærjæm</i>	<i>diruz</i>	<i>ketab</i>	<i>xærid.</i>
Maryam	yesterday	book	buy.PAST.3SG.

‘Maryam bought a book yesterday.’

b.

<i>Mærjæm</i>	<i>diruz</i>	<i>tʃi</i>	<i>xærid?</i>
Maryam	yesterday	what	buy.PAST.3SG.

‘What did Maryam buy yesterday?’

In fronted *wh*-questions, the listener can discern the sentence type as soon as the *wh*-phrase is uttered by the speaker. However, in *wh*-in-situ questions, the syntactic cue to sentence type occurs later in the sentence. In a conversation, the purpose of asking a question is to elicit a verbal response from the hearer. The listener needs to be made aware of this purpose as soon as possible to avoid leaving gaps between exchanges (Bolinger 1981). This suggests that when the syntactic interrogativity device occurs later in the sentence, as in *wh*-in-situ questions, listeners may draw upon other cues such as prosody to discern the sentence type before the *wh*-phrase is uttered by the speaker.

1.1 Background

1.1.1 Prosodic correlates of Persian *wh*-in-situ questions

In a previous study, Shiamizadeh et al. (2016) compared prosodic correlates of *wh*-in-situ questions and their declarative counterparts. They found that a higher level of pitch register, a higher F0 onset and a shorter duration distinguished the pre-*wh* part in *wh*-questions from declaratives. A steeper inclination of the F0 contour and a greater excursion size of the pre-*wh* words were two additional features that give rise to the prosodic markedness of the pre-*wh* part in *wh*-questions. Figure 1 and Figure 2 present an example of the pitch track of the data.

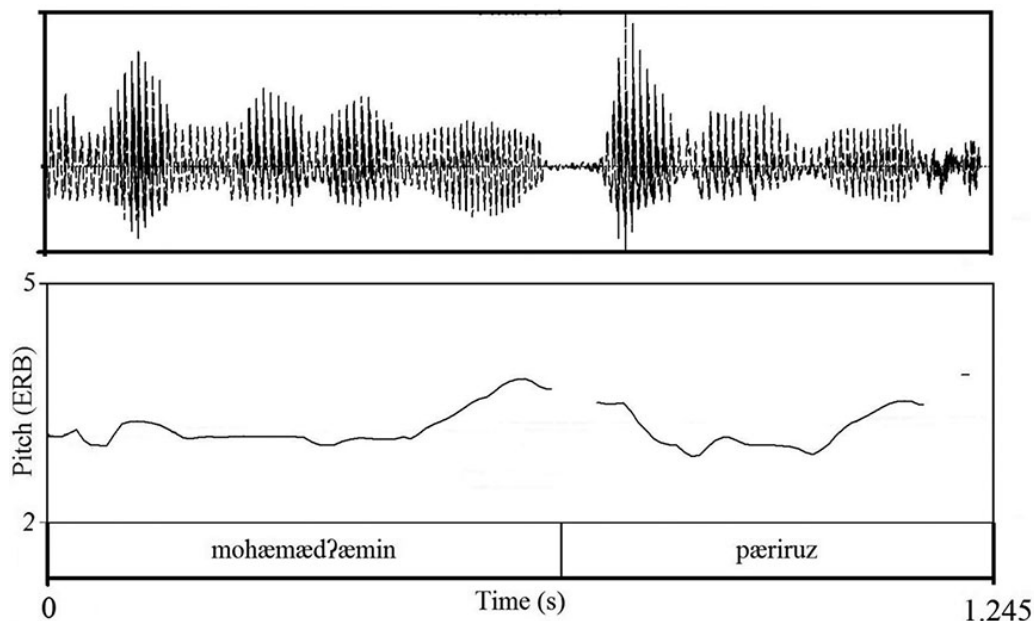


Figure 1: The pitch track of the pre-*wh* part of a declarative sentence. The solid line is the pitch contour. The tier represents the word boundaries. The first word is the subject and the second word is the adverb of time.

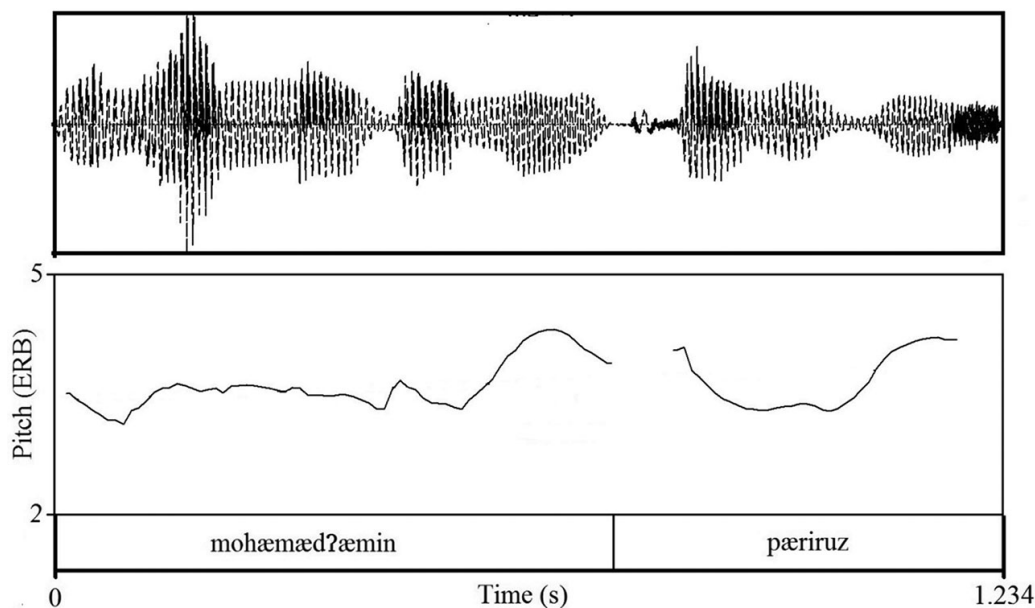


Figure 2: The pitch track of the pre-wh part of a wh-question. The solid line is the pitch contour. The tier represents the word boundaries. The first word is the subject and the second word is the adverb of time.

1.1.2 Theoretical background

Hermann (1942) claimed that a greater pitch height in questions can be regarded as a universal property of language. While it is now understood that this is not universal, it is quite common cross-linguistically for questions to have a greater pitch height.¹ Ohala (1984) explained the use of high pitch in questions in ethological terms as a phonologized remnant of animal behavior. In comparison to large (dangerous) creatures, small (harmless) creatures have higher pitches, and make faster movements. The questioner needs to be polite or subservient to the respondent as he is attempting to get the respondent to provide information. Hence, the speaker raises his/her pitch to pretend to be smaller than he/she physically is. The raised pitch is considered as a sign of submission or politeness to the hearer. The faster speech rate (shorter duration of utterances) may express the same message conveyed by high F0 to the hearer. It suggests that high pitch is tied with fast rate in signaling questions (Van Heuven and Van Zanten 2005). The association of high pitch with questions can be alternatively accounted for by Bolinger's (1989) claim that statements and questions are universally characterized by a dichotomy between relaxation (low, falling pitch) and tension (high, rising pitch), respectively.

According to the results of the previous studies on the relation between prosody and syntactic ambiguity resolution (e.g., Beach et al. 1996; Carlson et al. 2001), Snedeker and Trueswell (2003) proposed that native speakers possess an implicit knowledge on the relation between prosody and syntax and are capable of using this knowledge to guide their linguistic choices. This suggests that listeners possess implicit knowledge on the relation between prosody and sentence type and can draw on this knowledge to identify the sentence type.

1.1.3 Empirical background

To our knowledge, there are a limited number of sentence identification studies that investigated the role of prosody in the identification of interrogatives vs. declaratives (Baltazani et al. 2015; Sensui 1995; Truckenbrodt et al. 2009). Plain yes-no questions and statements in Greek are the same, differing only in intonation (Baltazani et al. 2015).² Higher level and earlier alignment of the peak of the pitch accent, a steeper slope of rise of the pitch accents, down-stepping of consecutive pitch accents (Baltazani et al. 2015), and L* nuclear pitch accent (Baltazani 2007) set plain yes-no questions apart from statements. The nuclear pitch accent (NPA) which is the last pitch accent in the sentence has a different tonal structure in declaratives and questions. (Baltazani et al. 2015) showed that Greek listeners can identify yes-no questions from statements drawing on the acoustic differences before hearing the NPA in 66.6% of the cases. The NPA was not made audible.

Spanish does not syntactically distinguish yes-no questions from declaratives. Sensui (1995) used filtered utterances as stimuli to investigate the role of the intonation pattern in the perception of interrogatives versus declaratives in Spanish. His experiment demonstrated accurate perception of sentence type on the part of Spanish native speakers based on only the prosodic information available in the stimuli.

Since neither syntax nor morphology regularly mark the distinction between yes-no questions and statements in Brazilian Portuguese (BP), the distinction between these sentence types chiefly rests on intonation (Truckenbrodt et al. 2009). According to De Moraes (1998), the opposition between declaratives and yes-no questions is mainly manifested in the final tonic. Final tonic is the last and the main³ pitch accent in the intonational phrase. It is called nuclear by Truckenbrodt et al. (2009). Focusing on the nuclear contour, Truckenbrodt et al. (2009) notated the statement contour as H + L* L% and the question contour as L + H* L%. In a perception study, Truckenbrodt et al. (2009) found that the intonational differences between questions and statements can cue recognition of questions vs. statements.

There are a number of studies which adopt the gating paradigm to investigate whether and how prosody guides identification of interrogatives as opposed to declaratives in Spanish, Neapolitan Italian, Northern Standard German, Dutch and French (Face 2005; Petrone and D'Imperio 2011; Petrone and Niebuhr 2014; Van Heuven and Haan 2000; Vion and Colas 2006).⁴ All of these investigations focused on yes-no questions or declarative questions. This section briefly reviews the results of these studies.

Castilian Spanish yes-no questions are not syntactically distinguished from declaratives. According to Face (2004), a raised F0 peak of pitch accents and a final F0 rise are the prosodic characteristics of yes-no questions in Castilian Spanish. A further prosodic feature that disambiguates yes-no questions from declaratives in this language is the presence vs. absence of the pitch accents; in declaratives every stressed word is associated with a pitch accent, while in questions only the first and the last stressed words are associated with pitch accents. Face (2005) designed a study employing gating paradigm to investigate whether the acoustic cues of prosody allow for correct perception of the sentence type. The results of his experiment showed that native speakers can correctly distinguish declaratives from yes-no questions in 95 % of the cases when the first prosodic distinction (the height of the initial F0 peak) occurs. Participants could perform with 100 % accuracy when the final rise was made audible.

The distinction between yes/no questions and statements in Neapolitan Italian rests on intonation only (D'Imperio 2000). According to Petrone and D'Imperio (2008), F0 peak of the NPA is aligned later in questions than in statements (L + H* in questions vs. L* + H in statements). F0 fall after the peak of the pitch accent preceding the NPA is shallower whereas the F0 rapidly falls from the peak of the prenuclear pitch accent to the end of the accented prosodic word. The boundary tone of both sentence types is L-L%. In a perception study based on gating paradigm, Petrone and D'Imperio (2011) investigated contribution of prenuclear region to sentence type categorization in Neapolitan Italian. The results revealed that prosody of the prenuclear region cue question identification (68 %) and accentual phrase boundary tone contributes significantly to question identification. Robust question recognition was achieved (above 90 %) upon the presentation of the complete sentence.

German questions can be signaled lexically, syntactically and intonationally (Petrone and Niebuhr 2014). According to Petrone and Niebuhr (2014), questions are not necessarily marked by a H% boundary tones in Northern Standard German. Rather, they can have an L% similar to statements. However, similar to Neapolitan Italian, there are prosodic differences between statements and questions in the area of the pitch accent preceding the NPA. Independent of the direction of the final F0 movement in questions, the rise of the prenuclear accent and its F0 peak are aligned later and its subsequent F0 fall takes longer and is less steep in questions. In a perception experiment based on gating method, Petrone and Niebuhr (2014) found that F0 differences in the prenuclear pitch accent region significantly contribute to identification of questions vs. statements in Northern Standard German.

According to Di Cristo and Hirst (1993), a final F0 rising movement and a sequence of lowered pitches preceding this sentence-final rise characterize yes-no questions containing more than two stress groups against their declarative counterparts in French. Vion and Colas (2006) applied the gating method to examine the role of these prosodic cues in the recognition of French yes-no questions. The reaction time was also measured in this experiment. Their results indicated that the lowered pitches preceding the sentence final rise contribute to the recognition of questions in 61 % of the cases. The accuracy percentage reaches the peak level (100 %) as soon as the participants hear the final gate which presents the final rise. They reported that the reaction time to declaratives is shorter than the reaction time to questions.

Dutch declarative questions are marked against declaratives by an upward trend of the declination line, the presence of a final rise, and a greater excursion size of the pitch accent associated with the object constituent of the sentence (Van Heuven and Haan 2000). Van Heuven and Haan (2000) designed a gating paradigm experiment to inspect the influence of these acoustic cues on the perception of declaratives versus declarative questions in Dutch. Their findings revealed that the prosodic cues before the final rise considerably contributes to declarative versus interrogative perception (90 %). The accuracy was raised to 100 % when participants were exposed to the final rise.

These studies converge in the conclusion that prosody cues correct perception of interrogatives as opposed to declaratives when the syntactic structure does not distinguish these two sentence types. An implication of

this is that the absence of the *wh*-phrase at the sentence-initial position may cause the acoustic cues of prosody to cue correct perception of *wh*-in-situ questions as opposed to declaratives in Persian.

2 Research questions, approach and hypothesis

It is repeatedly shown that prosody can cue identification of questions lacking lexico-syntactic features of interrogativity in different languages (c.f. Section 1.1.3). However, it is not known yet if prosody can contribute to the perception of questions in which lexico-syntactic interrogativity features appear later in the sentence, e.g. *wh*-in-situ questions. Therefore, this research tackles the following research question: does prosody of the pre-*wh* part of the sentence cue correct perception of Persian *wh*-in-situ questions as opposed to declaratives? To answer this question, a forced choice sentence identification task was designed in which Persian native speakers listened to the pre-*wh* part of declaratives and *wh*-in-situ questions. They were required to decide as quickly as possible what they heard is going to be a *wh*-question or a declarative. Based on Vion and Colas (2006) perception experiment, we expect shorter reaction times to declaratives than to *wh*-questions. Another reason for this expectation is a higher frequency of occurrence of declaratives than questions in daily conversation (Van Heuven and Haan 2000).

According to Ohala's frequency code, Bolinger's dichotomy theory of relaxation and tension, Snedeker and Trueswell (2003) proposal and based on earlier experimental results (Face 2005; Petrone and D'Imperio 2011; Petrone and Niebuhr 2014; Sensui 1995; Truckenbrodt et al. 2009; Van Heuven and Haan 2000; Vion and Colas 2006) we predict that listeners use their implicit knowledge that high pitch and fast speech rate represent an interrogative while low pitch and slow speech rate denote a declarative.⁵ Therefore, we hypothesize that Persian native speakers can discern *wh*-in-situ questions from declaratives drawing on the acoustic correlates of prosody of the pre-*wh* part of the sentence.

3 Methodology

3.1 Participants

Twenty-six native speakers of Persian, 13 males and 13 females, took part in the experiment. All participants were brought up in Tehran and came to the Netherlands in 2014 to continue their education at the Technology University of Delft. Their age range was between 26 and 40 years. None of them reported any hearing impairment.

3.2 Materials

3.2.1 Selection of speakers

The sentences produced in an earlier production experiment (Shiamizadeh et al. 2016) were used as the materials of the current experiment. To control for the potential effect of gender on the listeners' performance in the perception experiment we chose both a male and a female speaker.

Selecting those speakers who keep the two sentence types most distinct in their speech would limit the generalizability of the results to only these speakers. To make the results of the current experiment more generalizable we therefore selected speakers who are the best representatives of all participants of the production experiment. Speakers whose mean value of the acoustic measurements (cf. Section 1.1.1) were closest to the mean values in the production of all speakers were chosen.

3.2.2 Selection of stimuli

Twenty pairs of sentences elicited from a male and a female speaker in the production experiment by Shiamizadeh et al. (2016) comprise the stimuli of this experiment. The structure of the *wh*-questions and declaratives is illustrated in (3) and (4), respectively.

(3)
 Subj Adv Wh-phrase Verb

(4)
 Subj Adv ADO/ IDO/ AdjT/ AdjM/ AdjP⁶ Verb

Subject is abbreviated as Subj, adverb as Adv, animate direct object as ADO, inanimate direct object as IDO, adjunct of time as AdjT, adjunct of manner as AdjM and adjunct of place as AdjP. As (4) shows ADO, IDO, AdjT, AdjM and AdjP replace the wh-phrase in declaratives. Therefore, they will be referred to as declarative wh-phrase counterparts (DWC) in the remainder of the paper. An example of a declarative and a matching wh-question is given in (5a) and (5b).

(5)
 a.

<i>mohæmædʔæmin</i>	<i>pæriruz</i>	<i>ʔæsr</i>	<i>fena</i>	<i>kærd.</i>
Mohammadamin	two days ago	afternoon	swim	do. PAST. 3SG

‘Mohammadamin swam in the afternoon two days ago.’

b.

<i>mohæmædʔæmin</i>	<i>pæriruz</i>	<i>kej</i>	<i>fena</i>	<i>kærd?</i>
Mohammadamin	two days ago	when	swim	do. PAST. 3SG

‘When did Mohamadamin swim two days ago?’

Five different wh-phrases, two different nouns as the subjects, two words as the adverbs, two words in each category of DWC and five verbs⁷ were used as the sentence constituents of the original stimuli in the production experiment. The word constituents of the declaratives and wh-questions are presented in Table A. Repeated measures multivariate analysis of variance showed that the interaction effect between the nouns used as the subject and sentence type ($F(5,65) = 0.397, p > 0.05$; Wilk's $A = 0.970, \eta^2 = 0.030$) and between the words used as the adverb and sentence type ($F(6,12) = 0.432, p > 0.05$; Wilk's $A = 0.968, \eta^2 = 0.032$) on the dependent variables (acoustic features described in Section 2) was not significant. Therefore, we decided to include just one noun as the subject and one word as the adverb in the stimuli of this experiment. Variation in other sentence constituents is kept intact.

The total number of stimuli in this experiment equals 40 (1 subject x 1 adverb x 2 DWCs x 5 wh-phrases or matching verb x 2 conditions x 2 speakers).⁸ The pre-wh parts of the sentences were separated from the remainder of the sentence in Praat (Boersma and Weenink 2014) and were used as the stimuli of the current experiment. Table B presents descriptive statistics of the prosodic correlates (cf. Section 1.1.1) of the stimulus material used in this experiment.

3.3 Procedure

A forced-choice categorization task was designed in E-prime 2.0.10 software (Psychology Software Tools 2012). Participants were seated in front of a computer in a quiet room. The experiment started with the emergence of a written instruction on the computer screen. Participants could take as much time as they wanted to read the instruction. They were allowed to ask questions about the instruction if they had any. Then, they were introduced to the task by performing a practice session which included eight non-experimental items. The presentation order of the items of the practice session was the same for all participants. In the practice session, the pre-wh part of two declaratives and two wh-questions read by two different speakers were played to them over headphones (Sennheiser PC 141 headset). Participants were instructed to decide as soon as possible whether what they heard is going to be a wh-question or a declarative by pressing either M or Z on the keyboard within three seconds after hearing each stimulus. A full stop and a question mark along with the letters M and Z appeared on two opposite sides (left and right) of the screen at the same time a stimulus was played to them. Upon pressing the key, the next stimulus was played after 200 ms, during which a “+” sign was shown on the screen as a warning that the next stimulus is coming. If participants did not respond within three seconds, the experiment proceeded to the next stimulus automatically. Participants were allowed to do the practice session two times if they wanted. Having accomplished the practice session, participants embarked on the main part

of the experiment when they felt ready. To make participants ready for the main part of the experiment after the practice session, the main session started with a warming-up which consisted of 4 non-experimental items. The presentation order of the stimuli was randomized per participant. The order in which the full stop and the question mark were displayed on the screen was fixed for individual participants, whereas it was counterbalanced across participants. The procedure of the main session was identical to that of the practice session. The experiment lasted about 10 minutes for each participant.

3.4 Data analysis

The responses and the RT data were transferred from E-prime to SPSS version 22 (IBM SPSS, 2012). The response accuracy to declaratives and wh-questions were computed in terms of percentage correct and A' scores (Stanislaw and Todorov 1999). RT was calculated as the time between the response and the stimulus offset. Two separate Repeated Measures ANOVA (RM-ANOVA) were run on RT and accuracy data to investigate the effect of sentence type, speaker variation, wh-phrase type, and their interaction on response accuracy and RT.

4 Results

4.1 Accuracy

Table 1 presents the accuracy of sentence type perception per condition. As Table 1 illustrates, more correct responses were given to declaratives than to wh-questions. A RM-ANOVA was run with aggregated responses as the dependent variable and speaker variation, wh-phrase type and sentence type as independent variables. The result showed that there is no main effect of sentence type ($F(1,25) = 1.402$, Wilks' Lambda = 0.947, $p > 0.1$; $\eta^2 = 0.053$), speaker ($F(1,25) = 4.089$, Wilks' Lambda = 0.859, $p > 0.05$; $\eta^2 = 0.141$) and wh-phrase type ($F(4,22) = 1.229$, Wilks' Lambda = 0.817, $p > 0.1$; $\eta^2 = 0.183$) on response accuracy. The interaction between speakers and sentence type ($F(1,25) = 1.183$, Wilks' Lambda = 0.955, $p > 0.1$; $\eta^2 = 0.045$) and between wh-phrase type and sentence type ($F(4,22) = 0.541$, Wilks' Lambda = 0.910, $p > 0.1$; $\eta^2 = 0.090$) on response accuracy is not significant, either.

Table 1. Perception of intended sentence type per condition.

	Correct %	N	Incorrect %	N	Missing cases %	N
Declaratives	91.7	477	7.9	41	0.4	2
Wh-in-situ questions	88.7	461	11.2	58	0.2	1

In addition to calculating accuracy percentages, A' score was calculated to correct for a possible response bias (Stanislaw and Todorov 1999).⁹ A' values range from 0.5 to 1; 0.5 shows inability of the listeners to discriminate the two sentence types and 1 presents perfect performance. The average A' scores for all participants in both conditions was 0.937. This A' score reveals that the participants identified the correct sentence type with a high level of accuracy.

4.2 Reaction time

Reaction time was calculated as the time lapse between the stimulus offset and the response (all RT data are reported in seconds). When the response was given before the stimulus offset,¹⁰ the reaction time value was negative.

The mean RT to declaratives was shorter than the mean RT to questions (see Table 2). A RM-ANOVA revealed a main effect of sentence type ($F(1,25) = 14.740$, Wilks' Lambda = 0.629, $p < 0.01$; $\eta^2 = 0.371$). In contrast, the effect of the speaker ($F(1,25) = 0.036$, Wilks' Lambda = 0.999, $p > 0.1$; $\eta^2 = 0.001$), the wh-phrase type ($F(4,22) = 1.456$, Wilks' Lambda = 0.791, $p > 0.1$; $\eta^2 = 0.209$), the interaction between speaker and sentence type

($F(1,25) = 0.860$, Wilks' Lambda = 0.967, $p > 0.1$; $\eta^2 = 0.033$) and the interaction between sentence type and wh-phrase type ($F(4,22) = 1.546$, Wilks' Lambda = 0.781, $p > 0.1$; $\eta^2 = 0.219$) was not significant.

Table 2. Mean reaction time (and standard deviation) (in s) for declaratives and wh-questions.

	Mean
Declaratives	0.282 (0.312)
Wh-questions	0.430 (0.330)

5 Discussion and conclusion

The current study was designed to address the question whether 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 sentence-initial position. As it turns out, wh-questions can be correctly distinguished from declaratives in 90.30 % of the cases. The A' score was 0.937. Accordingly, it can be inferred that prosody plays an influential role in correct sentence type perception when the interrogativity marker occurs later in the sentence.

This result is consistent with the prediction that prosodic characteristics of the pre-wh part of the sentence cue identification of wh-in-situ questions as opposed to declaratives in Persian. The finding also confirms the result of previous perception experiments on the important role of prosody in the identification of interrogatives as opposed to declaratives in the absence of syntactic interrogativity markers (Face 2005; Petrone and D'Imperio 2011; Petrone and Niebuhr 2014; Sensui 1995; Truckenbrodt et al. 2009; Van Heuven and Haan 2000; Vion and Colas 2006).

In line with the results of Vion and Colas (2006), analysis of our data demonstrated that the reaction time to declaratives is shorter than the reaction time to questions. This can be explained by the higher frequency of occurrence of declaratives in comparison to questions in daily conversation (Van Heuven and Haan 2000). Speaker variation and wh-phrase type influences neither response accuracy nor RT.

Ohala's (1984) frequency code, Bolinger's (1989) dichotomy theory of relaxation and tension and Snedeker and Trueswell's (2003) proposal can be possible theoretical explanations for the ability of the listeners to perceive the correct sentence type. It can be argued that listeners identify the sentence type drawing on their implicit knowledge that high pitch and fast speech rate signal interrogativity whereas low pitch and slow speech rate denote declarativity. Based on Rialland (2004) work, the association of high pitch and fast rate with questions is not universal. This implies that Persian native speakers draw on language-specific implicit knowledge on the prosodic differences between Persian statements and wh-in-situ questions to identify the sentence type.

That listeners need not wait till the syntactic cue is uttered to predict the sentence type is plausible from processing and communicative points of view. Prediction can facilitate processing and decoding of sentence type and response preparation on the part of the listener (Grosjean 1983, Grosjean 1996). From the view of efficient communication, it is logical that sentence mode is designated as early as possible in an utterance (Bolinger 1981; Petrone and Niebuhr 2014). Sentence type identification in the absence of syntactic or lexical cue and based on prosody serves this communicative need.

The current inquiry yields evidence for the influential role of prosody in the identification of wh-in-situ questions in Persian. However, the relative contribution of duration and the various F0 features to sentence type perception is not evident yet. In a subsequent study, we will manipulate the F0 contour and duration and present the manipulated and resynthesized originals to Persian listeners in a sentence identification task, to inspect the relative contribution of the F0 and duration to the identification of declaratives vs. wh-in-situ questions.

Notes

¹Polarity questions in 18 languages (in a database of 80 African languages) are not characterized by high tones, rising intonation and shorter duration (Rialland 2004). This suggests that the association between high pitch and fast rate with questions is language-specific rather than universal.

²Plain yes-no questions are the questions that lack the question particle "mipos".

³The main pitch accent in BP falls on the last word of the sentence by default (Gebara, 1976).

⁴As a preliminary step, we need to know whether the prosodic characteristics we found in the production experiment (Shiamizadeh et al. 2016) are perceptible by Persian native speakers. Therefore, we conducted this simple perception study before running a gating experiment.

⁵We reiterate that we do not claim the universal association of high pitch and fast rate with questions. Since the production study on the prosody of Persian wh-in-situ questions (Shiamizadeh et al. 2016) indicated that the frequency code hypothesis (Ohala 1984) and linking high pitch and fast rate with interrogativity is the preferred association in Persian, we predict that Persian native speakers can draw on this language-specific implicit knowledge to discern wh-in-situ questions from declaratives.

⁶Different categories of DWC in example (3) are separated by slashes to illustrate that each declarative includes only one of the categories of DWC.

⁷As Appendix 1 shows, the verb varies along the wh-phrase.

⁸Though only the pre-wh part of the sentences comprises the stimuli of the current experiment, variation in DWCs, wh-phrases and their matching verbs are included in the formula to clarify how we arrived at 40 stimuli.

⁹A' score is not indicative of the response bias. Rather it shows performance of the participants excluding the effect of possible response bias.

¹⁰12.9% (f = 134) of declaratives and 6.1% (f = 63) of wh-questions were responded to before the stimulus offset.

References

- Abedi, F., A. Moizadeh & Z. Gharaei. 2012. WH-movement in English and Persian within the framework of government and binding theory. *International Journal of Linguistics* 4. 419–432.
- Adli, A. 2007. Constraint cumulativity and gradience: Wh-scrambling in Persian. *Lingua* 120. 2256–2294.
- Baltazani, M. 2007. Intonation of polar questions and the location of nuclear stress in Greek. In Gussenhoven C. & T. Riad (eds.), *Tones and tunes: Experimental studies in word and sentence prosody*, 387–405. Berlin: Mouton de Gruyter.
- Baltazani, M., E. Kainada, A. Lengeris & K. Nicolaidis. 2015. The prenuclear field matters: Questions and statements in standard modern Greek. In The Scottish consortium for ICPHS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, 10–14. Lund: Lund University.
- Beach, C. M., W. F. Katz & A. Skowronski. 1996. Children's processing of prosodic cues for phrasal interpretation. *Journal of the Acoustical Society of America* 99(2). 1148–1160.
- Boersma, P. & D. Weenink. 2014. Praat: Doing phonetics by computer (Version 6.0.04) [Computer program]. <http://www.praat.org/>.
- Bolinger, D. W. 1989. *Intonation and its uses*. Palo Alto: Stanford University Press.
- Bolinger, D. 1981. The place of intonation in a discourse model. In M. C. Loulth & M. Montgomery (eds.), *Studies in discourse analysis*, 146–157. London: Routledge & Kegan Paul.
- Carlson, K., C. Clifton & L. Frazier. 2001. Prosodic boundaries in adjunct attachment. *Journal of Memory and Language* 45. 58–81.
- Carnie, A. 2007. *Syntax: A generative introduction*, 2nd edn. Oxford: Blackwell.
- Chomsky, N. 1977. On wh-movement. In P. W. Culicover, T. Wasow & A. Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.
- D'Imperio, M. 2000. *The role of perception in defining tonal targets and their alignment*. Doctoral dissertation, Ohio State University https://etd.ohiolink.edu/etd.send_file?accession=osu1243021045&disposition=inline.
- De Moraes, J. A. 1998. Intonation in Brazilian Portuguese. In D. Hirst & A. Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 179–194. Cambridge: Cambridge University Press.
- Di Cristo, A. & D. J. Hirst. 1993. Prosodic regularities in the surface structure of French questions. In D. House & P. Touati (eds.), *Proceedings of the European Speech Communication Association Workshop on Prosody*, 268–271. Lund: Lund University.
- Face, T. L. 2004. The intonation of absolute interrogatives in Castilian Spanish. *Southwest Journal of Linguistics* 23(2). 65–79.
- Face, T. L. 2005. Fo peak height and the perception of sentence type in Castilian Spanish. *Revista Internacional de Linguística Iberoamericana* 2. 49–65.
- Gorjian, B., M. Naqhzadeh & P. Shahramiri. 2012. Making interrogative sentences in English and Persian language: A contrastive analysis approach. *Journal of Comparative Linguistics and Literature* 2. 120–124.
- Grosjean, F. 1983. How long is the sentence? Prediction and prosody in the on-line processing of language. *Linguistics* 21. 501–529.
- Grosjean, F. 1996. Using prosody to predict the end of sentences in English and French: Normal and brain damaged subjects. *Language and Cognitive Processes* 11. 1–2.
- Haan, J. 2003. *Speaking of questions: An exploration of Dutch question intonation*. Doctoral dissertation, Leiden University. www.lotpublications.nl/Documents/52_fulltext.pdf.
- IBM SPSS. 2012. IBM SPSS. IBM Software Business Analytics.
- Kahnemuyipour, A. 2009. *The syntax of sentential stress*. Oxford: Oxford University Press.
- Karimi, S. 2005. *A minimalist approach to scrambling: Evidence from Persian*. Berlin: Mouton de Gruyter.
- Karimi, S. & A. Taleghani. 2007. Wh-movement, interpretation, and optionality in Persian. In S. Karimi, V. Samiiian & W. K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation*, 167–187. Amsterdam: Benjamins.
- Lieberman, M. & J. Pierrehumbert. 1984. Intonational in-variance under changes in pitch range and length. In M. Aronoff & R. T. Oehle (eds.), *Language sound structure*, 157–233. Cambridge: MIT Press.
- Lotfi, A. R. 2003. Persian Wh-riddles. In C. Boeckx & K. K. Grohmann (eds.), *Multiple wh-Fronting*, 161–186. Amsterdam: Benjamins.
- Megerdoomian, K. & S. Ganjavi. 2000. Against optional wh-Movement. In V. Samiiian (eds.), *Proceedings of the Western Conference on Linguistics: WECOL*, 358–370. Fresno: California University.
- Mirsaeedi, A. 2006. *Wh-movement in Persian language* Isfahan University of Isfahan MA thesis.
- Ohala, J. 1984. An ethological perspective on common cross-language utilization of Fo of voice. *Phonetica* 41. 1–16.
- Petrone, C. 2008. *Le rôle de la variabilité phonétique dans la représentation des contours intonatifs et de leur sens*. Doctoral dissertation, Université de Provence <https://tipa.revues.org/410>.

- Petrone, C. & M. D'Imperio. 2008. edited by P. A. Basbosa, S. Madureira & C. Reis *Proceedings of the Fourth Conference on Speech Prosody*, São PauloCapes Tonal structure and constituency in Neapolitan Italian: Evidence for the accentual phrase in statements and questions. 301–304.
- Petrone, C. & M. D'Imperio. 2011. From tones to tunes: Effects of the fo prenuclear region in the perception of Neapolitan statements and questions. In S. Frota, G. Elordieta & P. Prieto (eds.), *Prosodic categories: Production, perception and comprehension*, 207–230. Dordrecht: Springer.
- Petrone, C. & O. Niebuhr. 2014. On the intonation of German intonation questions: The role of the prenuclear region. *Language and Speech* 57(1). 108–146.
- Psychology Software Tools. 2012. E-Prime. Pittsburg, PA.
- Rialland, A. (2004). Paper presented at the International Conference on Tone and Intonation, Greece, Santorini *A typology of question prosody in African languages*.
- Rietveld, A.C.M. & C. Gussenhoven. 1987. Perceived speech rate and intonation. *Journal of Phonetics* 15. 273–285.
- Rietveld, T. & P. Vermillion. 2003. [Cues for perceived pitch register](#). *Phonetica* 60. 261–272.
- Sensui, H. 1995. Percepción de la entonación interrogativa del español: Un estudio experimental. *Sophia Lingüística* 38. 1–23.
- Shiamizadeh, Z., J. Caspers & N. O. Schiller. 2016. *Do Persian native speakers prosodically mark wh-in-situ questions?*. Manuscript submitted for publication.
- Snedeker, J. & J. Trueswell. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 48(1). 103–130.
- Stanislaw, H. & N. Todorov. 1999. Calculation of signal detection theory measures. *Behavior Research Methods, Instruments and Computers* 31. 137–149.
- Toosarvandani, M. 2008. Wh-movement and syntax of sluicing. *Journal of Linguistics* 44. 677–722.
- Truckenbrodt, H., F. Sandalo & B. Abaurre. 2009. Elements of Brazilian Portuguese intonation. *Journal of Portuguese Linguistics* 8(1). 75–114.
- Van Heuven, V. & J. Haan. 2000. Phonetic correlates of statement versus questions intonation in Dutch. In A. Botinis (eds.), *Intonation, analysis, modeling and technology*, 119–143. Dordrecht: Kluwer.
- Van Heuven, V. & E. Van Zanten. 2005. Speech rate as a secondary prosodic characteristic of polarity questions in three languages. *Speech Communication* 47. 87–99.
- Vion, M. & A. Colas. 2006. Pitch cues for the recognition of yes-no questions in French. *Journal of Psycholinguistics Research* 35. 427–445.

Appendix

Table A: Structure and word constituents of the sentences from which the stimuli of this experiment is extracted. Only the pre-wh part (subject and the adverb) comprise the stimuli of this experiment.

Subject	Adverb	3rd constituent of the sentence Wh-phrase	DWC ^a	Verb
1) mohæmædʔæmin (Mohamadamin)	1) pæriruz (two days ago)	1) tʃi (what)	Inanimate DO ^b 1) kæfʃ (shoe) 2) kif (bag)	tæʔmir kærdæn (repair)
2) mohæmædʔæli (Mohamadali)		2) ki (who)	Animate DO 1) jas (Yas) 2) naz (Naz)	dæʔvæt kærdæn (invite)
		3) kodʒa (where)	Adjunct of Place ^c 1) dʒæŋgæl (jungle) 2) kutʃe (street)	bazi kærdæn (play)
		4) kej (when)	Adjunct of Time 1) zohr (noon) 2) ʔæsɾ (afternoon)	ʃena kærdæn (swim)
		5) tʃetori (how)	Adjunct of Manner 1) ba ʔænduh (sadly) 2) ba deqæt (carefully)	sohbæt kærdæn (talk)

^aDWC refers (declarative wh-phrase counterpart) to the categories which replace the wh-phrase in declaratives.

^bDO refers to direct object. The object marker “ra” occurs after direct object in declaratives and wh-in-situ questions.

^cThe preposition “tu” which means “at” precedes the adjunct of place in declaratives but not in wh-in-situ questions.

Table B: Descriptive statistics of the prosodic correlates of the stimuli across sentence types.

Variable	M SD (Decl)	M SD (Wh-q)	N
----------	----------------	----------------	---

F0 onset (ERB)	5.082 1.303	5.928 0.774	40
Subject Pitch Excursion (ERB)	1.025 0.420	1.740 0.354	40
Adverb Pitch Excursion (ERB)	0.622 0.109	1.430 0.222	40
Slope of regression line of the pre-wh part (ERB/sec)	0.143 0.102	0.954 0.211	40
F0 mean of the pre-wh part (ERB)	4.903 1.120	5.666 0.951	40
Duration of the pre-wh part (in sec)	1.360 0.072	1.043 0.069	40