

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

Zohreh Shiamizadeh, Johanneke Caspers & Niels O. Schiller

To cite this article: Zohreh Shiamizadeh, Johanneke Caspers & Niels O. Schiller (2018) When is a wh-in-situ question identified in standard Persian?, *Language, Cognition and Neuroscience*, 33:9, 1168-1183, DOI: [10.1080/23273798.2018.1463444](https://doi.org/10.1080/23273798.2018.1463444)

To link to this article: <https://doi.org/10.1080/23273798.2018.1463444>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 18 Apr 2018.



[Submit your article to this journal](#)



Article views: 216



[View Crossmark data](#)

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

Zohreh Shiamizadeh^{a,b}, Johanneke Caspers^{a,b} and Niels O. Schiller ^{a,b}

^aLeiden University Centre for Linguistics, Leiden University, Leiden, the Netherlands; ^bLeiden Institute for Brain and Cognition, LUMC, Leiden University, Leiden, the Netherlands

ABSTRACT

Previous literature demonstrated the influential role of prediction in processing speech [Brazil, 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; Grosjean, 1983. How long is the sentence? Prediction and prosody in the on-line processing of language. *Linguistics*, 21, 501–529, 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; Snedeker & Trueswell, 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, 48, 103–130], and of prosody in predicting the eventual syntactic structure of ambiguous sentences [e.g. Snedeker & Trueswell, 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, 48, 103–130]. Wh-in-situ questions contain temporary syntactic ambiguity. One of the languages characterised by wh-in-situ questions is Persian. The current research adopted the gating paradigm [Grosjean, 1980. Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics*, 28, 267–283] to investigate when distinctive prosodic cues of the pre-wh part enable correct identification of wh-in-situ questions in Persian. A perception experiment was designed in which gated stimuli were played to Persian native speakers in a forced-choice sentence identification task. In line with our expectation, correct identification responses were given from the beginning of the sentence. The result is discussed in the context of proposals regarding the need to integrate prosody and prediction into models of language and speech processing [Beach, 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; Grosjean, 1983. How long is the sentence? Prediction and prosody in the on-line processing of language. *Linguistics*, 21, 501–529, 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].

ARTICLE HISTORY

Received 2 August 2017
Accepted 30 March 2018

KEYWORDS



Persian; wh-in-situ questions; prosody; gating

1. Introduction

Processing conversational speech is part of language processing. According to Grosjean (1983, 1996a), listeners draw on any source of information that can facilitate and accelerate the processing of a conversation. They use past and present information to process sentences up to the point uttered by the speaker and to predict forthcoming information. Prediction can be helpful to the listener in several ways, for example it can focus listeners' attention by reducing the set of possibilities, or it can give listeners time for other activities that can accelerate processing and communication, such as integrating information, storing it and preparing a response. 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 (Grosjean,

1983, 1996a). One source of information in speech processing and prediction of upcoming events is prosody. According to Grosjean (1983, 1996a), 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.

Previous studies on the role of prosody in sentence comprehension (e.g. Snedeker & Trueswell, 2003) indicate that speakers and listeners not only share some implicit knowledge about the correspondence between prosody and syntax, but also can utilise this knowledge to guide their interpretation of syntactically ambiguous sentences. Efficient use of prosody in processing syntactically ambiguous sentences has been demonstrated by multiple researchers (e.g. Beach, 1991; Beach, Katz, & Skowronski,

CONTACT Zohreh Shiamizadeh  zshiamizadeh@gmail.com, z.shiamizadeh@hum.leidenuniv.nl
 Supplemental data for this article can be accessed at <https://doi.org/10.1080/23273798.2018.1463444>

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

1996; Carlson, Clifton, & Frazier, 2001; Kjølgaard & Speer, 1999; Nagel, Shapiro, & Nawy, 1994; Snedeker & Trueswell, 2003; Warren, Grabe, & Nolan, 1995). These studies have revealed that listeners can efficiently use prosody to predict the eventual syntactic structure of sentences that have local or global syntactic ambiguity. In situations of global syntactic ambiguity, the sentence remains syntactically ambiguous even after all lexical information of the sentence has been presented, as in the sentence “I saw a man in the garden with a binocular”. The syntactic ambiguity is local if the information in the early parts of the sentence is compatible with several possible structures of the current input, but the information in a later portion of the sentence assigns only one possible grammatical interpretation to the sentence (Beach, 1991).

The current study focuses on a type of interrogative which typically has local syntactic ambiguity, i.e. *wh-in-situ* questions. *Wh*-questions can be divided into two categories: (i) fronted *wh*-questions, and (ii) *in-situ wh*-questions. Fronted *wh*-questions are constructed by the obligatory movement of the *wh*-phrase. This obligatory movement is a syntactic process which results in the movement of the *wh*-phrase to sentence-initial position (Carnie, 2007; Chomsky, 1977; see example 1).

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

English is an example of a fronted *wh*-question language. However, in some other languages the *wh*-phrase does not undergo movement, but it remains in the same site where the non-*wh* phrase is expected to occur in its declarative counterpart. This phenomenon is known as *wh-in-situ*: the *wh*-phrase does not move but stays in the original position where its declarative counterpart is expected to appear. One language characterised by *in-situ wh*-questions is Persian (Abedi, Moinzadeh, & Gharaei, 2012; Adli, 2007; Gorjian, Naghizadeh, & Shahramiri, 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdooomian & Ganjavi, 2000; Mirsaeeadi, 2006; Toosarvandani, 2008). In Persian, *wh*-questions are *in-situ* by default (see 2b).¹

- (2) a *mæɾjæm* *ʔæsr* *bazi- kærd.*
 Maryam afternoon play- do.PAST.3SG.
 “Maryam played in the afternoon.”
 b *mæɾjæm* *kej* *bazi- kærd ?*
 Maryam when play- do.PAST.3SG.
 “When did Maryam play?”

As (2b) illustrates, the syntactic feature relating to the clause type in *wh-in-situ* questions, namely the *wh*-phrase, occurs later in the sentence. Therefore, *wh-in-situ* questions typically have local syntactic ambiguity.

Engaging in a conversation requires the smooth exchange of information. Asking a question is

tantamount to eliciting a verbal response from the addressee and people rarely leave long gaps between turns (Brazil, 1981; Sacks, 2004; Sacks, Schegloff, & Jefferson, 1974; Schegloff, 2006; Stivers et al., 2009). Combining the proposal of minimising gaps between turns (Brazil, 1981; Sacks, 2004; Sacks et al., 1974; Schegloff, 2006; Stivers et al., 2009) and the purpose of asking a question, we can suggest that listeners need to be made aware of the purpose of the speaker to have enough time to process the sentence and prepare a response. Early awareness of the purpose of the speaker facilitates and accelerates sentence processing and response preparation. In other words, the earlier the listeners can predict the syntactic structure of the sentence the more time they will have to prepare a response. The results of the perception study by Shiamizadeh, Caspers, and Schiller (2017a) suggest that the prosody of the pre-*wh* part of a sentence can help predict sentence type (*wh-in-situ* questions vs. declaratives) in Persian in the absence of the *wh*-phrase at the sentence-initial position (see Section 1.1.2). The result of that perception study 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?

1.1. Background

1.1.1. Definition of the acoustic correlates

This section presents the definition of the acoustic features reported in the current study. F0 onset is the first voiced frame of a segment (phoneme, word, sentence) (Haan, 2001). Declination (vs. inclination) is the gradual time-dependent lowering of F0 in the course of an utterance or a text (e.g. Cohen & ‘t Hart, 1967; Liberman & Pierrehumbert, 1984). A pitch accent is “a local feature of a pitch contour – usually but not invariably a pitch range, and often involving a local maximum and minimum – which signals that the syllable with which it is associated is prominent in the utterance” (Ladd, 2008, p. 42). H* (high) and L* (low) tones or tone complexes like H*L or LH* are examples of pitch accents in which the starred tone associates with an accented syllable (Gussenhoven, 2004). Pitch excursion can be defined as “the size of a local pitch movement” (Ladd, 2008, p. 69). Boundary tones are tones that appear at the edge of prosodic constituents like the intonational phrase (Gussenhoven, 2004). L% (low) and H% (high) are examples of boundary tones.

1.1.2. Production and perception of Persian *wh-in-situ* questions

In a related study, Shiamizadeh, Caspers, and Schiller (2018) conducted a production experiment in which

they compared the prosodic correlates of Persian *wh*-in-situ questions with their declarative counterparts. They investigated whether acoustic correlates of the pre-*wh* part mark *wh*-in-situ questions as opposed to declaratives in the absence of the *wh*-phrase at the beginning of *wh*-questions. In their production experiment, Shiamizadeh et al. (2018) elicited declarative and *wh*-in-situ question stimuli from native speakers of Persian. They find that a higher level of pitch mean, a higher F0 onset and a shorter duration of the pre-*wh* part contribute to the prosodic distinction of the pre-*wh* part in *wh*-questions as opposed to declaratives. Steeper inclination of the pitch contour and a greater pitch excursion of the pitch accents realised on the pre-*wh* words are the two additional features that give rise to the prosodic markedness of the pre-*wh* part in *wh*-questions. The pitch accents in Persian are described as H* or L+H* in which the starred tone is associated with the stressed syllable (Mahjani, 2003; Sadat Tehrani, 2008). Figures 1 and 2 demonstrate the prosodic differences explained in this section. For the details of the measurement method of

the prosodic features the readers are referred to Shiamizadeh et al. (2018).

Following the production study, Shiamizadeh et al. (2017a) ran a perception experiment to investigate whether 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. A sentence identification task was designed in which the pre-*wh* part of *wh*-in-situ questions and their matching declaratives were presented to the participants at once (one-time stimulus presentation). As it turned out, *wh*-questions can be correctly distinguished from declaratives in 90.3% of the cases.

1.1.3. Empirical background

Some perception studies adopt the gating paradigm (Grosjean, 1996a) to try to determine the amount of acoustic-phonetic information required to identify a stimulus, for example a sentence type. The gating technique enables us to limit the amount of information

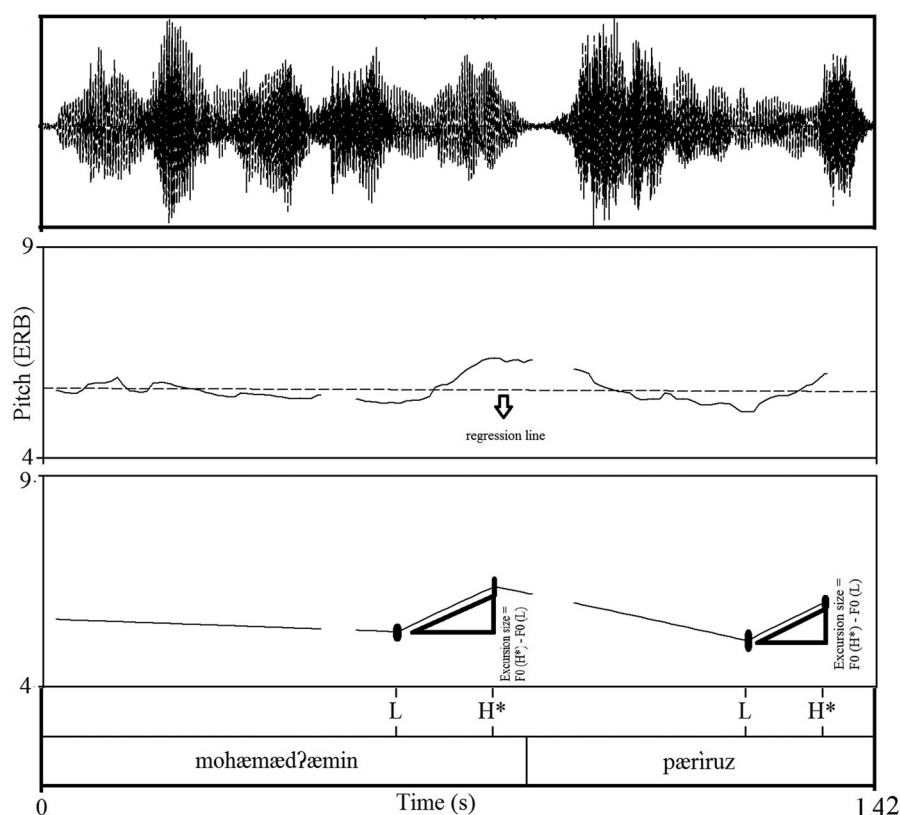


Figure 1. The acoustic correlates measured in the pre-*wh* part of a declarative sentence. In the second panel, the solid line is the pitch contour and the dotted line is the regression line. “L” and “H*” represent the valleys and the peaks of the realised pitch accents. The second tier represents the word boundaries. In the pitch stylised panel, only the points designating L and H* are kept and the irrelevant points are deleted. The vertical side of the triangle shows the excursion size of the pitch accents which is computed by subtracting the F0 value of H* (the peak of the accent) from the F0 value of L (the valley of the accent). The non-stylised pitch contour is presented along with the regression line.

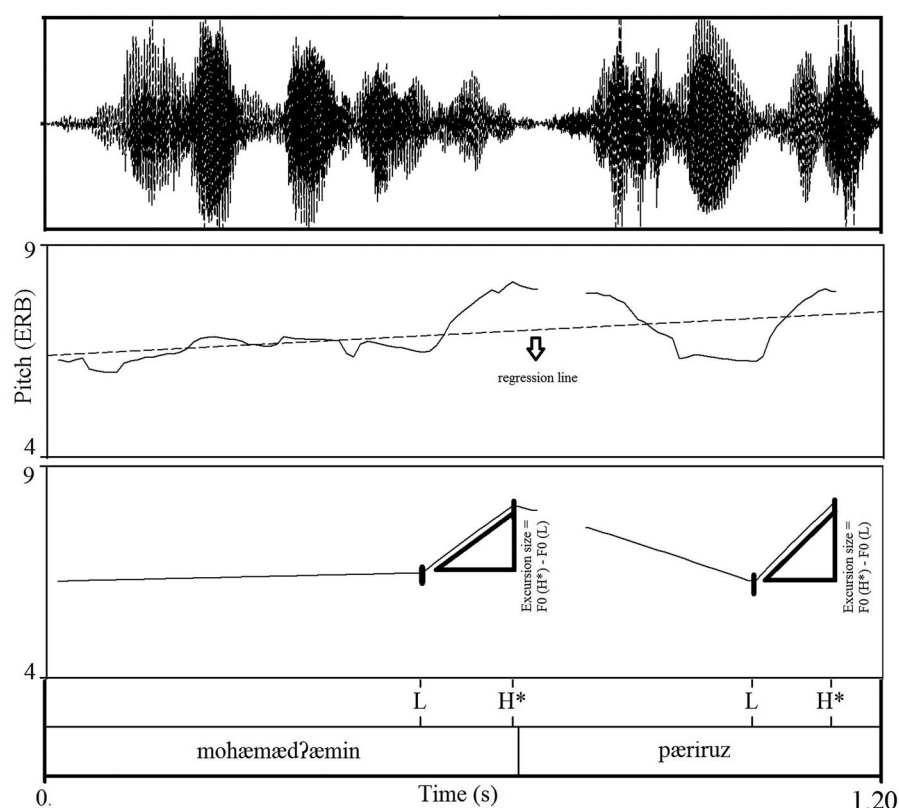


Figure 2. The acoustic correlates measured in the pre-wh part of a question. In the second panel, the solid line is the pitch contour and the dotted line is the regression line. “L” and “H*” represent the valleys and the peaks of the realised pitch accents. The second tier represents the word boundaries. In the pitch stylised panel, only the points designating L and H* are kept and the irrelevant points are deleted. The vertical side of the triangle shows the excursion size of the pitch accents which is computed by subtracting the F0 value of H* (the peak of the accent) from the F0 value of L (the valley of the accent). The non-stylised pitch contour is presented along with the regression line.

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 competitors² and possibly lead to the correct prediction of the target (Beach, 1991). The gating technique also helps us to assess whether prediction improves as the listener progresses through the acoustic signal (Grosjean, 1983, 1996b).

As far as we know, no gating study has been conducted on the role of prosody in identifying Persian interrogative sentences, including wh-in-situ questions. However, there are gating studies that investigate whether and how prosody guides the identification of interrogatives as opposed to declaratives in other languages, namely Castilian Spanish, Neapolitan Italian, Northern Standard German, Dutch, French and Mandarin Chinese (Face, 2005; Gryllia, Yang, Pablos, Doetjes, & Cheng, 2016 September; Petrone & D’Imperio, 2011; Petrone & Niebuhr, 2014; Van Heuven & Haan, 2000; Vion & Colas, 2006; Yang, 2018). The studies by Gryllia et al. (2016 September) and Yang (2018) are on wh-in-situ questions and the other studies focused on yes-no

questions or declarative questions.³ In this section, we will briefly review the results of these studies.

Castilian Spanish yes-no questions do not syntactically differ from declaratives, but they have recognisable prosodic characteristics, namely a raised F0 peak (the high tone in the pitch accent) height in pitch accents and a final F0 rising movement (final F0 rise) (Face, 2004). Another prosodic feature that disambiguates yes-no questions from declaratives in Castilian Spanish is the presence of pitch accents; in questions, only the first and the last word are associated with pitch accents, while in declaratives every stressed word is associated with a pitch accent. Face (2005) designed a gating paradigm study to investigate whether the acoustic cues of prosody enable listeners to perceive the correct sentence type. The results of his experiment showed that native speakers can correctly distinguish declaratives from yes-no questions in 95% of cases where the first prosodic distinction (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). The nuclear pitch accent (NPA) is the last pitch accent in a sentence. According to Petrone and D'Imperio (2008), NPA is aligned later in questions than in statements, in the form $L + H^*$ in questions but $L^* + H$ in statements. The F0 fall after the peak of the pitch accent preceding the NPA is shallower in questions, whereas the F0 falls rapidly from the peak of the pre-nuclear pitch accent to the end of the accented prosodic word in statements. The boundary tone of both sentence types is low; L-L%. In a perception study based on the gating paradigm, Petrone and D'Imperio (2011) investigate the contribution of the pre-nuclear region to sentence type categorisation in Neapolitan Italian. The results revealed that the prosody of the pre-nuclear region cues question identification (68%) and the accentual phrase boundary tone contributes significantly to question identification. Robust question recognition (above 90%) was achieved upon the presentation of the complete sentence.

German questions can be signalled lexically, syntactically and intonationally (Petrone & Niebuhr, 2014). According to Petrone and Niebuhr (2014), questions are not necessarily marked by a high boundary tone (H %) in Northern Standard German. Rather, they can have an L% similar to statements. However, similar to Neapolitan Italian, there are prosodic differences between the 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 pre-nuclear 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 the gating method, Petrone and Niebuhr (2014) found that F0 differences in the pre-nuclear pitch accent region significantly contribute to identification of questions as opposed to statements in Northern Standard German.

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

Van Heuven and Haan's (2000) study showed that 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. They designed a gating experiment to inspect the influence of acoustic cues in the perception of declaratives versus declarative questions in Dutch. Their findings revealed that the prosodic cues before the final rise considerably contribute to declarative versus interrogative perception (90%). The accuracy was raised to 100% when the participants were exposed to the final rise.⁴

Wh-phrases in Mandarin Chinese wh-questions appear in the same position as their non-interrogative counterpart in statements (Gryllia et al., 2016 September). According to Gryllia et al. (2016 September), F0, duration and intensity differentiate wh-in-situ questions from declaratives in Mandarin Chinese. They ran a gating experiment to investigate whether prosody cues identification of the clause type (declarative vs. wh-in-situ questions) before the appearance of the wh-phrase. They found that listeners could indeed identify the sentence type based on prosody from the first gate on, i.e. response accuracy to declaratives and questions was 59.6% and 64.6%, respectively. The authors suggested that listeners drew on F0 and duration to decide on the sentence type.

In a production study on Mandarin Chinese wh-in-situ questions, Yang (2018) reported that Mandarin Chinese wh-in-situ questions in which the wh-phrase is preceded by "dianr" can have an interrogative and a non-interrogative interpretation. The production experiment showed that prosodic features differentiate the declarative interpretation from the question interpretation: a) the pre-wh part in wh-questions has a shorter duration than declaratives, and b) the post-wh part in wh-questions has a higher pitch but a smaller F0 range in comparison to the post-wh part in declaratives. Following the production study, Yang (2018) conducted a gating experiment to investigate at what point prosody cues identification of sentence type. In this experiment, only the part of the sentence preceding the wh-phrase was presented. The results showed that listeners can identify the intended sentence type at the first gate, i.e. response accuracy is 59.0% for declaratives and 54.6% for questions. The response accuracy increases to 72.1% for declaratives and 62.1% for questions upon the presentation of the last gate (pre-wh part).

Yes-no questions in Castilian Spanish, Neapolitan Italian, Northern Standard German, and French, as well as declarative questions in Dutch are typical of sentences with global syntactic ambiguity and Mandarin Chinese wh-in-situ questions are sentences with local syntactic ambiguity. The results of the studies by Face (2005),

Gryllia et al. (2016 September), Petrone and D'Imperio (2011), Petrone and Niebuhr (2014), Van Heuven and Haan (2000), Vion and Colas (2006) and Yang (2018) all suggest that prosodic features available in the early parts of the sentence can cue the correct perception of interrogatives with global and local syntactic ambiguity. The efficient role of prosody in the identification of yes-no questions, declarative questions and even *wh*-in-situ questions in one or several languages, implies (however, not necessarily) that the prosodic correlates of the pre-*wh* part in Persian *wh*-in-situ questions could also cue prediction of *wh*-in-situ questions as opposed to declaratives. As such, a separate study is required to investigate the role of prosody in the perception of *wh*-in-situ questions vs. declaratives in Persian.

Previous gating studies mainly concentrated on yes-no questions and declarative questions as globally ambiguous sentences in languages other than Persian. As to the knowledge of the authors, gating studies on prosody-driven perception of *wh*-in-situ questions vs. declaratives are limited to Mandarin Chinese (Gryllia et al., 2016 September; Yang, 2018). This suggests that *wh*-in-situ question is an understudied question type with respect to the role that prosody plays 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 study on the perception of Persian *wh*-in-situ questions vs. declaratives. Therefore, investigating the role of prosody in the identification of an understudied question type in an understudied language is the added value of the current study.

2. Research questions, approach and hypotheses

2.1. Research questions and approach

In the perception study by Shiamizadeh et al. (2017a) the pre-*wh* part was presented to the participants based on one-time stimulus presentation (cf. Section 1.1.2). In contrast with the gating paradigm, one-time stimulus presentation cannot delineate how the preliminary decision whether a sentence is going to be a statement or a question develops over time and cannot determine the amount of acoustic-phonetic information needed to identify a stimulus (Grosjean, 1996a). This study applied the gating paradigm to answer the following research question: at what point can Persian native speakers use prosodic correlates to predict *wh*-in-situ questions before the *wh*-phrase is made audible?⁵ The answer to this question can improve current understanding of how prosody guides syntactic interpretation, in particular

temporary syntactic ambiguity resolution. It could also contribute to the evaluation of the proposal that integrates prediction into language processing models (Grosjean, 1983; Hale, 2001; Levy, 2008; Snedeker & Trueswell, 2003), and whether processing models need to account for the fact that a prediction can be reset as more prosodic information becomes available to the listener (Grosjean, 1983, 1996a).

To answer the research 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 statement or a *wh*-question. Participants were also asked to show how confident they were about their response by indicating on a confidence scale from one to five.

2.2. Hypotheses

From a descriptive point of view, prosodic correlates differentiate *wh*-in-situ questions from declaratives from the beginning of the sentence, since the F0 onset is higher in questions in comparison with declaratives (cf. Section 1.1.1). We hypothesise that Persian native speakers could start sentence type prediction from the beginning of the sentence, 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 (Snedeker & Trueswell, 2003). Such evidence includes the fact that high F0 onsets represent questions while low F0 onsets characterise statements. Along the same lines, we predict that identification improves as the amount of discriminating prosodic information increases. Thus, we expect higher rates of correct prediction upon the presentation of the pitch accents which are associated with the pre-*wh* words.

3. Methodology

3.1. Participants

Twenty native speakers of Persian, ten males and ten females, took part in this experiment. All participants were brought up in Tehran. They came to the Netherlands in 2014 or 2015 to continue their education at the Technology University of Delft. Their age range was between 26 and 40. None of them reported any hearing impairment. The data were collected in February 2016.

3.2. Materials

3.2.1. Speaker selection

Some of the sentences produced by native speakers of Persian who participated in the production experiment on the prosodic correlates of Persian *wh*-in-situ questions (Shiamizadeh et al., 2018) (see Section 1.1.1) were used as the materials for this experiment. To control for the effect of gender on the listeners' performance in the perception experiment, we chose both a male and a female speaker.

Selecting the 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 generalisable, we chose speakers who are the best representatives of 18 participants of the production experiment. Therefore, we selected a male and a female speaker whose mean value of the acoustic measurements were closest to the mean value of the acoustic measurements (cf. Section 1.1.1) in the production of all speakers.

3.2.2. The stimuli

3.2.2.1. Selection of the stimuli. Part of the sentences elicited in the production experiment by Shiamizadeh et al. (2018) (see Section 1.1.1) comprises the stimuli of the current perception experiment. The structure of the *wh*-question and declarative stimuli of the production experiment is illustrated in (3) and (4) respectively. Since the stimuli of the current perception experiment are chosen from the stimuli of the production experiment, (3) and (4) represent the structure of *wh*-question and declarative stimuli of the current experiment as well.

(3) Subj ⁶	Adv	Wh-phrase	Verb
(4) Subj	Adv	ADO/IDO/AdjT/AdjM/AdjP	Verb

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). Part of the sentence preceding the *wh*-phrase in *wh*-questions and the DWC in declaratives, i.e. the subject and the adverb, is referred to as the *pre-wh part*. The words in the *pre-wh part*, i.e. the subject and the adverb, will be referred to as *pre-wh words*. An example of a declarative and a matching *wh*-question is given in (5a) and (5b).

- (5) a. *mohæmædʔæmin pæriruz ʔæsr fæna- kærd.*
 Mohamadamin two days ago afternoon swim- do.PAST.3SG
 'Mohammadamin swam in the afternoon two days ago.'
- b. *mohæmædʔæmin pæriruz kej fæna- kærdʔ*
 Mohamadamin two days ago when swim- do.PAST.3SG
 'When did Mohammadamin 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 verbs were used as sentence constituents of the original stimuli in the production experiment by Shiamizadeh et al. (2018). The word constituents of the declaratives and *wh*-questions are presented in Appendix 1.

The subjects and the adverbs in *wh*-questions and declaratives were associated with a pitch accent, regardless of the *wh*-phrase and DWC (see Sections 1.1.1 & 4.1). Two separate repeated measures multivariate analysis of variance were run to investigate the effect of variation in the words used as the subject and the adverb on the difference between the acoustic features of declarative and *wh*-question stimuli elicited in the production experiment (c.f. Section 1.1.1). The result of repeated measures multivariate analysis of variances showed that the interaction effect between the nouns used as the subject and the sentence type ($F(5,65) = 0.397, p > .05$; Wilk's $A = .970, \eta_p^2 = .030$) and between the words used as the adverb and the sentence type ($F(6,12) = 0.432, p > .05$; Wilk's $A = .968, \eta_p^2 = .032$) on the acoustic features of declarative and *wh*-question stimuli elicited in the production experiment was not significant. Therefore, we decided to include just one noun as the subject and one word as the adverb in the stimuli of the current perception experiment. Variation in other sentence constituents is constant.

The *pre-wh part* of the sentences was separated from the remaining part of the sentence in Praat version 6.0.04 (Boersma & Weenink, 2014) and was used as the basic stimulus for the current experiment. The process of gating the stimuli is explained in Section 3.2.3. The complete version of each stimulus was played to the participants at the end of the experiment. The complete versions are syntactically unambiguous.

3.2.2.2. Number of the stimuli. Forty pairs of sentences elicited from a male and a female speaker (twenty pairs per speaker) in the production experiment by Shiamizadeh et al. (2018) comprise the stimuli of this experiment.

The total number of the stimuli of this experiment equals 320 (1 subject \times 1 adverb \times 2 DWCs \times 5 *wh*-phrases and the matching verbs \times 2 sentence types \times 2 speakers \times 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.

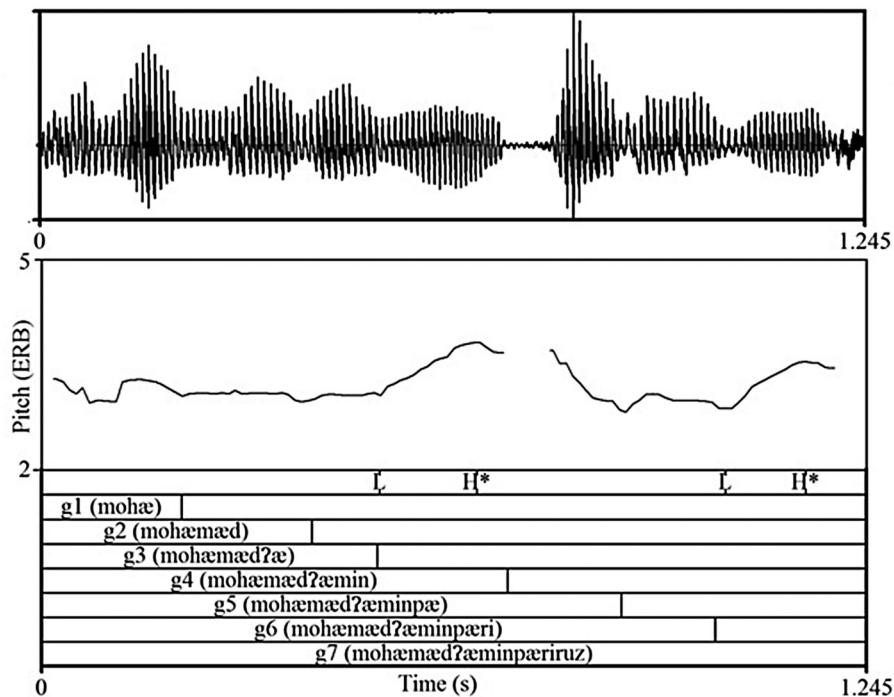


Figure 3. The seven gates of a declarative stimulus. The “L” and “H*” represent the valleys and the peaks of the realised pitch accents. The other tiers represent the gate boundaries. The letter g represents the word gate and the number designates the gate number.

3.2.3. Gating procedure

The pre-wh part of the sentence was 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 6 and Figures 3 and 4). 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. Example (6a) presents an example of a stimulus and (6b) illustrates the gates and boundaries. Figures 3 and 4 illustrate the pitch contour and the gates of both a declarative and a question stimulus. The term gate will be abbreviated as “g” in the remainder of the paper.

- (6) a. *mohæmædʔæmin pæriruz*
 “Mohammadamin two days ago”
 b. mohæ | mæd | ʔæ | min | pæ | ri | ruz
 g1 | g2 | g3 | g4 | g5 | g6 | g7

The truncation point of the gates corresponds with syllable boundaries. Using Praat, syllable boundaries were indicated manually in Praat, then each gate was extracted from the original sound file by running a script.

As (6b) and Figures 3 and 4 demonstrate, at gate 7 the pre-wh part which is ambiguous with regard to sentence type is completely presented. The complete unambiguous version of each item (see 5) was also played. However, it was not presented immediately after gate 7 (i.e. the pre-wh part) of the corresponding item. All of the complete unambiguous versions 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 complete unambiguous version of an item immediately after hearing the pre-wh part of the same item can be practice for the participants in identifying the sentence type. Hearing the complete unambiguous version 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. The beginning and the end of the sentences were manually determined in Praat.

3.3. Procedure

A forced-choice sentence categorisation task was designed in E-prime 2.0.10 (Psychology Software Tools, 2012). Participants were seated in front of a computer in a quiet room. The experiment started with the emergence of the written instruction on the computer screen. Participants could take as much time as they wanted to read the instructions, and were allowed to ask questions about them if necessary. Next, they were familiarised with the task by means of a practice session. The practice session included two non-experimental items, i.e. two sets of seven gates generated as described in Section 3.2.3. The items were a declarative and a question read by one of the speakers from the

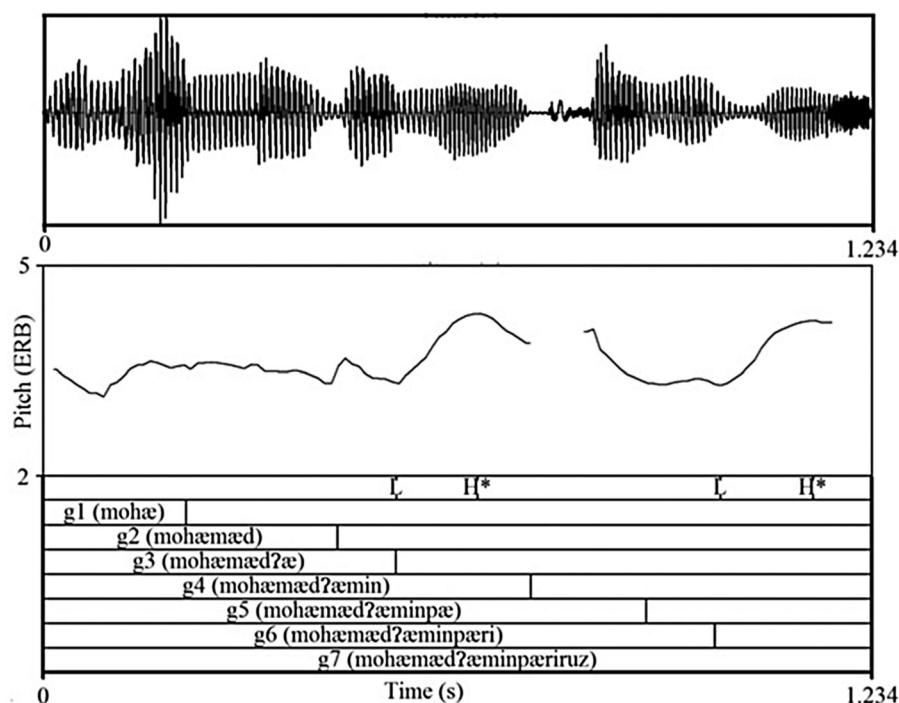


Figure 4. The seven gates of a question stimulus. The “L” and “H*” represent the valleys and the peaks of the realised pitch accents. The other tiers represent the gate boundaries. The letter g represents the word gate and the number designates the gate number.

production task (Shiamizadeh et al., 2018). The stimuli were played to the participants through Sennheiser PC 141 Headset headphones. For each sentence, the first seven gates were played 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. At the end of the practice session participants were presented with the complete unambiguous versions of the same stimuli. Participants were instructed to decide whether what they heard is going to be a wh-question or a declarative. After hearing each stimulus, they had four seconds to opt for either a wh-question or a declarative by pressing either V or M on the keyboard. To help participants to remember which key they needed to press for declaratives and wh-questions, a full stop (for declaratives) and the letter V and a question mark (for wh-questions) and the letter M appeared on two opposite sides (left and right) of the screen at the same time a stimulus was played to them. The right side of the screen corresponds with the M key on the keyboard while the left side of the screen corresponds with the V key of the keyboard. The order in which the full stop and the question mark and the corresponding letters (M or V) were displayed on the screen was fixed for individual participants, whereas it was counterbalanced across participants. After having decided on a sentence type, a question asking how confident the participants were about their response and a five-point confidence scale

appeared on the screen, where one means “not sure at all” and five “completely sure”. They had four seconds to indicate their confidence by choosing a number from one to five. Two seconds passed as the inter-stimulus interval. If participants did not give a response within four seconds, the experiment proceeded to the next stimulus automatically after two seconds. The presentation order of the items of the practice session was the same for all participants. They were allowed to do the practice session twice if they wanted. Having accomplished the practice session, participants embarked on the main part of the experiment when they felt ready. The main session of 320 items were divided into five blocks. Each of the first four blocks included 70 stimuli, comprising 10 sentences divided into seven gates. The final block contained the complete unambiguous version of the items presented in the previous four blocks. Therefore, block five included 40 stimuli. Participants were instructed to take at least a three-minute break between each two blocks. After the break, they were asked to press the space bar to continue with the next block. Every block started with a warm-up which consisted of two non-experimental items. The purpose of including warm-up items was to prepare participants for the new block after the break. The sequence in which the first four blocks were presented was randomised per participant. However, the fifth block was always presented at the end of the experiment to avoid a

Table 1. Perception of intended sentence type across gates and sentence type.

Gate number	Sentence type	Correct		Incorrect		Missing cases		Total	
		N	%	N	%	N	%	N	%
Gate 1	Declaratives	325	81.25	69	17.25	6	1.5	400	100
	Wh-questions	279	69.75	119	29.75	2	0.5	400	100
Gate 2	Declaratives	352	88.00	47	11.75	1	0.25	400	100
	Wh-questions	317	79.25	82	20.50	1	0.25	400	100
Gate 3	Declaratives	359	89.75	40	10.00	1	0.25	400	100
	Wh-questions	329	82.25	71	17.75	0	0	400	100
Gate 4	Declaratives	369	92.25	29	7.25	2	0.5	400	100
	Wh-questions	345	86.25	55	13.75	0	0	400	100
Gate 5	Declaratives	363	90.75	31	7.75	6	1.5	400	100
	Wh-questions	346	86.50	54	13.50	0	0	400	100
Gate 6	Declaratives	379	94.75	21	5.25	0	0	400	100
	Wh-questions	345	86.25	53	13.25	2	0.5	400	100
Gate 7	Declaratives	384	96.00	16	4.00	0	0	400	100
	Wh-questions	348	87.00	51	12.75	1	0.25	400	100
CUV	Declaratives	398	99.50	2	0.50	0	0	400	100
	Wh-questions	390	98	8	2	0	0	400	100

Note: CUV = Complete unambiguous version of the stimuli.

practice effect on sentence modality identification, as indicated above. The presentation order of the items within all blocks were randomised per participant. The procedure of the main session was identical to that of the practice session. The experiment took about 40 min to complete.

3.4. Data analysis

The responses, reaction time (RT) and confidence rating data were transferred from E-prime to SPSS version 22 (IBM SPSS, 2012). The response accuracy to declaratives and wh-questions was computed in terms of percentage correct and A' (Stanislaw & Todorov, 1999). Reaction times were calculated in terms of the time lapse between the stimulus offset and the response (all RT data are reported in seconds). Three separate two-way repeated measures ANOVAs (RM-ANOVA) were run on the accuracy, RT and confidence rating data in order to investigate the effect of sentence type, gate, and their interaction. The assumptions of the RM-ANOVA were met.

4. Results

4.1. Response accuracy

Table 1 gives 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 (75.50%) is above chance level ($t(19) = 29.417$, $p < 0.01$).

Responses are transformed to A' to correct for a possible response bias (Stanislaw & Todorov, 1999). The mean A' score for each gate is presented in Figure 5.

To investigate the effect of gate number, sentence type, as well as the interaction between gates and

sentence type on response accuracy, a two-way RM-ANOVA was run, with aggregated response as the dependent variable and gate number and sentence type as independent variables. The multivariate test demonstrated that the main effect of gate ($F(7,13) = 12.249$, $p < .001$; Wilks' Lambda = .135, $\eta_p^2 = .865$; response accuracy increases as the gate number increases) and sentence type ($F(1,19) = 7.577$, $p < .02$; Wilks' Lambda = .715, $\eta_p^2 = .285$; response accuracy to declaratives is higher than to questions) is significant. On the other hand, the interaction effect of sentence type and gates ($F(7,13) = 1.617$, $p > .05$; Wilks' Lambda = .535, $\eta_p^2 = .465$) on response accuracy is revealed to be insignificant. Pairwise comparison tests using Bonferroni correction (see Table 2) demonstrated that the differences between all gates except for gate 2 and 3, 3 and 4, 3 and 5, 4 and 5, 4 and 6, 5 and 6, 5 and 7, 4 and 7, 6 and 7 were significant ($p < .01$).⁷

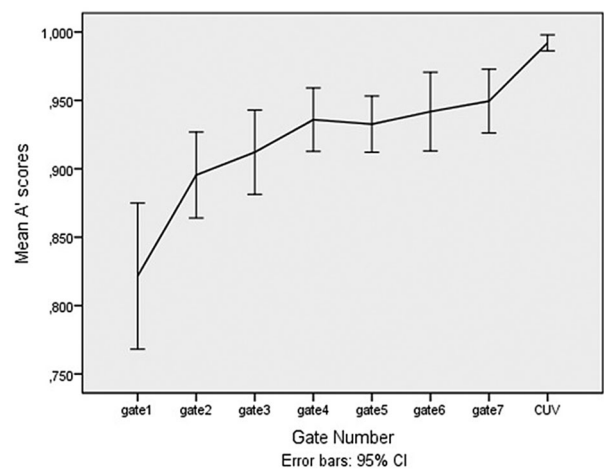


Figure 5. Mean A' scores across gates. CUV stands for complete unambiguous version of the stimuli.

Table 2. Results of pairwise comparison tests for response accuracy differences between gates (the result is based on Bonferroni correction test).

Gates	<i>p</i> value							CUV
	Gate 1	Gate 2	Gate 3	Gate 4	Gate 5	Gate 6	Gate 7	
Gate 1		.000*	.000*	.000*	.000*	.000*	.000*	.000*
Gate 2	.000*		.154	.006*	.005*	.001*	.000*	.000*
Gate 3	.000*	.154		.663	.607	.013*	.006*	.000*
Gate 4	.000*	.006*	.663		1.000	1.000	1.000	.001*
Gate 5	.000*	.005*	.607	1.000		1.000	.092	.000*
Gate 6	.000*	.001*	.013*	1.000	1.000		1.000	.030*
Gate 7	.000*	.000*	.006*	1.000	.092	1.000		.036*
CUV	.000*	.000*	.000*	.001*	.000*	.030*	.036*	

Note: * $p < 0.05$.

CUV = Complete unambiguous version of the stimuli.

4.2. Reaction time analysis

RT is calculated as the time lapse between the stimulus offset and the response (all RT data are reported in seconds). In cases where the response was given before the stimulus offset, we have a negative reaction time.⁸

As Table 3 illustrates, the RT to declaratives was shorter than the RT to wh-questions within each gate. According to Figure 6, the RT to stimuli decreases as the gate number increases, likely reflecting the increased availability of prosodic information as the gate number increases.

All RT data were submitted to a two-way RM-ANOVA with sentence type and gate as independent variables. According to the multivariate test, sentence type ($F(1,19) = 11.583$, $p < .01$; Wilks' Lambda = .621, $\eta_p^2 = .379$; longer RT to questions than to declaratives), gate ($F(7,13) = 38.080$, $p < .001$; Wilks' Lambda = .047, $\eta_p^2 = .953$; RT decreases as gate number increases) and the interaction of sentence type and gate ($F(7,13) = 4.512$, $p < .01$; Wilks' Lambda = .292, $\eta_p^2 = .708$; at gates 4, 5 and 7, RT to declaratives was significantly shorter than to

questions) 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 ($p > 0.5$). The p -value was adjusted for multiple comparisons using a Bonferroni correction.

4.3. Confidence rating

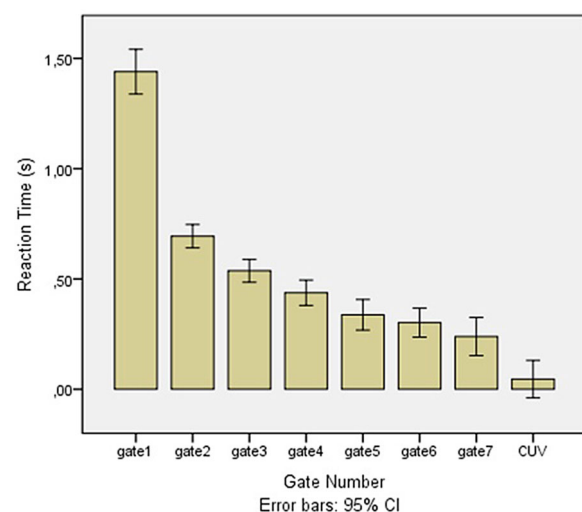
As observable in Figure 7, participants' confidence in their responses increased as the gate number also increased. This is in line with the results regarding response accuracy and RT to different gates, namely that response accuracy increased and RT decreased as the gate number increased.

A two-way RM-ANOVA was administered with sentence type and gate as independent variables and confidence rating as the dependent variable. The main effect of gate ($F(7,13) = 20.872$, $p < .001$; Wilks' Lambda = .082, $\eta_p^2 = .918$; confidence increases as the gate number increases) was revealed to be significant. However, the main effect of sentence type ($F(1,19) = 0.162$, $p > .05$;

Table 3. Mean reaction time (and standard deviation) (in s) for declaratives and wh-questions across gates.

Gate number	Sentence type	Reaction time Mean(SD)
Gate 1	Declaratives	1.40 (0.32)
	Wh-questions	1.48 (0.32)
Gate 2	Declaratives	0.67 (0.15)
	Wh-questions	0.71 (0.18)
Gate 3	Declaratives	0.53 (0.16)
	Wh-questions	0.54 (0.17)
Gate 4	Declaratives	0.41 (0.18)
	Wh-questions	0.46 (0.17)
Gate 5	Declaratives	0.28 (0.22)
	Wh-questions	0.39 (0.21)
Gate 6	Declaratives	0.29 (0.23)
	Wh-questions	0.31 (0.19)
Gate 7	Declaratives	0.16 (0.30)
	Wh-questions	0.31 (0.22)
CUV	Declaratives	0.02 (0.29)
	Wh-questions	0.07 (0.25)

Note: CUV = Complete unambiguous version of the stimuli.

**Figure 6.** Mean reaction time (in seconds) across gates. CUV stands for the complete unambiguous version of the stimuli.

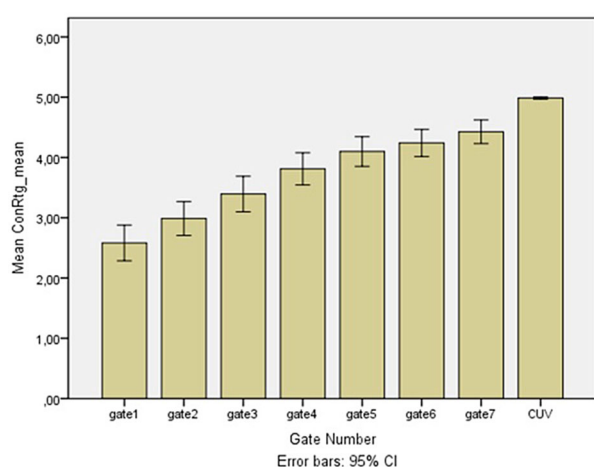


Figure 7. Mean confidence rating across gates. CUV stands for complete unambiguous version of the stimuli.

Wilks' Lambda = .992, $\eta_p^2 = .008$), and the interaction of sentence type and gate ($F(7,13) = 0.276$, $p > .05$; Wilks' Lambda = .871, $\eta_p^2 = .129$) were non-significant. Pairwise comparison tests indicated that the difference between all gates with respect to confidence rating is significant $p < .01$. The p -value was adjusted for multiple comparisons using a Bonferroni correction.

5. Discussion

The aim of this study was to investigate at which point in the pre-wh part of a sentence the distinctive prosodic correlates to sentence modality contrast enable participants to predict the sentence type. 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.

As Figures 3 and 4 illustrate, only two syllables are presented at gate 1. According to Shiamizadeh et al. (2018), the prosodic characteristic of questions available at gate 1 is the higher F0 onset. The significant difference between response accuracy to gate 1 and the other gates can be explained by the prosodic information available at gate 1. The significant difference in accuracy between gate 1 and gate 2 and gate 1 and gate 3 might be explainable by the steeper inclination and the decreased duration of the questions, which is perceptible when more syllables are audible.

The subject of the sentence was completely presented at gate 4 (see Figures 3 and 4). The pitch accent associated with the subject is presented at this gate. At gate 7, the adverb is entirely presented and the pitch accent realised on it is made audible (see also Figures 3 and 4). The larger excursion size of the pitch

accents realised on the subject and the adverb are the other prosodic features that characterise the pre-wh part in Persian wh-in-situ questions (Shiamizadeh et al., 2018). This can account for the significant difference in the response accuracy between gates 1 and 4, 1 and 5, 1 and 6 and 1 and 7. The difference in the accuracy between gates 2 and 4, 2 and 5, 2 and 6 can be possibly explained by the emergence of the subject pitch accent at gate 4. The audibility of the pitch accent on the subject and adverb, the shorter duration and the steeper inclination of questions can explain the significant difference between the response accuracy to gates 2 and 7.

The insignificant difference between the response accuracy to gates 3 and 4, 3 and 5 suggests that the larger excursion size of the subject pitch accent could not be the only reason for the difference between gates 3 and 6, and 3 and 7. Since the pitch accent on the adverb is not audible at gate 6 and the difference between gates 4 and 7 is not significant, the larger excursion of the adverb pitch accent cannot be mentioned as the only reason for the difference between gates 3 and 6, 3 and 7 as well. Therefore, we can suggest that the combination of the differences in inclination, duration and pitch accent excursion are the possible justifications for the differences between the response accuracy to gates 3 and 6, and 3 and 7.

Based on the insignificant increase in response accuracy from gate 4 to gate 7 (2.3%, $p > 0.5$; see Table 2), we can propose that the prosodic information until gate 4 provides a strong cue to sentence type identification.

The difference between RT to gate 5 and 6 is not significant. The pitch movement on the first two syllables of the adverb *pæriruz* "two days ago" can account for this insignificant decrease in RT. Gate 6 presents the pre-wh part of the sentence until the end of the second syllable of the adverb (see Figures 3 and 4). The first two syllables of the adverb *pæri* represent a female name in Persian. Since the word *pæri* is a content word, a pitch accent must be associated with its second syllable *-ri* (Mahjani, 2003; Sadat Tehrani, 2008). The syllable *-ri* is presented at gate 6. However, since the word *pæri* is part of the content word *pæriruz*, no pitch accent is realised on *-ri* (see Figures 3 and 4). It can be proposed that not hearing a pitch accent on *ri* makes listeners uncertain about the sentence type. This uncertainty implies that the participants need more time to decide on the sentence type.

Although sentence type identification was high (89.20%) at gate 4 and there was no significant increase in identification responses from gate 4 to gate 7, the highest confidence rating (4.45 on a scale of 5) in sentence type recognition is achieved at gate 7. This

implies that although listeners could correctly predict the sentence type at early gates, they may only confidently focus their attention on the process of response preparation at a later gate (that is, later in the utterance), when they are highly confident of the sentence type. Another implication of this finding is that it is possible that prediction can be reset as the listener progresses through the acoustic signal. This is in line with sentence processing theories (e.g. Hale, 2001; Levy, 2008) which state that predictions change as the sentence unfolds. Possible support for the resetting of predictions lies in the significant increase in confidence rating as the sentence unfolds, along with the presentation of gates. In other words, more distinctive prosodic correlates are presented as the sentence unfolds in gates.

Response accuracy to declaratives was shown to be higher than response accuracy to questions. Higher response accuracy to declaratives has been reported in earlier perception studies (Shiamizadeh et al., 2017a; Vion & Colas, 2006). In line with the results of other perception studies on the role of prosody in sentence type identification (Shiamizadeh et al., 2017a; Shiamizadeh, Caspers, & Schiller, 2017b; Vion & Colas, 2006), declaratives also have shorter reaction times in comparison to questions. A possible reason for the decreased RTs and the higher response accuracy to declaratives could be the higher frequency of declaratives in comparison to questions in daily conversation (as suggested by Van Heuven and Haan (2000) and Vion and Colas (2006)).

Comparing the result of the current experiment with that of the previous studies confirms Lindblom's (1990) hyper- and hypo-theory of speech production. According to Van Heuven and Haan (2000), the hyper- and hypo-theory of speech production (Lindblom, 1990) suggests that prosodic interrogativity marking will be weaker when lexico-syntactic interrogativity markers are available in the sentence, whereas prosodic interrogativity cues will be stronger when lexico-syntactic features of interrogativity are absent in the sentence. Similar to the distinction between yes-no questions and statements in Castilian Spanish, French and Neapolitan Italian and between declarative questions and statements in Dutch, the distinction between the pre-wh part in wh-in-situ questions and declaratives in Persian rests on intonation only (see Section 1.1.2). However, the accuracy of the identification of yes-no or declarative questions vs. statements in Castilian Spanish, Dutch, French and Neapolitan Italian upon the presentation of the complete sentence is higher than the accuracy of the identification of wh-questions vs. declaratives in Persian upon the presentation of the pre-wh part.⁹ This difference suggests that the prosodic cues will be stronger in globally ambiguous yes-no and declarative

questions in comparison to locally ambiguous wh-in-situ questions.

In addition to Lindblom's (1990) theory, the general result of this research corroborates several other proposals suggested in the literature. First, prosody plays a prominent role in processing syntactically ambiguous sentences (e.g. Beach, 1991; Beach et al., 1996; Carlson et al., 2001; Kjelgaard & Speer, 1999; Nagel et al., 1994; Snedeker & Trueswell, 2003; Warren et al., 1995) and models of spoken sentence processing may need to integrate the (online) use of prosody in interpreting these constructions (Beach, 1991). 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 (Snedeker & Trueswell, 2003). Third, prediction can be reset as more prosodic information is provided to the listener (Grosjean, 1983).

The first distinctive prosodic feature in Persian wh-in-situ questions is the F0 onset which has an efficient role in sentence type perception at the first gate (the first two syllables of the sentence). However, the studies presented in Section 1.1.2 do not report significant differences between the F0 onset of questions and statements. Upon the presentation of the first distinctive prosodic feature, the accuracy of identification of wh-questions as opposed to declaratives in this study is higher than the accuracy of the perception of interrogatives vs. declaratives in Dutch, French and Mandarin Chinese. Therefore, it may 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 other languages reported in the literature.

6. Conclusion

This study illustrates that the distinctive prosodic correlates to sentence modality contrast enable participants to predict the sentence type early in the utterance. The results corroborate some of the theories and proposals on the role of prosody in speech production (Lindblom, 1990) and the role of prosody (e.g. Snedeker & Trueswell, 2003) and prediction (e.g. Grosjean, 1983; Levy, 2008) in sentence processing.

There are some limitations on the generalizability of the results of the current experiment to models of language processing. Listeners will draw upon any and all information that may facilitate language processing (Grosjean, 1983). The amount of attentional resources that a listener can allocate to the process of perceiving a particular source of information is limited (cf. Norman

& Bobrow, 1975) and is different across different sources of information (Wales & Taylor, 1987). Wales and Taylor (1987) argued that fewer attentional resources are allocated to processes of intonation perception than processes of lexical or syntactic encoding. The stimuli of the current study have no lexical or syntactic cues to sentence type. Other cues to sentence type, e.g. visual cues (House, 2002), are absent as well since this experiment is conducted in a laboratory setting. This may lead listeners to devote more attentional resources to perception of prosody than the amount of attentional resources usually allocated to prosody when processing language outside of the laboratory (see also Vion & Colas, 2006).

Though the gating paradigm can determine whether prosodic information can be used by the listener to give a response in a laboratory setting, it cannot demonstrate if listeners use this information during online processing (Grosjean, 1983). Therefore, 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; see e.g. Liu, Chen, and Schiller's (2016) study in which a P300 effect was found to distinguish between question and statement processing in Mandarin Chinese. Furthermore, due to their fine-grained time resolution, EEG experiments can give additional insights into the time course of prosody processing.

Notes

1. The *wh*-phrase can optionally move to the earlier parts, including the beginning of the sentence (Abedi et al., 2012; Adli, 2007; Gorjian et al., 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdumian & Ganjavi, 2000; Mirsaedi, 2006; Toosarvandani, 2008) for non-syntactic reasons. These authors claim that the movement of the *wh*-phrase to earlier parts of the sentence is not triggered by the syntactic (+*wh*) feature. Therefore, Persian cannot be categorized as a *wh*-movement language. Adli (2007), Kahnemuyipour (2001), Karimi (2005), Karimi and Taleghani (2007), Lotfi (2003) and Toosarvandani (2008) claim that the *wh*-phrase moves to earlier parts of the sentence to receive contrastive focus. (1) is an example of a sentence in which the *wh*-phrase "*chi*" (what) moves to the beginning of the sentence to receive contrastive focus. This question occurs when the respondent answers the speaker's question "when did Maryam play?" with a response unrelated to the *wh*-phrase "when", e.g. Maryam played in the garden. Then the questioner may repeat his/her question as in (1). The declarative

and *wh-in-situ* question counterparts of (1) are given in (2a) and (2b) within the text.

- | | | | |
|----|---|-------------------------|---|
| 1. | <i>kej</i>
when
"When did Maryam play?" | <i>mærjæm</i>
Maryam | <i>bazi- kærd?</i>
play-do.PAST.3SG. |
|----|---|-------------------------|---|

2. In this paper, the competitors are statements and *wh*-questions in Persian.
3. The declarative question is sometimes considered as a subtype of yes-no questions (Quirk, Greenbaum, Leech & Svartvik, 1972). Formal markers of interrogativity are absent in the declarative question. Thus, lexico-syntactically, it is identical to the corresponding declarative sentence.
4. The values of accuracy percentage reported here are based on the figure that presents the percentage of correct response across sentence type in Van Heuven and Haan (2000).
5. The successive presentation nature of the gating paradigm may make listeners stick to their earlier choice, even when additional information to the identification of the target is provided (Craig & Kim, 1990; Grosjean, 1996b) and facilitate the identification of the target (see e.g. Cotton & Grosjean, 1984; Salasoo & Pisoni, 1985). This means that a gating experiment does not necessarily yield the same results as one-time stimulus presentation (e.g., Salasoo & Pisoni, 1985). Therefore, the current perception experiment based on the gating paradigm was preceded by the perception experiment based on one-time stimulus presentation (Shiamizadeh et al., 2017a).
6. 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.
7. A separate RM-ANOVA was run with gate number as the independent variable and A' scores as the dependent variable. A main effect of gate on A' was found ($F(7,13) = 7.698$, $p < .003$; Wilks' Lambda = .217, $\eta_p^2 = .783$). The result of the pairwise comparison tests using a Bonferroni correction was similar to the result of the pairwise comparison tests of the effect of gates on percentage of response accuracy: the differences between all gates except for gate 2 and 3, 3 and 4, 3 and 5, 4 and 5, 4 and 6, 5 and 6, 5 and 7, 4 and 7, 6 and 7 ($p < .01$).
8. 16.9% ($f = 1084$) of the stimuli (18.2% ($f = 581$) of declaratives and 15.7% ($f = 503$) of *wh*-questions) were responded to before the stimulus offset.
9. The exact value of the response accuracy upon the presentation of the complete sentence in German and Neapolitan Italian is not reported, but an inspection of the figures in the relevant papers suggests that response accuracy is around or above 95%.

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 Pls of that project, Lisa Lai-Shen Cheng and Jenny Doetjes. Two anonymous reviewers are thanked for their valuable comments. Special thanks are due to Daan van de Velde for his assistance in writing the E-prime script.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Niels O. Schiller  <http://orcid.org/0000-0002-0392-7608>

References

- Abedi, F., Moinszadeh, A., & Gharai, Z. (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. International Review of General Linguistics. Revue internationale De Linguistique Generale*, 120, 2256–2294.
- 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.
- Beach, C. M., Katz, W. F., & Skowronski, A. (1996). Children's processing of prosodic cues for phrasal interpretation. *The Journal of the Acoustical Society of America*, 99, 1148–1160.
- Boersma, P., & Weenink, D. (2014). *Praat: Doing phonetics by computer (version 6.0.04) [computer program]*. Retrieved from <http://www.praat.org/>
- 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.
- Carlson, K., Clifton, C., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, 45, 58–81.
- Carnie, A. (2007). *Syntax: A generative introduction* (2nd ed.). Oxford: Blackwell.
- Chomsky, N. (1977). On wh-movement. In P. W. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal syntax* (pp. 71–132). New York, NY: Academic Press.
- Cohen, A., & 't Hart, J. (1967). On the anatomy of intonation. *Lingua. International Review of General Linguistics. Revue internationale De Linguistique Generale*, 19, 177–192.
- Cotton, S., & Grosjean, F. (1984). The gating paradigm: A comparison of successive and individual presentation formats. *Perception and Psychophysics*, 35, 41–48.
- Craig, C. H., & Kim, B. W. (1990). Effects of time gating and word length on isolated word-recognition performance. *Journal of Speech Language and Hearing Research*, 33, 808–815.
- Di Cristio, A., & Hirst, D. J. (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* (pp. 268–271). Lund: Lund University.
- D'Imperio, M. (2000). The role of perception in defining tonal targets and their alignment (Doctoral dissertation), Ohio State University. Retrieved from https://etd.ohiolink.edu/etd.send_file?accession=osu1243021045&disposition=inline
- Face, T. L. (2004). The intonation of absolute interrogatives in castilian spanish. *Southwest Journal of Linguistics*, 23, 65–79.
- Face, T. L. (2005). F0 peak height and the perception of sentence type in castilian spanish. *Revista Internacional de Lingüística Iberoamericana*, 2, 49–65.
- Gorjian, B., Naghizadeh, M., & Shahramiri, P. (2012). Making interrogative sentences in English and Persian language: A contrastive analysis approach. *Journal of Comparative Linguistics and Literature*, 2, 120–124.
- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics*, 28, 267–283.
- Grosjean, F. (1983). How long is the sentence? Prediction and prosody in the on-line processing of language. *Linguistics*, 21, 501–529.
- 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.
- Grosjean, F. (1996b). Gating. *Language and Cognitive Processes*, 11, 597–604.
- 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, Spain.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Haan, J. (2001). *Speaking of questions: An exploration of Dutch question intonation (Doctoral dissertation)*, Leiden University. Retrieved from www.lotpublications.nl/Documents/52_fulltext.pdf
- Hale, J. (2001). A probabilistic early parser as a psycholinguistic model. In *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.
- House, D. (2002). *Intonational and visual cues in the perception of interrogative mode in Swedish*. Proceedings of the seventh international conference on spoken language processing (ICSLP), (pp. 1957–1960).
- IBM SPSS. (2012). *IBM SPSS. IBM Software Business Analytics*.
- Kahnemuyipour, A. (2001). On wh-questions in Persian. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, 46, 41–61.
- 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., & Taleghani, A. (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* (pp. 167–187). Amsterdam: Benjamins.
- 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.
- Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge: Cambridge University Press.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106, 1126–1177.

- Liberman, M., & Pierrehumbert, J. (1984). Intonational in-variance under changes in pitch range and length. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure* (pp. 157–233). Cambridge, MA: MIT Press.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403–439). Dordrecht: Kluwer.
- Liu, M., Chen, Y., & Schiller, N. O. (2016). Online processing of tone and intonation in mandarin: Evidence from ERPs. *Neuropsychologia*, 91, 307–317.
- Lotfi, A. R. (2003). Persian Wh-riddles. In C. Boeckx, & K. K. Grohmann (Eds.), *Multiple wh-fronting* (pp. 161–186). Amsterdam: Benjamins.
- Mahjani, B. (2003). *An instrumental study of prosodic features and intonation in modern Persian* (Unpublished master's thesis), University of Edinburgh.
- Megerdooomian, K., & Ganjavi, S. (2000). Against optional wh-movement. In V. Samiian (Ed.), *Proceedings of the western conference on linguistics: WECOL* (pp. 358–370). Fresno: California University.
- Mirsaeedi, A. (2006). *Wh-movement in Persian language* (Unpublished master's thesis), University of Isfahan.
- Nagel, H. N., Shapiro, L. P., & Nawy, R. (1994). Prosody and the processing of filler-gap sentences. *Journal of Psycholinguistic Research*, 23, 473–485.
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44–64.
- Petrone, C., & D'Imperio, M. (2008). Tonal structure and constituency in neapolitan Italian: Evidence for the accentual phrase in statements and questions. In P. A. Basbosa, S. Madureira, & C. Reis (Eds.), *Proceedings of the fourth conference on speech prosody* (pp. 301–304). São Paulo: Capes.
- 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.
- Petrone, C., & Niebuhr, O. (2014). On the intonation of German intonation questions: The role of the prenuclear region. *Language and Speech*, 57, 108–146.
- Psychology Software Tools. (2012). *E-Prime*. Pittsburg, PA.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A grammar of contemporary English*. Harlow: Longman.
- Sacks, H. (2004). An initial characterization of the organization of speaker turn-taking in conversation. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 35–42). Philadelphia, PA: John Benjamins.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735.
- Sadat Tehrani, N. (2008). *Intonational grammar of Persian* (Doctoral dissertation), University of Manitoba. Retrieved from www.collectionscanada.gc.ca/obj/s4/f2/dsk3/MWU/TC-MWU-2839.pdf
- Salasoo, A., & Pisoni, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory and Language*, 24, 210–231.
- Schegloff, E. A. (2006). Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In N. J. Enfield & S. C. Levinson, (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 70–96). Oxford: Berg Publishers.
- Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2017a). The role of prosody in the identification of Persian sentence types: Declarative or wh-question? *Linguistics Vanguard*. doi:10.1515/lingvan-2016-0085
- Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2017b). The role of F0 and duration in the identification of wh-in-situ questions in Persian. *Speech Communication*, 93, 11–19.
- Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2018). Do Persian native speakers prosodically mark wh-in-situ questions? *Language and Speech*. doi:10.1177/0023830917753237
- 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.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, 31, 137–149.
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., ... Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106, 10587–10592.
- Toosarvandani, M. (2008). Wh-movement and syntax of sluicing. *Journal of Linguistics*, 44, 677–722.
- 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.
- Vion, M., & Colas, A. (2006). Pitch cues for the recognition of yes-no questions in French. *Journal of Psycholinguistic Research*, 35, 427–445.
- Wales, R., & Taylor, S. (1987). Intonation cues to questions and statements: How are they perceived? *Language and Speech*, 30, 199–211.
- Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes*, 10, 457–486.
- Yang, Y. (2018). *The two sides of wh-indeterminates in mandarin: A prosodic and processing account* (Doctoral dissertation), Leiden University, Utrecht.