



Universiteit
Leiden
The Netherlands

Prosody and processing of wh-in-situ questions in standard Persian
Shiamizadeh, Z.

Citation

Shiamizadeh, Z. (2018, April 4). *Prosody and processing of wh-in-situ questions in standard Persian*. Retrieved from <https://hdl.handle.net/1887/60911>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/60911>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The following handle holds various files of this Leiden University dissertation:
<http://hdl.handle.net/1887/60911>

Author: Shiamizadeh, Z.

Title: Prosody and processing of wh-in-situ questions in standard Persian

Issue Date: 2018-04-04

**PROSODY AND PROCESSING
OF WH-IN-SITU QUESTIONS
IN STANDARD PERSIAN**

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
<http://www.lotschool.nl>

Cover illustration: Zohreh Shiamizadeh using Boersma, P. & D. Weenink. (2014).
Praat: Doing phonetics by computer (Version 6.0.04) [Computer program].
<http://www.praat.org/>.

ISBN: 978-94-6093-279-3
NUR 616

Copyright © 2018: Zohreh Shiamizadeh. All rights reserved.

**Prosody and processing of
wh-in-situ questions in standard
Persian**

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 4 april 2018
klokke 16:15 uur

door
Zohreh Shiamizadeh
geboren te Ghom, Iran
in 1984

Promotor: Prof. dr. N. O. Schiller (Leiden University)

Co-promotor: Dr. J. Caspers (Leiden University)

Promotiecommissie: Prof. dr. Y. Modarresi (Institute for Humanities and Cultural Studies, Tehran, Iran)

Prof. dr. V. J. J. P. van Heuven (University of Pannonia)

Prof. dr. J. Doetjes (Leiden University)

Dr. Y. Chen (Leiden University)

To my mother

Contents

Contents	vii
Acknowledgements	xi
1 Introduction	1
1.1 Preliminaries	1
1.2 Intonation of questions	1
1.3 Wh-questions	2
1.4 Background on Persian	3
1.5 Theoretical background	5
1.6 Research questions and approach	6
1.7 Hypotheses	8
1.8 Outline of the dissertation	8
2 Do Persian native speakers prosodically mark wh-in-situ questions?	9
Abstract	9
2.1 Introduction	10
2.2 Prosody of interrogatives	11
2.2.1 Interrogatives in different languages	11
2.2.2 Wh-questions in Persian	13
2.3 Research questions, hypotheses and approach	16
2.3.1 Research questions and approach	16
2.3.2 Hypotheses	18
2.4 Experiment	18
2.4.1 Method	18
2.4.1.1 Participants	18
2.4.1.2 Material	18
2.4.1.3 Procedure	20
2.4.2 Data analysis	21
2.4.2.1 Segmentation	21
2.4.2.2 Acoustic correlates	21
2.4.2.2.1 Definition of acoustic correlates	21

2.4.2.2.2 Measurement of the acoustic correlates	21
2.4.2.3 Statistical analysis	24
2.4.3 Results	25
2.5 Discussion and conclusion	27
3 The role of prosody in the identification of Persian sentence types:	
Declarative or wh-question?	31
Abstract	31
3.1 Introduction	32
3.1.1 Background	32
3.1.1.1 Prosodic correlates of Persian wh-in-situ questions	32
3.1.1.2 Theoretical background	33
3.1.1.3 Empirical background	34
3.2 Research questions, approach and hypotheses	37
3.3 Methodology	38
3.3.1 Participants	38
3.3.2 Material	39
3.3.2.1 Selection of the speakers	39
3.3.2.2 Selection of the stimuli	39
3.3.3 Procedure	40
3.3.4 Data analysis	41
3.4 Results	41
3.4.1 Accuracy	41
3.4.2 Reaction time	42
3.5 Discussion and conclusion	42
4 When is a wh-in-situ question identified in Persian?	45
Abstract	45
4.1 Introduction	46
4.1.1 Background	48
4.1.1.1 Prosodic correlates of Persian wh-in-situ questions	48
4.1.1.2 Empirical background	48
4.2 Research questions, approach and hypotheses	51

4.2.1 Research questions and approach	51
4.2.2 Hypotheses	51
4.3 Methodology	52
4.3.1 Participants	52
4.3.2 Material	52
4.3.2.1 Speaker selection	52
4.3.2.2 The stimuli	52
4.3.2.2.1 Selection of the stimuli	52
4.3.2.2.2 Number of the stimuli	54
4.3.2.3 Gating procedure	54
4.3.3 Procedure	56
4.3.4 Data analysis	57
4.4. Results	58
4.4.1 Response accuracy	58
4.4.2 Reaction time analysis	61
4.4.3. Confidence rating	63
4.5 Discussion and conclusion	65
5 The role of F0 and duration in the identification of wh-in-situ questions in Persian	67
Abstract	67
5.1 Introduction	68
5.1.1 Background	69
5.1.1.1 Production and perception of prosodic correlates of Persian wh-in-situ questions	69
5.1.1.2 Empirical background	70
5.2 Research questions, approach and hypotheses	70
5.2.1 Research questions and approach	70
5.2.2 Hypotheses	71
5.3 Methodology	72
5.3.1 Participants	72
5.3.2 Material	72

5.3.2.1 Speaker selection	72
5.3.2.2 Selection of the stimuli	72
5.3.2.3 Stylizing the pitch contour	74
5.3.2.4 F0 manipulation	76
5.3.2.5 Duration manipulation	77
5.3.2.6 F0 and duration manipulation	77
5.3.2.7 Resynthesized original sounds	78
5.3.3 Procedure	78
5.3.4 Data analysis	79
5.4. Results	80
5.4.1 Response accuracy	80
5.4.2 Reaction time analysis	82
5.5 Discussion and conclusion	83
6 Discussion and conclusion	85
6.1 Introduction	85
6.2 Theoretical implications	86
6.3 Suggestions for further research	88
References	91
Appendices	101
Appendix I	101
Appendix II	102
Appendix III	104
Appendix IV	106
List of figures and tables	107
List of figures	107
List of tables	108
English Summary	109
Samenvatting in het Nederlands	113
Summary in Persian	117
Curriculum vitae	119

Acknowledgements

Conducting the research reported in this thesis would not have been possible without the company and support from a number of people. First of all, I would like to express my gratitude to my supervisors Niels O. Schiller and Johanneke Caspers. I am grateful to Niels who helped me to get admission to LUCL, who placed trust in my ability to conduct research at PhD level and who encouraged me by his positive attitude. I would like to thank Johanneke for her continuous academic and mental support at all stages of my research, for the time she allotted to train me to conduct research and to write journal articles and this thesis. Johanneke, thank you for listening to me during the difficulties I experienced.

Special thanks to Jos Pacilly, Daan van de Velde, Kate Bellamy, Hamed Rahmani, Mortaza Taheri Ardali and Prof. Mahmoud Bijankhan. Jos, thank you for your technical support in writing Praat scripts. Daan, I am grateful to you for your advice on statistics, writing E-prime scripts and for translating the summary of this thesis to Dutch. Kate, my thanks go to you for proofreading my manuscripts and this thesis. Hamed, Mortaza and Prof. Bijankhan, I would like to give thanks to you for the inspiring discussions we had about the literature on Persian prosody. The anonymous reviewers of the manuscripts and the committee members are thanked for their valuable comments.

I would not be where I am without the encouragement and the support of my family, especially my mother. *Maman*, I could not have even dreamed of doing my PhD without your dedicated mental and financial support. My life in the Netherlands would not have been possible without the support and company of my aunt, my uncle and their children. My deepest thanks go to you for your invaluable support.

I would also like to give thanks to Behazin Bijani and his wife for their mental support and their advice during the harsh moments I had in my personal life during these years. My thanks go to the friends I met in the Netherlands and their families: Yasemin, Canan, Hatice, Gonul, Ayse, Fatma Zahra, Haji Lotfi, Haji Zolkof, Mina, Maryam, Tanaz, Zohreh, Fatima, Negin, Elaheh, Elektra, Sima, Zahra, Bahareh and Bahar. Thank you for your encouragement and support during these years. Last but not the least, I am grateful to all the participants in my experiments.

The experiments reported in this dissertation are related to the project “Understanding Questions” funded by the Netherlands Organization for Scientific Research (NWO). For discussion, I would like to thank the PIs of that project, Lisa Lai-Shen Cheng and Jenny Doetjes.

Chapter one

Introduction

1.1 Preliminaries

All human languages allow their speakers to ask questions (Chisholm, Milic, & Greppin, 1984). The speech act of questioning is prevalent in human communication. The typical purpose of asking a question is to elicit a response from the hearer, whereas a statement is uttered to present facts or beliefs. Utterances require no sign to be recognized as statements. However, questions may be signaled lexically, for example by the insertion of *wh*-phrases and/or other syntactic devices, such as inversion of the subject and finite verb. Lexico-syntactic features are not the sole indicators of interrogativity (Haan, 2001). The literature on question intonation reveals that intonation plays an influential role in characterizing questions (e.g. Hermann, 1942; Haan, 2001).

1.2 Intonation of questions

Though there is cross-linguistic variation in the lexico-syntactic devices marking interrogativity, intonation is suggested to be the near-universal defining property of questions (Haan, 2001). Having conducted a survey of 177 languages, Hermann (1942) argued for the presence of a high pitch somewhere in the sentence as a cross-linguistic property of question intonation.¹ The results of previous studies on question intonation across languages lead to the conclusion that this high pitch can be observed both locally, that is in the initial, medial or final portion of the sentence (Benkirane, 1998; Cheng & Rooryck, 2000; Gryllia, Doetjes & Cheng, 2016 September; Haan, 2001; Haan, Van Heuven, Pacilly, & Van Bezooijen, 1997; Hadding-Koch, 1961; Hadding-Koch & Studdert-Kennedy, 1964; Ishihara, 2005; Iivonen, 1998; Kitagawa, 2005; Lee, 2005; Mahjani, 2003; Sadat Tehrani, 2011; Shen, 1990; Uldall, 1962; Yang, Gryllia, Doetjes & Cheng, 2016a September) and globally, namely in the form of raised register and/or the absence of an F0 downtrend (Bolinger, 1982; Geluykens, 1986; Haan, 2001; Haan et al., 1997; Inkelas & Leben, 1990; Lindsey, 1985; Sadat Tehrani, 2011; Thorsen, 1980; Vaissière, 1983). In addition to pitch features, durational properties, such as shorter duration, have been reported as the defining features of interrogatives (Cangemi &

¹ According to Rialland (2004), polarity questions in 18 languages (in a database of 80 African languages) are not characterized by high tones and rising intonation. Thus, in this dissertation I do not consider the correlation between high pitch and questions as universal. Rather, we take it as language-specific.

2 Prosody and processing of wh-in-situ questions in standard Persian

D’Imperio, 2013; Lindsey, 1985; Jorissen, 2014; Van Heuven & Van Zanten, 2005; Yang et al., 2016a September).

1.3 Wh-questions

The syntactic literature offers different question types, one of which is the wh-question. They are called wh-questions because they are typically identified by phrases that start with the letters “wh” in English, e.g. *what*, *when*, *where* (Carnie, 2007). Wh-questions are considered a near universal question type (Sadock & Zwicky, 1985). They 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 bought a book.
b. What did Mary buy?

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 (see 2). One language characterized by in-situ wh-questions is Persian. In Persian, wh-questions are in-situ by default (Abedi, Moinzadeh, & Gharaei, 2012; Adli, 2010; Gorjian, Naghizadeh, & Shahramiri, 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdoomian & Ganjavi, 2000; Mirsaedi, 2006; Toosarvandani, 2008).²

² The wh-phrase can optionally move to earlier parts, including the beginning, of the sentence (Abedi et al., 2012; Adli, 2010; Gorjian et al., 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdoomian & 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. As such, Persian cannot be categorized as a wh-movement language. Adli (2010), 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. Example (1) constitutes a sentence in which the wh-phrase *chi* ‘what’ moves to the beginning of the sentence to receive contrastive focus. The declarative and wh-in-situ question counterpart of it is given in (2a) and (2b) within the text.

1. tʃi mæɾjæm diruz xæɾid?
what Maryam yesterday buy.PAST.3SG.
“What did Maryam buy yesterday?”

- (2) a. mæɾjæm diruz ketab xærid.
 Maryam yesterday book buy.PAST.3SG.
 “Maryam bought a book yesterday.”
- b. mæɾjæm diruz tʃi xærid?
 Maryam yesterday what buy.PAST.3SG.
 “What did Maryam buy yesterday?”

To reiterate, example (1b) is a wh-question in English while (2b) is a wh-question in Persian. The internal syntactic structure of (1b) and (2b) is presented in (3) and (4), respectively (Karimi, 2005).

- (3) $[_{CP} \text{What}_y \ [_{C'} \text{did}_i \ [_{TP} \text{Mary}_j \ [_{T'} t_i \ [_{VP} t_j \ [_{V'} \text{buy } t_y]]]]]]?$
- (4) $[_{CP} [_{TP} \text{mæɾjæm} \ [_{\text{diruz}} \ [_{VP} t_{\text{tʃi}} \ \text{xærid}]]]]]? \ [_{CP} [_{TP} \text{Maryam} \ [_{\text{yesterday}} \ [_{VP} \text{what} \ \text{buy.PAST.3SG}]]]]]? \ \text{“What did Maryam buy yesterday?”}$

The structural difference between wh-movement in (3) and wh-in-situ in (4) is evident: the wh-phrase in (3) occurs at the very beginning of the wh-question while in (4) it remains in its base-generated position. The complementizer phrase (CP) in both (3) and (4) carries a wh-feature (Carnie 2007; Cheng, 1991). This feature is checked by the overt movement of the wh-phrase to the specifier of the CP in (3) while in (4) the wh-feature is checked by the covert movement of the wh-phrase to the specifier of the CP (Karimi & Taleghani, 2007).

1.4 Background on Persian

Persian, also known as Farsi, belongs to the Iranian language family within the Indo-Iranian branch of the Indo-European family (Dabir-Moghaddam, 1982; Windfuhr, 1991; Karimi, 2003, 2005). Persian has several dialects, the three main representatives of which are spoken in Iran, Afghanistan and Tajikistan (Windfuhr, 1991; Toosarvandani, 2004). The Persian dialect spoken in Iran, Afghanistan and Tajikistan is known as Farsi, Dari and Tajiki respectively (Karimi, 2005). Farsi is the official language of Iran (Mahjani, 2003). Similar to Dari and Tajiki, Farsi has its own dialectal divisions (Windfuhr, 1991). The focus of this study is on standard Persian spoken in Iran and we will refer to it as Persian in this dissertation.

Persian words have prominence on a specific syllable. This prominence has been traditionally described as stress (e.g. Ferguson, 1957; Kahnemuyipour, 2003; Lazard, 1992; Mahjani, 2003; Mahootian, 1997; Rohany Rahbar, 2012; Sadat

4 Prosody and processing of wh-in-situ questions in standard Persian

Tehrani, 2008, 2009).³ Location of stress in Persian words has been discussed in several works (e.g. Kahnemuyipour, 2003; Lazard, 1992; Mahootian, 1997; Rohany Rahbar, 2012; Sadat Tehrani, 2008, 2009).⁴ Persian has been described as an intonation language (Esposito & Barjam, 2007; Mahjani, 2003; Mahootian, 1997; Sadat Tehrani, 2008; Towhidi, 1974). Towhidi (1974) studied the Persian intonation system within the theoretical framework of British tradition (Crystal, 1969). Esposito and Barjam (2007) used Tone and Break Indices (ToBI; Silverman, Beckman, Pitrelli, Ostendorf, Wightman, Price, Pierrehumbert, & Hirschberg, 1992) and Mahjani (2003) and Sadat Tehrani (2008) applied the autosegmental-metrical framework to study Persian prosody. The details of these studies are presented in Chapter 2.

Persian is described as an SOV language which allows changes in word order (Adli, 2010; Dabir-Moghaddam, 1982; Karimi, 1999, 2003, 2005; Lotfi, 2003; Mahjani, 2003; Rahmani et al., 2018; Sadat Tehrani, 2008). No study has been devoted to the reasons for changes in word order in Persian. Inspection of the data presented in Karimi (1999, 2005) shows that receiving contrastive focus is one of the reasons of changes in word order in Persian. However, to receive contrastive focus, a word does not necessarily move, it can remain in-situ and be interpreted as contrastive focus by changes in prosody (Karimi, 1999, 2005; Abolhasanizadeh et al., 2012; Mahjani, 2003; Rahmani et al., 2018; Sadat Tehrani, 2008, 2009; Taheri Ardali & Xu, 2012). Topic is mentioned as another reason for changes in word order by Karimi (1999, 2005).⁵

Similar to the reasons for changes in word order in Persian, as far as we know, no study has investigated how information structure is conveyed in Persian. Inspection of the data and the results of the studies by Abolhasanizadeh, et al. (2012), Mahjani (2003), Rahmani et al. (2018), Sadat Tehrani (2008, 2009) and Taheri Ardali and Xu (2012) suggest that contrastive focus in Persian can be conveyed by changes in prosody. Abolhasanizadeh et al. (2012), Rahmani et al. (2018), Tehrani (2008, 2009) and Taheri Ardali and Xu (2012) suggest that contrastive focus is conveyed by changes in the acoustic features of the pitch accent realized on the focused element. Inspection of the data presented in Karimi (1999, 2005) suggests that changes in word order and use of the morpheme “*fagat*” (only) can also convey information structure in Persian.⁶

³ Abolhasanizadeh, Bijankhan and Gussenhoven (2012) and Rahmani, Rietveld and Gussenhoven (2018) argued that the nature of prominence in Persian words is accent rather than stress. Since the focus of this dissertation is not on stress, we do not present the detail of the arguments.

⁴ A summary of the position of stress in Persian words is as follows. In nouns, adjectives and most adverbs, the final syllable carries the stress. In prefixed verbs, the stress is on the prefix. Simple verbs are stressed on the final syllable of the main constituent, i.e., the present stem, the past stem, or the participle. Compound verbs comprise of a non-verbal element and a verb. In compound verbs, the non-verbal element carries the stress.

⁵ Karimi (1999, 2005) argued that the moved element can be a topic or contrastive focus. She suggested that the stress that the moved element receives determines if the word is a focus or a topic, i.e. the focused element receives a stronger stress than the topic. However, she does not explain what she means by stress. She did not present empirical or quantitative data on the role of stress in the interpretation of the moved element as the focus or topic.

⁶ Recently, there has been some dispute about the conditions in which the morpheme “*fagat*” can convey information structure (in this case focus) in Persian (c.f. Rahmani et al., 2018).

1.5 Theoretical background

Fronted *wh*-questions, such as (1b), are syntactically marked by the presence of the *wh*-phrase in sentence-initial position. In contrast, *wh*-in-situ questions such as (2b) are not marked syntactically sentence-initially: the syntactic structure does not designate the clause type at the beginning of the sentence. The absence of lexico-syntactic markers (i.e. a *wh*-phrase) at the beginning of *wh*-in-situ questions raises questions regarding the role of prosody in the production and perception of *wh*-in-situ questions. Zubizarreta (1998) proposed that the difference between fronted and in-situ *wh*-questions lies in the way they are marked: while fronted *wh*-questions are syntactically marked by the overt movement of the *wh*-phrase to the specifier of the CP, in-situ *wh*-questions are prosodically marked.

Prosody is the defining feature of questions in the absence of a syntactic interrogativity marker (e.g. Bolinger, 1978; Ohala, 1983, 1984). The hyper- and hypo-theory of speech production (Lindblom, 1990) predicts that in interrogatives with lexico-syntactic features of interrogativity the prosodic cues will be weaker, whereas prosodic interrogativity marking will be stronger when lexico-syntactic features of interrogativity are absent or are fewer in the sentence. Gunlogson's (2008) and Lindblom's (1990) proposals can be extended to interrogatives where the lexico-syntactic interrogativity feature occurs later in the sentence, as in *wh*-in-situ questions: prosody marks the production of *wh*-in-situ questions in the absence of the *wh*-phrase in sentence-initial position.

Bolinger (1989) argued that the contrast between statements and questions is universally characterized by a dichotomy between relaxation (low, falling pitch), which marks statements, and tension (high, rising pitch), which characterizes questions. Hermann (1942) claimed that cross-linguistically questions tend to have raised pitch levels in comparison to statements. Similarly, Ohala's (1983, 1984) frequency code (FC) considers high pitch as a universal characteristic of questions in human languages. Ohala's FC suggests that the questioner is asking a question to prompt the respondent to provide some piece of information, so he/she needs to show his/her dependence on the addressee's knowledge and cooperation. Therefore, the questioner raises his/her pitch to signal his/her submissiveness or politeness to the respondent. Bolinger (1964, 1978, 1989), Bartels (1997) and lately Gunlogson (2008) have argued for a possible association of high F₀ values and questions. Similar to Ohala (1983, 1984), Bartels (1997) and Gunlogson (2008) proposed that rising intonation shows the questioner's dependency on the respondent. While the questioners use high F₀ to show their dependence on the respondent, the respondents use low F₀ to assert their authority.

According to Van Heuven and Van Zanten (2005), a faster speech rate (i.e. a shorter utterance duration) can convey the same message to the hearer as that expressed by a high F₀. They suggested that high pitch can be tied with fast speech rate in signaling questions. The association between acoustic correlates (high vs. low pitch and fast vs. slow speech rate) and sentence modality contrast (statements vs. questions) can be more prominent in sentences where the syntactic or lexical markers to sentence type are absent or appear later in the sentence (Lindblom, 1990). The implication is that prosody may play an important role in characterizing the

6 Prosody and processing of wh-in-situ questions in standard Persian

production of wh-in-situ questions in the absence of the lexico-syntactic interrogativity marker in sentence-initial position.

Processing conversational speech is an integral part of language processing. Listeners utilize any source of information that can facilitate the processing of conversation (Grosjean, 1983, 1996). Prosody is one of the multiple sources of information available in the processing of conversational speech. Grosjean's (1983, 1996) studies suggest that the role of prosody in processing becomes prominent in the absence of other sources of information, such as syntactic information regarding the clause type. In a normal conversation, the questioner raises a question with the purpose of eliciting a response from the addressee. The interlocutors who are engaged in a conversation rarely leave gaps between exchanges (Brazil, 1981; Sacks, 2004; Sacks, Schegloff, & Jefferson, 1974; Schegloff, 2006; Stivers, Enfield, Brown, Englert, Hayashi, Heinemann, Hoymann, Rossano, De Ruiter, Yoon, & Levinson, 2009). The implication here is that listeners draw on past and present information to process the sentence up to the point uttered by the speaker, as well as to predict upcoming events. In the absence of the syntactic interrogativity device in sentence initial position, prosody can aid the listener in predicting the sentence type and start the process of preparing a response. This suggests that prosody can cue the perception of wh-in-situ questions before the occurrence of the wh-phrase. Furthermore, several studies have indicated that language users may have implicit knowledge about the relationship between prosody and syntax, and are capable of using this knowledge to guide linguistic choices (e.g. Carlson, Clifton, & Frazier, 2001; Snedeker & Trueswell, 2003).

Several prosodic correlates, namely F0, intensity and duration, closely cooperate in the production of prosodic contrasts (Lehiste, 1970; Peng, Chatterjee, & Lu, 2012). Cooperation of the same cues is assumed to contribute to the perception of prosodic contrasts (Hazan & Rosen, 1991; Lehiste, 1970, 1976; Peng et al., 2012). Prosodic correlates do not contribute equally to the production and perception of prosodic contrasts, which are mainly conveyed through variations in F0, but this variation is often accompanied by variation in intensity and duration patterns (Cooper & Sorensen, 1981; Ladd, 1996; Lehiste, 1970, 1976; Peng et al., 2012; Van Heuven & Van Zanten, 2005). Similarly, the recognition of prosodic contrasts is mainly achieved by relying on F0 variation as the primary cue while other cues, such as duration and/or intensity, are considered secondary (Cangemi & D'Imperio, 2013; Lehiste, 1976; Peng et al., 2012). Perception studies (e.g. Fry, 1955; Hazan and Rosen, 1991) reported improved perception of prosodic contrasts in matching cue conditions in comparison to mismatching cue conditions. These studies suggest that prosodic correlates may not contribute equally to the perception of sentence modality contrast in Persian.

1.6 Research questions and approach

Four research questions are addressed in this dissertation:

1) Does the prosody of the pre-wh part, the post-wh part and the complete sentence distinguish Persian wh-in-situ questions from declaratives?

- 2) Does the prosody of the pre-wh part cue the correct perception of Persian wh-in-situ questions?
- 3) At which point in the pre-wh part can Persian wh-in-situ questions be identified?
- 4) Does the F0 or the durational cues or their cooperation contribute more to the identification of wh-in-situ questions in Persian?

Four separate experiments were conducted to investigate these research questions. The first experiment was a production experiment in which declaratives and wh-questions were elicited from Persian native speakers. In this experiment, the prosodic correlates of the pre-wh part, the post-wh part and the complete sentence in questions and declaratives were measured and compared.

Following the production experiment, three perception experiments were run. A forced choice sentence identification task was designed for each experiment. Part of the sentences elicited in the production experiment formed the base of the material of the perception experiments. The same material was used in all the perception experiments. However, since each perception experiment had different objectives, different procedures were adopted to prepare the stimuli for each perception experiment.

The first perception experiment was run to tackle the second research question. This experiment will be referred to as the sentence identification experiment. The purpose of this experiment was to investigate whether the prosodic differences between sentence types are perceivable by listeners. In this experiment the pre-wh part of declaratives and wh-in-situ questions was played to Persian native speakers in a forced choice sentence identification task. They were required to decide whether what they heard was going to be a wh-question or a declarative as quickly as they could.

To answer the third research question, a gating paradigm perception experiment was run (cf. Grosjean, 1980). This question was posed to investigate at which point in the pre-wh part Persian native speakers are able to identify the sentence type. To prepare the stimuli of this experiment, the pre-wh part of the sentences was divided into several gates. The gated stimuli were played to Persian native speakers in a forced choice sentence identification task. After hearing each gate, they were instructed to a) opt for either a declarative or a wh-question as the sentence type as quickly as possible, and b) report how confident they were about their decision by choosing a number between 1 and 5, where 1 is not sure at all and 5 is completely sure. This experiment will be referred to as the gating experiment in this dissertation.

Another perception experiment was designed to explore the final research question. This experiment will be called the manipulation experiment in the remainder of the dissertation. The F0 and durational features of the pre-wh part of the utterances were manipulated. The goal of this experiment was to weigh the relative contribution of the F0 features, the durational features and their cooperation to the identification of sentence modality contrast. Similar to the sentence identification experiment, Persian native speakers listened to the pre-wh part of declaratives and wh-in-situ questions. Upon hearing each stimulus, they were asked to decide whether what they heard was going to be a wh-question or a declarative as quickly as possible.

1.7 Hypotheses

On the basis of the theoretical framework (Bartels, 1997; Bolinger, 1964, 1978; Bolinger's dichotomy theory of relaxation and tension, 1989; Gunlogson, 2008; Hermann, 1942; Lindblom's hyper- and hypo-theory of speech production, 1990; Ohala's FC, 1983, 1984; Zubizarreta, 1998) and the findings of previous investigations on the prosody of interrogatives in different languages (Benkirane, 1998; Cangemi & D'Imperio, 2013; Cheng & Rooryck, 2000; Gryllia et al., 2016 September; Geluykens, 1986; Haan, 2001; Haan et al., 1997; Hermann, 1942; Inkelas & Leben, 1990; Iivonen, 1998; Jorrisen, 2014; Lindsey, 1985; Sadat Tehrani, 2011; Vaissière, 1983; Van Heuven & Van Zanten, 2005; Yang et al., 2016a September) we expect prosody to characterize the pre-wh part, post-wh part and the complete sentence in wh-questions as opposed to declaratives.

Drawing on the work of Grosjean (1983, 1996), Brazil (1981), and Snedeker and Trueswell (2003) and the results of previous studies on the role of prosody in the identification of interrogatives in different languages (e.g., Gryllia, Yang, Pablos, Doetjes & Cheng, 2016 September; Petrone & D'Imperio, 2011; Yang, Gryllia, Pablos, Doetjes & Cheng, 2016b September), I hypothesize that Persian native speakers can differentiate wh-in-situ questions from declaratives on the basis of prosodic correlates of the pre-wh part of the sentence.

Finally, we expect F0 variation to be the dominant contributor to the identification of wh-questions in comparison to statements (Cangemi & D'Imperio, 2013; Lehiste, 1976; Peng et al., 2012). An augmented effect of matching cue condition on sentence type categorization in comparison to the mismatching cue condition is also expected (e.g. Fry 1955; Hazan & Rosen, 1991).

1.8 Outline of the dissertation

This dissertation focuses on a series of experiments on the production and perception of a sentence modality contrast in Persian, namely statements as opposed to wh-in-situ questions. Chapters two to five are written as independent articles and have their own introduction and conclusion. Thus, there may be a certain amount of overlap between the introduction and the conclusion or other sections of these chapters and the general introduction and the conclusion of the dissertation. Chapter two deals with the production experiment on the prosodic correlates of declaratives and wh-in-situ questions. Chapter three focuses on the sentence identification experiment that investigates the role of prosody in the identification of wh-in-situ questions. The gating experiment investigating the point in the pre-wh part of the sentence where categorization is feasible is explained in chapter four. Chapter five presents the manipulation experiment on the relative contribution of the F0 and the duration to sentence modality contrast identification. The sixth and final chapter provides the general summary of the dissertation, the discussion of the overall results, the limitations of the study and suggestions for further research.

Chapter two

Do Persian native speakers prosodically mark wh-in-situ questions?⁷

Abstract

It has been shown that prosody contributes to the contrast between declarativity and interrogativity, notably in interrogative utterances lacking lexico-syntactic features of interrogativity (e.g. Van Heuven & Haan, 2000). Accordingly, it may be proposed that prosody plays a role in marking wh-in-situ questions in which the interrogativity feature (the wh-phrase) does not move to sentence-initial position, as, for example, in Persian. This chapter examines whether or not prosody distinguishes Persian wh-in-situ questions from declaratives in the absence of the interrogativity feature in sentence-initial position. To answer this question, a production experiment was designed in which wh-questions and declaratives were elicited from Persian native speakers. On the basis of the results of previous studies, we hypothesize that prosodic features mark wh-in-situ questions as opposed to declaratives at both the local (pre- and post-wh part) and global level (complete sentence). The results of the current study confirm our hypothesis that prosodic correlates mark the pre-wh part as well as the complete sentence in wh-in-situ questions. The results support theoretical concepts such as the frequency code (Ohala, 1984), the universal dichotomous association between relaxation and declarativity on the one hand and tension and interrogativity on the other (Bolinger, 1989), the relation between prosody and pragmatics (Bartels, 1997; Gunlogson, 2008), and the relation between prosody and encoding and decoding of sentence type (Grosjean, 1983, 1996; Snedeker & Trueswell, 2003).

⁷ This chapter is based on Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2018). Do Persian native speakers prosodically mark wh-in-situ questions? *Language and Speech*. doi: <https://doi.org/10.1177/0023830917753237>.

2.1 Introduction

Wh-questions are expressions that use wh-phrases to enquire about desired information. There are two types of wh-questions: fronted wh-questions and wh-in-situ questions. In fronted wh-questions, which occur in languages such as English, the wh-phrase moves to the beginning of the sentence to form a wh-question (Carnie, 2007; Chomsky, 1977; see example 1). On the other hand, there are languages, including Persian, in which the wh-phrase is not required to move to sentence-initial position in order to form a wh-question. In Persian, wh-questions are in-situ by default (Abedi, Moïnzadeh, & Gharaei, 2012; Adli, 2010; Gorjian, Naghizadeh, & Shahramiri, 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdoomian & Ganjavi, 2000; Mirsaedi, 2006; Toosarvandani, 2008).⁸ In wh-in-situ questions, the wh-phrase occurs at the same site as its declarative counterpart is expected to occur (2).

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

- (2) a mæɾjæm diruz ketab xæɾid.
 Maryam yesterday book buy.PAST.3SG.
 “Maryam bought a book yesterday.”
- b mæɾjæm diruz tʃi xæɾid?
 Maryam yesterday what buy.PAST.3SG.
 “What did Maryam buy yesterday?”

In fronted wh-questions (1b), the fronted wh-element indicates the clause type at the very beginning of the sentence. This way, the syntactic structure reveals the sentence type once the first word of the fronted wh-question is uttered. However, in the case of wh-in-situ questions (2b), the syntactic structure does not designate the clause type at the beginning of the sentence.

⁸ The wh-phrase can optionally move to the earlier parts, including the beginning of the sentence (Abedi et al., 2012; Adli, 2010; Gorjian et al., 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdoomian & 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 (2010), Kahnemuyipour (2001), Karimi (2005), Karimi & 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. The declarative and wh-in-situ question counterparts of it are given in (2a) and (2b) within the text.

1. tʃi mæɾjæm diruz xæɾid?
 what Maryam yesterday buy.PAST.3SG.
 “What did Maryam buy yesterday?”

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 or are fewer in the sentence. In addition, Ohala (1983, 1984), Bolinger (1978) and lately Gunlogson (2008) state that prosody is the defining property of a declarative question. Ohala's (1983, 1984), Bolinger's (1978), Gunlogson's (2008), and Lindblom's (1990) proposals can be applied to interrogatives where the lexico-syntactic interrogativity feature occurs later in the sentence, as in wh-in-situ questions. The absence of the wh-phrase in sentence-initial position in Persian wh-in-situ questions raises the question, therefore, as to whether or not the prosody of a sentence is indicative of the sentence type. In addition, the occurrence of the interrogativity marker at a later point in the sentence brings focus to the role of prosody in characterizing the pre-wh part (i.e. the part of the sentence before the wh-phrase) as compared to the post-wh part (i.e. the part of the sentence after the wh-phrase) of the sentence.

The organization of the chapter is as follows. Section 2 provides an overview of the available literature on the prosody of interrogatives in different languages and on wh-questions in Persian. Section 3 contains the details of the experiment. Section 4 elucidates the analysis of the data and Section 5 discusses the results and concludes the paper.

2.2 Prosody of interrogatives

This section first presents an overview of the literature on the prosody of interrogatives in different languages (Section 2.2.1) and after that on the prosody of Persian wh-in-situ questions (Section 2.2.2).

2.2.1 Interrogatives in different languages

Research on the acoustic features of questions in different languages has concentrated mainly on the prosodic features of yes-no questions and declarative questions. Some work has been done on the acoustic correlates of the prosody of fronted and wh-in-situ questions.

Hermann (1942) argued that interrogatives are universally marked by the presence of a high pitch somewhere in the sentence. Subsequent studies demonstrated that high pitch may manifest itself both locally, e.g. in the initial, medial or final portion of the sentence (Benkirane, 1998; Cheng & Rooryck, 2000; Haan, 2001; Haan, Van Heuven, Pacilly, & Van Bezooijen, 1997; Gryllia, Cheng & Doetjes, 2016; Hadding-Koch, 1961; Hadding-Koch & Studdert-Kennedy, 1964; Ishihara, 2005; Iivonen, 1998; Kitagawa, 2005; Sadat Tehrani, 2008, 2011; Thorsen, 1980; Uldall, 1962; Yang, Gryllia, Doetjes & Cheng, 2016a September) and globally, e.g. in the form of a raised register and/or the absence of F0 downtrend (Bolinger, 1982; Geluykens, 1986; Haan, 2001; Haan et al., 1997; Inkelas & Leben, 1990; Lindsey, 1985; Sadat Tehrani, 2008, 2011; Vaissière, 1983).

12 Prosody and processing of wh-in-situ questions in standard Persian

According to Hadding-Koch (1961), Hadding-Koch and Studdert-Kennedy (1964) and Uldall (1962), interrogatives can be distinguished from declaratives by the presence of a terminal rise in Swedish and American English. The studies conducted by Bolinger (1982) and Geluykens (1986) revealed that a raised pitch register level differentiates interrogatives from declaratives in American English. Thorsen (1980) showed that the absence of F0 downtrend differentiates interrogatives from declaratives in Danish.

Different types of questions in Dutch are marked by several (different) prosodic features as opposed to statements: a) a higher level of pitch register marks yes-no, declarative and wh-questions, b) a steeper F0 downtrend marks wh-questions and an F0 inclination marks yes/no and declarative questions, c) a higher F0 onset marks yes-no and wh-questions, d) a greater excursion size of pitch accents marks yes-no and declarative questions, e) a final rise marks yes-no and declarative questions and some instances of wh-questions (Haan, 2001; Haan et al., 1997) and f) a shorter duration marks declarative questions (Van Heuven & Van Zanten, 2005).

Sadat Tehrani (2008, 2011) reported that echo questions and tag questions have a high boundary tone. Based on qualitative analysis, Sadat Tehrani (2008, 2011) concludes that Persian leading yes-no questions have a higher pitch register as opposed to declaratives and yes-no questions are marked by a final rise, a greater excursion size of the pitch accents, less declination, a higher pitch register and final lengthening.

In Finnish (Iivonen, 1998) and Western Arabic (Benkirane, 1998) questions are reported to have higher onsets than the corresponding statements. Questions in Hausa have a raised pitch register level (Inkelas and Leben, 1990) and shorter duration in comparison to declaratives (Lindsey, 1985). In Manado Malay and Orkney English yes-no questions have a decreased duration in comparison to declaratives (Van Heuven & Van Zanten, 2005). Cangemi and D'Imperio (2013), in their study on a Neapolitan regional variety of Italian, found that durational differences between questions and statements are localized in specific portions of the sentence (i.e. the first and the last segment), but that the global duration of questions does not differ from that of declaratives.

According to Lee (1990), wh-in-situ questions in Korean usually end with a fall, but a rise can also occur at the end. The wh-phrase in Korean wh-in-situ questions carries the last accentual phrase of the sentence and the post-wh part is de-accentuated (Jun & Oh, 1996).

Similar to Korean wh-in-situ questions, the wh-phrase in Tokyo Japanese wh-in-situ questions carries the emphatic accent (Ishihara, 2005; Kitagawa, 2005). This emphatic accent consists of a sharp rise in F0 followed by a fall on the wh-phrase. The post wh-phrase in Japanese wh-in-situ questions is characterized by a post-focus F0 reduction, which suppresses all lexical accents and continues until the end of the sentence. Ishihara (2005) calls this prosodic pattern Focus Intonation (FI), i.e. the F0-peak of the wh-phrase is boosted (focus F0-boosting), while the F0-peaks of the post-wh-phrase are significantly reduced (post-focus F0-reduction). The end of wh-in-situ questions in Japanese is marked by a rise in F0 (Ishihara, 2005; Kitagawa, 2005).

According to Yang et al. (2016a September), several prosodic features differentiate Mandarin wh-in-situ questions containing the wh-phrase preceded by

“dianr” from their matching declaratives, namely: 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 declaratives. The authors suggested that the smaller F0 range in the post-wh part of wh-questions can be an indication of post-focal compression caused by the wh-phrase. Jorrissen’s (2014) data on wh-in-situ questions in Mandarin Chinese demonstrated a shorter duration of the complete sentence compared to their declarative counterparts.

French wh-in-situ questions exhibit a rising intonation, similar to the intonation of yes-no questions in the same language (Cheng & Rooryck, 2000). However, several other researchers (e.g. Beyssade, Delais-Roussarie, & Marandin, 2007; Hamlaoui, 2008; Hamlaoui, 2011; Zubizarreta, 2003) have argued against the idea that French wh-in-situ questions are required to have a rising intonation contour. Similar to Mandarin Chinese wh-in-situ questions, the pre-wh part (i.e. the subject and the verb) in French wh-in-situ questions is marked by F0 and durational properties as opposed to the same part in declaratives (Gryllia et al., 2016). The first syllable of declaratives with monosyllabic subjects has a longer duration in comparison to the matching wh-question. In addition, the low tone (L) realized on the verb is higher and its penultimate syllable is longer in declaratives. There is also a pitch rise on the final syllable of the verb in declaratives while pitch remains flat and low on the corresponding syllable in wh-questions. Finally, the wh-word carries an emphatic accent.

Summarizing the results of studies on interrogatives in different languages, we can state that there appears to be considerable cross-linguistic evidence that interrogativity is associated with local and global prosodic characteristics that set this sentence type apart from declaratives. As such, this study intends to investigate whether the prosodic encoding of interrogatives can be generalized to Persian wh-in-situ questions.

2.2.2 Wh-questions in Persian

A few studies have been conducted on the intonation system of Persian, all of which revolve around the phonological description of the accentual structure of different sentence types, including wh-questions and declaratives. This section presents a summary of these phonological studies.

Towhidi (1974) studied the Persian intonation system within the theoretical framework of British tradition (Crystal, 1969). He defined Persian intonation based on prosodic and paralinguistic features, dividing speech into Tone Units which consist of at least one syllable. Each Tone Unit is divided into prehead, head, nucleus and tail. The presence of prehead, head and tail is optional while one nucleus must be present in each Tone Unit. He identified three major categories of tones for Persian, i.e. simple, compound and complex.

Mahootian (1997) identified five types of intonation patterns for Persian: rising-falling, mid-rising, low-rising, mid-falling and high-falling. High-falling intonation is typical of wh-questions. It starts at a high level, and falls at the end of the question. According to Mahootian, the intonation peak in wh-questions is on the

wh-phrase. A possible reason for the occurrence of the intonation peak on the wh-phrase might be the focus nature of the wh-phrase in wh-questions.

Mahjani (2003) applied the autosegmental-metrical (AM) framework to study the prosody of different types of Persian sentences. He stipulated an accentual phrase (AP), an intermediate phrase (iP) and an intonation phrase (IP) as the levels of the Persian intonation system. The AP is the smallest unit of the Persian prosodic system, consisting generally of one content word. The AP is represented by either a L+H* or a H* pitch accent. Mahjani proposed that a L+H* pitch accent is assigned to polysyllabic words which are not initially stressed. Initially stressed content words and monosyllabic content words are assigned the H* pitch accent. He posited that the starred tone of the AP is assigned to the accented (stressed) syllable of the word and the L tone is associated with the syllables preceding the accented syllable. Stress is usually realized on the last syllable of words in Persian (Kahnemuyipour, 2003; Sadat Tehrani, 2008). AP and its boundary tone account for the F0 movement within a content word. Based on Mahjani, one or more APs form the iP. The end of an iP is demarcated by low (L-) or high (H-) tones, which are called phrase accents. The phrase accent characterizes the pitch from the last pitch accent to the end of the iP (for a definition see Sadat Tehrani, 2008).⁹ The IP as the highest level of the Persian intonation system is composed of at least one iP. It can have either a low (L%) or a high (H%) boundary tone (Mahjani, 2003). IP boundary tones define the F0 contour from the end of the iP boundary tone to the end of the IP (Sadat Tehrani, 2008).

Esposito and Barjam (2007) used Tone and Break Indices (ToBI; Silverman, Beckman, Pitrelli, Ostendorf, Wightman, Price, Pierrehumbert, & Hirschberg, 1992) labeling conventions to investigate prosody of different types of questions including wh-questions in Persian. They proposed a two-level intonational structure for questions: an accentual phrase (AP) and an intonation phrase (IP). The domain of an AP is a single content word. Each AP contains a pitch accent and ends with either a low or a high tone. The pitch accent is realized as L+H*, L+^H* (^H* denoting an extra high rise), H* or ^H *. The starred tone is associated with the stressed syllable of the word and the unstarred tone is associated with the syllables preceding the stressed syllable. One or more APs form an IP, the end of which is marked by an L-L%, L-H%, H-^H% or an H-L% boundary tone. An AP accounts for the F0 movement within a content word and the IP and its boundary tone define the F0 contour from the end of the last AP to the end of the IP.

Sadat Tehrani (2008) examined the intonation of Persian sentences within the autosegmental-metrical framework. He argued for a two-level hierarchy for the Persian intonation system. This two-level hierarchy is composed of an accentual phrase (AP) and an intonational phrase (IP). Tehrani considered the AP as the smallest unit of Persian prosody. The AP can consist of a content word or a compound noun. He proposed that the regular AP of Persian has the pitch accent L+H* or H* and every content word in Persian is associated with a pitch accent. The L+H* pitch accent is associated with polysyllabic words with the stress on the final syllable. The H* pitch accent is assigned to monosyllabic content words and initially

⁹ Sadat Tehrani's (2008) model for Persian intonation does not include the level iP, but he does provide a clear definition of iP.

stressed words. The right edge of the AP is marked by either a low (l) or a high (h) boundary; if the AP contains the nuclear pitch accent (NPA, the last pitch accent in an IP), it carries a low (l) boundary tone, otherwise it carries a high (h) boundary tone. The pitch accent of an AP accounts for the F0 changes within a content word and the boundary tone of an AP handles the F0 changes between the last tone of the pitch accent and the end of the AP. The next level of Persian prosody is the IP. It is composed of one or more APs, one of which contains the NPA. The right edge of an IP is marked by a low (L%) or high (H%) boundary tone. The prosodic structure proposed by Sadat Tehrani (2008) does not include the level intermediate phrase and its tone, the phrase accent. He suggested that the AP boundary tone can account for the part of the F0 contour that Mahjani (2003) referred to as iP and its tone: the F0 contour from the last pitch accent to the end of the IP. Figure 2.1 presents an example of the prosodic structure of a simple Persian sentence.

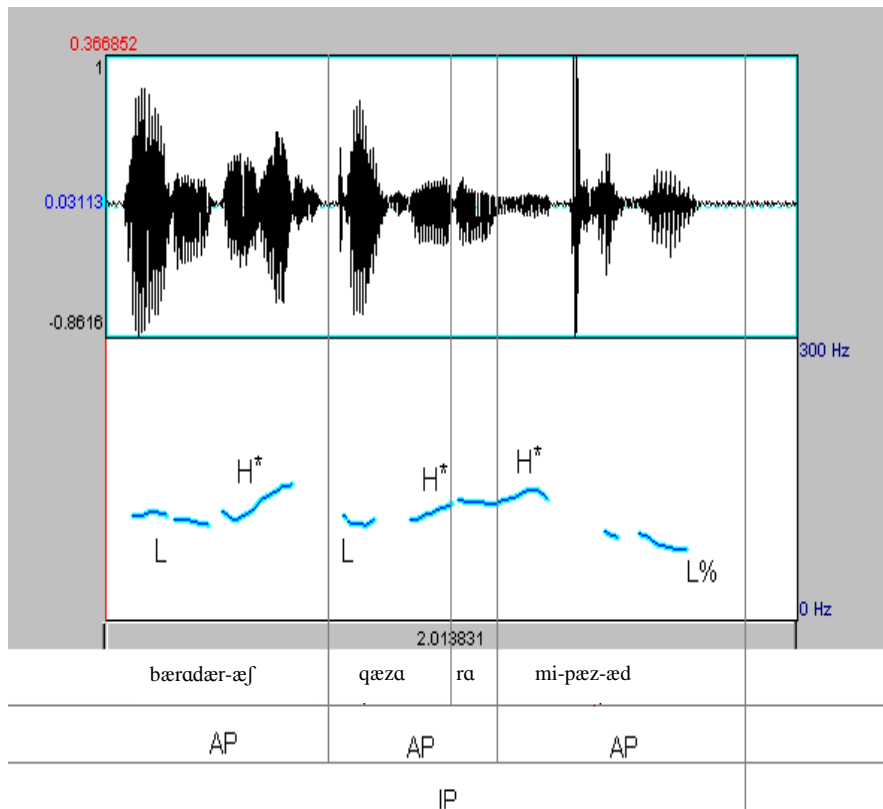


Figure 2.1. Prosodic structure of a simple sentence in Persian (from Mahjani (2003, p. 38)). Three accentual phrases (AP) comprise an iP and the iP comprises an IP. Mahjani (2003) does not mark the iP in this figure since there is only one iP in the prosodic structure of this sentence which is dominated by an IP. The English equivalent of the sentence is “His brother cooks the food”.

According to Mahjani (2003), Esposito and Barjam (2007) and Sadat Tehrani (2008), the intonational structure of wh-questions is similar to that of declaratives in that the IPs of both sentence types contain a series of APs and end with a L% boundary tone. The place of the NPA in a declarative sentence depends on the syntactic structure of the sentence and the verb type (for details of the place of the NPA in declaratives the reader is referred to Sadat Tehrani, 2008). However, the wh-word in wh-questions attracts the NPA as it is the most prominent word regarding the information structure of the sentence. The part of the sentence following the NPA in both declaratives and wh-questions is de-accented. According to Sadat Tehrani (2008), in wh-questions and declaratives with a contrastive focus item, the NPA is assigned to the contrastive focus item. Therefore, the wh-phrase in a wh-question with a contrastive focus item will not carry the NPA unless it is the contrastive focus of the sentence.

Previous studies on Persian prosody did not provide a comparative account of the acoustic features of wh-in situ questions and declaratives. The current study is an attempt to systematically explore and compare global and local acoustic features of Persian statements and wh-questions with a focus on the pre- and post-wh part of the sentence. Studying the differences in the acoustic correlates of the pre- and post-wh part in declaratives and interrogatives is novel and important. The prosody of wh-in-situ questions is an understudied area of research in general, and there is no study on their prosodic correlates in Persian. The present study can reveal if prosodic properties of wh-in-situ questions tend to localize in the pre-wh part only, since prosodic information is most needed where the syntactic ambiguity is maintained, or whether they are distributed throughout the sentence.

2.3 Research questions, hypotheses and approach

2.3.1 Research questions and approach

The present study is concerned with the intonational properties of interrogativity in Persian wh-in-situ questions, with the main aim of investigating whether the prosodic correlates of interrogativity characterize the pre-wh part in wh-questions. The interrogative marker in Persian wh-in-situ questions (i.e. the wh-phrase), which is the lexical cue to sentence type, occurs in the same position as the non-interrogative phrase in a declarative sentence (see example 2). The pre-wh parts in wh-in-situ questions and matching declaratives are lexically and syntactically identical. The pre-wh part precedes the lexical cue to the sentence type. This means that sentence type ambiguity resides in the pre-wh part. The main focus of the experiment is on the role of prosody in characterizing the pre-wh part in the production of wh-in-situ questions as opposed to declaratives. On the basis of the dominant focus of the experiment, the first research question (RQ1) is formulated as: do global and local F0 properties and the duration of the pre-wh part mark wh-in-situ questions as compared to declaratives? The global F0 properties refer to pitch register and the overall F0 trend and the local F0 properties refer to F0 onset and the size of the pitch accents of the pre-wh words. The answer to RQ1 not only sheds light on the role of prosody in encoding sentence type, but also lays the ground for

further investigation into the role of prosody in decoding sentence type before the appearance of the wh-phrase.

Similar to the pre-wh parts, the post-wh parts in wh-in-situ questions and matching declaratives are lexically and syntactically the same. However, the post-wh part follows the lexical cue to sentence type. This means that sentence type ambiguity is not maintained in the post-wh part, rather that ambiguity is resolved as soon as the wh-phrase occurs. RQ1 taps into the role of prosody in disambiguating the syntactically ambiguous part of the sentence. However, it cannot illustrate if prosodic characteristics of wh-in-situ questions are distributed throughout the sentence or tend to localize in the pre-wh part. Therefore, another research question (RQ2) was added: do F0 properties and the duration of the post-wh part characterize wh-in-situ questions?¹⁰ The F0 properties in this question refer to pitch register, overall trend of F0 and F0 offset. The presence of the wh-phrase causes de-accentuation in the post-wh part in wh-questions, but there can be pitch accent(s) in the post-wh part in declaratives depending on the non-interrogative phrase (Sadat Tehrani, 2008). Therefore, the excursion sizes of the pitch accents in the post-wh part are not compared. RQ1 and RQ2 enable us to compare the role of prosody in characterizing wh-in-situ questions before and after the occurrence of the lexical cue to the sentence type.

Investigating the prosodic correlates of the pre- and post-wh parts (RQ1 and RQ2) gives information on the role of prosody in characterizing questions at a more local level. However, if and how prosodic correlates characterize questions at a more global level is also unclear. Therefore, the final research question (RQ3) was formulated: do global F0 properties and the duration of the complete sentence distinguish in-situ wh-questions from declaratives? The global F0 properties in RQ3 refer to pitch register and overall trend of F0.

RQ2 and RQ3, along with RQ1, provide a complete picture of the prosodic informativity of the entire sentence. This picture contributes to the investigation of the applicability of the hyper- and hypo-theory of speech (Lindblom, 1990) to interrogatives where lexico-syntactic features of interrogativity appear later in the sentence.

Though the pre- and post-wh parts in declaratives and matching wh-questions are lexically and syntactically the same, the wh-phrase and the non-interrogative phrase necessarily consist of different words. As such, the acoustic features of the wh-phrase in wh-questions and the non-interrogative counterpart in declaratives are not compared.

To answer the three research questions, a question-response task was designed in which 40 declaratives and 40 wh-questions were elicited from 18 Persian native speakers.

¹⁰ This study originally investigated only RQ1. RQ2 and RQ3 were added as additional RQs to investigate local (i.e. F0 offset) and global prosodic correlates of the post-wh part and the complete sentence, following the comments of two anonymous reviewers.

2.3.2 Hypotheses

On the basis of the findings of previous studies on the prosody of interrogatives in different languages, we expect wh-questions to be prosodically marked by global prosodic features of (i) less declination (Haan, 2001; Haan et al., 1997; Sadat Tehrani, 2011; Vaissière, 1983), (ii) a raised level of pitch register (Bolinger, 1982; Gluykens, 1985; Haan, 2001; Haan et al., 1997; Hermann, 1942; Inkelas & Leben, 1990; Lindsey, 1985; Sadat Tehrani, 2011), (iii) a shorter duration (Jorrisen, 2014; Lindsey, 1985; Van Heuven & Van Zanten, 2005), and local prosodic features of (i) a higher F0 onset (Benkirane, 1998; Haan, 2001; Hermann, 1942; Iivonen, 1998), and (ii) a higher F0 offset (Haan, 2001) in the absence of the interrogativity marker in sentence-initial position. A greater excursion size of the pitch accents is expected to be associated with the pre-wh words in wh-questions according to Haan's (2001) and Haan et al.'s (1997) study on Dutch declarative questions and Sadat Tehrani's (2008, 2011) qualitative report on Persian yes-no questions.

We further expect both the pre- and post-wh parts to be characterized by (i) a shorter duration (Jorrisen, 2014; Lindsey, 1985; Van Heuven & Van Zanten, 2005; Yang et al., 2016a September), and (ii) a raised level of pitch register. Based on the general trend of F0 in questions reported in Haan (2001), Haan et al. (1997), Thorsen (1980) and Vaissière (1983), less declination is predicted for the pre-wh part. However, based on the de-accentuation that is reported for the post-wh part in different languages, including Persian (Mahjani, 2003; Sadat Tehrani, 2008, 2011), a steeper declination is expected for the post-wh part.

2.4 Experiment

2.4.1 Method

2.4.1.1 Participants

Eighteen native speakers of standard Persian (9 males and 9 females) between the ages of 24 and 42 years participated in the production experiment. All of them were university students or university lecturers at different universities in Tehran.

2.4.1.2 Material

The materials of this experiment represent two main conditions: (i) wh-in-situ questions and (ii) declaratives. We composed a corpus of 40 sentences for each condition, which means that each participant produced 80 sentences. The structure of the wh-questions and declaratives used in this experiment is presented in (3) and (4), respectively.

(3) Subj	Adv	Wh-phrase	Verb
(4) Subj	Adv	ADO/IDO/AdjT/AdjM/AdjP	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. ADO, IDO, AdjT, AdjM and AdjP will be referred to as declarative wh-phrase counterparts (DWCs). Different categories of DWCs in example (4) are separated by slashes to illustrate that each declarative includes only one of the categories of DWCs. An example of a declarative and a matching wh-question is given in (5a) and (5b). Stressed syllables are underlined in the phonetic transcription.¹¹

(5) a. mohæmædʔæmin pæriruz ʔæsɾ ʃenɑ-kærd.
 Mohamadamin two days ago afternoon swim- do.PAST.3SG
 ‘Mohamadamin swam in the afternoon two days ago.’

b. mohæmædʔæmin pæriruz kej ʃenɑ-kærd?
 Mohamadamin two days ago when swim- do.PAST.3SG
 ‘When did Mohamadamin swim two days ago?’

The sentences in both conditions were structured so as to be minimally different in order to provide the best comparison across conditions. Moreover, sentences were composed of the same number of words and syllables in both conditions. In order to avoid the effect of pitch perturbation caused by obstruents (Ladd, 2008), we tried to use words consisting of sonorants as much as possible. However, in some cases it was not possible to avoid certain stops, e.g. /k/ in the word *kodza* ‘where’.

As illustrated in examples (3) and (4), the questions and declaratives are both constructed of four constituents, three of which are the same (the subject, the adverb and the final verb), and only the third constituent differs: it is either a wh-phrase or a DWC. Appendix I presents the structure and the list of the words used in both sentence types. There are two variations for each category of the DWC in declaratives (2 Subj × 2 Adv × 2 DWCs × 5 verbs), hence there are 40 statements. The DWCs were varied so the result will not be limited to only one variation of argument or adjunct. In wh-in-situ questions, we have the same two subjects, two adverbs, five wh-phrases and matching verbs, and all sentences are presented twice (2 Subj × 2 Adv × 5 wh-phrases × 2 repetitions) resulting in 40 wh-question stimuli. The total number of stimuli elicited from all participants is 1440 (40 stimuli per sentence type × 2 sentence types × 18 speakers).

¹¹ For more information on the position of stress in Persian words, see e.g. Kahnemuyipour (2003), Lazard (1992), Mahootian (1997), Rohany Rahbar (2012) and Sadat Tehrani (2008).

2.4.1.3 Procedure

Participants were recorded using a microphone (Sennheiser PC 141 Headset) and a digital recorder (M-Audio MicroTrack II) in a quiet room. The target sentences were elicited from participants in a question-answer task. Declaratives and wh-questions were put in two separate blocks. These blocks were presented in a counter-balanced order, whereby half of the participants received the questions first while the other half were presented with the declaratives first. The sequence in which the sentences within each block were elicited was randomized per participant. Prior to conducting the main experiment, participants took part in a practice session. They were instructed that they would see a question and the target sentence on the computer screen. Simultaneously, they heard the question (also shown on the computer screen) read to them by a trained speaker, namely the first author of this article. The participants were asked to read the sentence (either a declarative or a wh-question) in response to the question they heard. The basic form of the questions asked by the experimenter to elicit wh-question stimuli from participants is given in (6).

(6) You know that Mohamadamin swam two days ago. In order to know when he swam, what would you ask?

The wh-question produced by the participants in response to (6) is (5b). All of the declaratives were produced in response to the question “what happened?”.¹² Each stimulus accompanied by a spoken question was presented one at a time. Participants were asked to act as naturally as possible when producing the sentences. They were also told that if they made a mistake, such as an unusual pause when producing the sentence or incorrect pronunciation of the words, they should produce the sentence again. The entire session took about 25 minutes for each participant and they were given a five-minute break in the middle of their session.

¹² As a result of the elicitation method, the declarative stimuli are all-new information focus sentences. In wh-questions the wh-phrase, as the information focus of the sentence, attracts the NPA and this causes deaccentuation in the post-wh part, i.e. the verb. However, in the declaratives, the NPA can be on the non-interrogative phrase or on the verb (Sadat Tehrani, 2008). In Persian every content word is produced with a pitch accent (see Section 2.2.2). An inspection of Taheri Ardali and Xu (2012) and Sadat Tehrani's (2009) results suggests that information structure does not affect the presence of pitch accents on the words preceding the NPA. Rahmani, Rietveld and Gussenhoven (2018) show that every content word preceding the NPA of the sentence retains its pitch accent, even if it contains given information. Therefore, the words in the pre-wh part of both declaratives and wh-in-situ questions have pitch accents and can thus be compared (see Section 2.3.1).

2.4.2 Data Analysis

2.4.2.1 Segmentation

The sentences were analyzed and segmented in Praat version 6.0.04 (Boersma & Weenink, 2014). The beginning and the end of the sentences were determined and the boundaries of the pre- and post-wh parts were indicated manually. The subject plus the adverb construct the pre-wh part of the sentence in both declaratives and wh-questions. Next, the pre-wh part was segmented into the subject and the adverb. The post wh-part was not segmented.

2.4.2.2 Acoustic correlates

2.4.2.2.1 Definition of the acoustic correlates

This section presents the definition of the acoustic features which are the target of the current study. F0 onset and offset are the F0 of the first and the last voiced frame of a segment (phoneme, word, sentence; Haan, 2001). Pitch register is defined as the pitch mean of the sentence following Rietveld and Vermillion (2003). Declination 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). The slope of the regression line is taken as a measure of declination (Lieberman, Katz, Jongman, Zimmerman, & Miller, 1985; Swerts, Strangert, & Heldnert, 1996). 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). The pitch accents in Persian are described as H* or L+H* (see Section 2.2.2), in which the starred tone is associated with the stressed syllable. Pitch excursion can be defined as “the size of a local pitch movement” (Ladd, 2008, p. 69). The following section explains how these phonetic correlates were measured.

2.4.2.2.2 Measurement of the acoustic correlates

According to the findings of the literature on interrogatives in several languages and based on our predictions, separate scripts were run in Praat to extract the slope of the regression line, the pitch mean and the duration of the pre-wh part (see Figures 2.2 and 2.3), the post-wh part and the complete sentence.¹³ The excursion size of the

¹³ An anonymous reviewer commented that since the post-wh parts do not tonally match in the two sentence types, their prosodic correlates cannot be compared in detail. S/he suggested using narrow information focus declaratives (with the NPA on the non-interrogative phrase) instead of all-new information focus declaratives. This would indeed solve the problem that in some cases a pitch accent is realized after the non-interrogative phrase in the declaratives. However, the post-wh parts of the two sentence types would still not be completely comparable prosodically. The reason is that in a small number of cases ($f = 97, 13.47\%$) the phonological shape of the post-wh part in wh-questions is different from its matching declarative, i.e. declaratives all end in L% while some wh-questions have a H%

pitch accents associated with the pre-wh words (see Figures 2.2 and 2.3) and the F0 onset and the F0 offset were also measured. The term "pre-wh word" refers to the subject and the adverb in both declaratives and wh-questions. All F0 values are expressed in Equivalent Rectangular Bandwidth (ERB), which is held to be the psychophysically most relevant F0-scale in intonation languages (Hermes & Van Gestel, 1991).

Regression techniques were adopted to inspect the global trend of the F0 of the sentence as these techniques are reproducible and can be objectively verified (Van Heuven & Haan, 2000; Lieberman et al., 1985; Swerts et al., 1996). A regression line was automatically calculated using all F0 points of the pitch contour of the pre-wh part, the post-wh part and the complete sentence as a function of time. The slope of the regression line was computed as a measure of the global trend of the F0 of the pre-wh part, post-wh part and the complete sentence. The boundaries of the pre-wh part, the post-wh part and the complete sentence correspond with the beginning and end of the regression line computed for them.

With respect to the F0 onset and offset, taking the first and the last measurable F0 point of the sentence respectively can be subject to estimation errors. Therefore, regression techniques were applied to also calculate the F0 onset and F0 offset of the sentence. The F0 level at the beginning and the end of the regression line was considered as the F0 onset and F0 offset of the sentence, respectively. Relying on the regression line of the pre-wh part to measure the F0 onset can cause the F0 onset measurement to be more heavily affected by the excursion size of the pitch accents of the pre-wh words. Similarly, relying on the regression line of the post-wh part to measure the F0 offset can cause the F0 offset measurement to be more heavily affected by the de-accentuation in the post-wh part of the sentence. The pitch means of the pre-wh part, post-wh part and the complete sentence were calculated as a measure of pitch register of the relevant part, following Rietveld & Vermillion (2003).

To accurately determine the location and F0 value of the valleys and peaks of the pitch accents, two steps were taken. Firstly, the pitch contour was stylized using a 1 semitone frequency resolution. Every subject and every adverb carried a L+H* pitch accent, as predicted by the phonological account of Persian intonation (cf. Mahjani, 2003; Sadat Tehrani, 2008). The L and H* points were indicated by hand in the stylized pitch contour. Then the points designated as the valley (L) and the peak (H*) were automatically moved to the maximum F0 and the minimum F0 in the original F0 contour using Praat functions (this method was adopted from Sadat Tehrani, 2009). The F0 value of the peak was subtracted from the F0 value of the valley of each pitch accent by running a script in Praat. The subtraction provided us with the excursion size of the pitch accents realized on the pre-wh words.

boundary tone. Therefore, even if the declarative stimuli were produced with narrow focus, we still could only make global comparisons for the post-wh part (and the complete sentence).

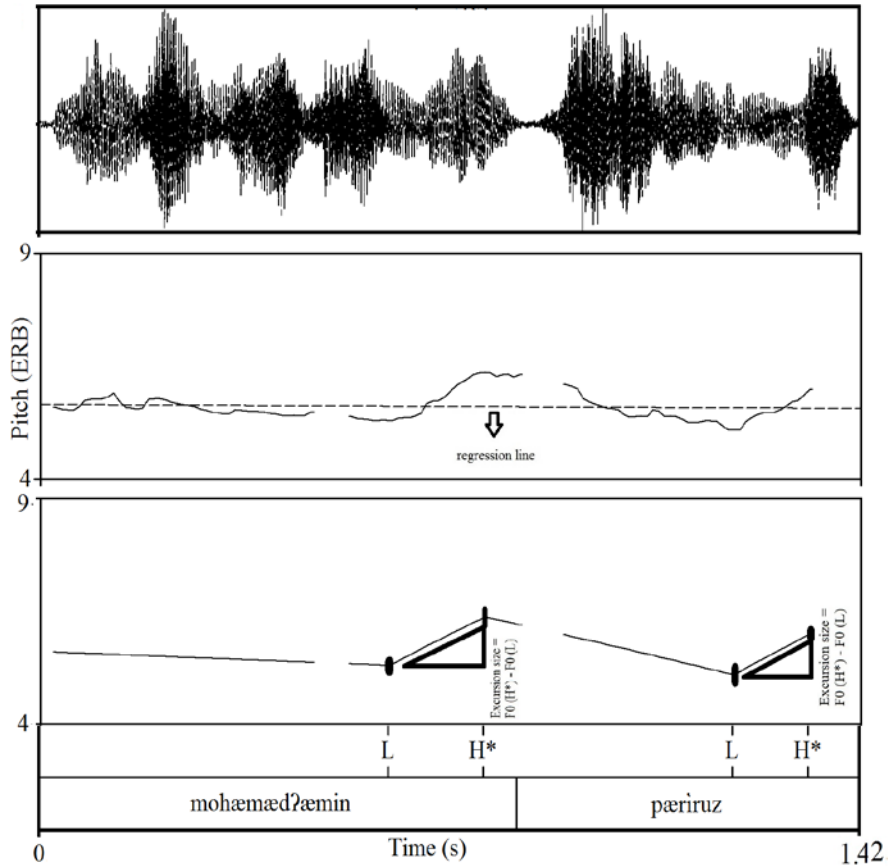


Figure 2.2. The acoustic correlates measured in the pre-wh part of a declarative sentence. In the second panel, the solid line is the pitch contour and the dotted line is the regression line. “L” and “H*” represent the valleys and the peaks of the realized pitch accents. The second tier represents the word boundaries. In the pitch stylized 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-stylized pitch contour is presented along with the regression line.

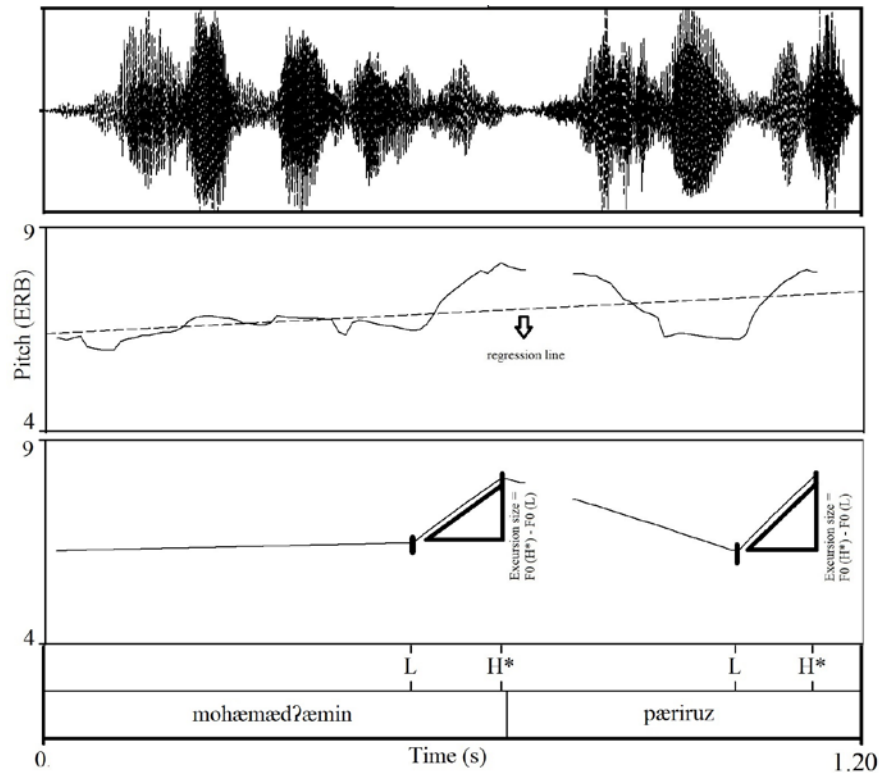


Figure 2.3. The acoustic correlates measured in the pre-wh part of a question. In the second panel, the solid line is the pitch contour and the dotted line is the regression line. “L” and “H*” represent the valleys and the peaks of the realized pitch accents. The second tier represents the word boundaries. In the pitch stylized 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-stylized pitch contour is presented along with the regression line.

2.4.2.3 Statistical analysis

Before running any kind of statistical analysis, all dependent variables were checked for normality of distribution. To answer the research questions, a multivariate repeated measures analyses of variance (RM-MANOVA) was conducted, with sentence type as the independent variable and the slope of the regression line, the pitch mean and the duration of the complete sentence, the excursion size of the pre-wh words, the pitch mean and the slope of the regression line of the pre-wh part and the slope of the regression line, the pitch mean and the duration of the post-wh part

as the dependent variables. The assumptions of repeated measures analysis of variance were met.

One of the conditions of the RM-MANOVA is that the correlation between dependent variables must not be higher than 0.80 (Field, 2009). Since the pitch mean of the complete sentence and the F0 onset ($r = 0.918$, $p < .001$), the duration of the pre-wh part and that of the complete sentence ($r = 0.849$, $p < .001$) and the pitch mean of the complete sentence and the F0 offset ($r = 0.831$, $p < .001$) were highly correlated, the duration of the pre-wh part, the F0 onset and the F0 offset were not included in the RM-MANOVA. Relying on the correlation between the pre-wh part and the sentence duration, the result of the univariate analysis for the sentence duration holds true for the univariate analysis of the pre-wh part duration. Based on the same logic, the result of the univariate analysis for the pitch mean of the complete sentence is applicable to the univariate analysis for the F0 onset and the F0 offset (Field, 2009).

2.4.3 Results

The multivariate test revealed a significant effect of the sentence type on the acoustic features of the pre-wh part, the post-wh part and the complete sentence, ($F(10,8) = 22.325$, $p < .001$; Wilks' Lambda = .035, $\eta_p^2 = .965$). As Table 1 illustrates, univariate tests for the individual variables indicated a significant effect of sentence type on the F0 excursion size of the pre-wh words, the slope of the regression line and F0 mean of the pre-wh part and the duration and the F0 mean of the complete sentence. The effect sizes (η_p^2) suggest that there is a substantial effect of the independent variable on all dependent variables (effect sizes larger than .13 count as significant (Pallant, 2013)).

According to Table 2.1, a larger pitch excursion size can be attributed to the subject and the adverb in wh-questions. The upward trend of the F0 (the slope of the regression line) of the pre-wh part is steeper in questions than in declaratives.¹⁴ The larger pitch excursion size of the subject and the adverb in wh-questions can explain this steeper upward trend of F0 of the pre-wh part in questions in comparison to the same part in declaratives. A higher F0 mean is another characteristic of the pre-wh part in wh-questions.

The complete sentence in wh-questions is characterized by a higher F0 mean and a shorter duration in comparison to declaratives (cf. Table 2.1).¹⁵

¹⁴ This upward trend can be an artifact of fitting a single regression line through the pitch points, as a result of the size of the pitch accents. Using two separate regression lines – one fitted through the low and one through the high pitch points – leads to somewhat different results for the pre-wh part of declaratives. The slope of the lower regression line of the pre-wh part in questions shows inclination – so we can suggest that the reported uptrend in the pre-wh part of wh-questions is not an artifact of the size of the pitch accents – while the pre-wh part of the declaratives shows declination. For the post-wh part and the complete sentence the results are comparable to those reported above. The reader is referred to appendix II for the relevant data.

¹⁵ A high final boundary tone (H%) in questions may cause the slope of the single regression line of the complete sentence to go up. The number of wh-questions ending in a high boundary tone in the current dataset is 13.47% ($N = 97$). Since the prevailing pattern of the final boundary tone in wh-questions is L%, i.e. 86.53% ($N = 623$), reanalysis of the regression line excluding the final boundary tone was not done.

Table 2.1. Results of the univariate tests and descriptive statistics for the acoustic correlates of the pre-wh part, the post-wh part and the complete sentence for the two sentence types.

Variable	F	df	p	η_p^2	M SD (Decl)	M SD (Wh-q)	N
Subject pitch excursion (ERB)	28.536	(1,17) ^a	.000**	.627	0.991 0.327	1.478 0.517	1440
Adverb pitch excursion (ERB)	60.471	(1,17)	.000**	.781	0.710 0.210	1.325 0.339	1440
Slope of pre-wh part regression line (ERB/sec)	48.535	(1,17)	.000**	.741	0.068 0.139	0.585 0.354	1440
Slope of post-wh part regression line (ERB/sec)	0.281	(1,17)	.603	.016	-1.155 0.633	-2.573 1.985	1440
Slope of sentence regression line (ERB/sec)	4.363	(1,17)	.052	.204	-0.179 0.157	-0.318 0.304	1440
F0 mean of the pre-wh part (ERB)	88.966	(1,17)	.000**	.840	4.830 1.402	5.720 1.587	1440
F0 mean of the post-wh part (ERB)	0.000	(1,17)	.999	.000	4.440 1.198	5.183 1.395	1440
F0 mean of the complete sentence (ERB)	80.883	(1,17)	.000**	.826	4.731 1.332	5.582 1.505	1440
Duration of the post-wh part (sec)	0.635	(1,17)	.437	.036	0.831 0.085	0.816 0.087	1440
Duration of the sentence (sec)	143.486	(1,17)	.000**	.894	2.637 0.201	2.213 0.190	1440
Duration of the pre-wh part (sec) ^b	-	-	-	-	1.362 0.100	1.107 0.112	1440
F0 onset (ERB) ^b	-	-	-	-	4.955 1.462	5.923 1.697	1440
F0 offset (ERB) ^b	-	-	-	-	4.458 1.141	5.187 1.364	1440

Note. Decl = declaratives; Wh-q = wh-in-situ questions.

^a The values in brackets give the degrees of freedom for condition and participants.

^b Since the duration of the pre-wh part, the F0 onset and the F0 offset were not included in the RM-MANOVA (see Section 2.4.2.3), only the descriptive statistics of these variables are reported.

**p < .01.

However, to see if the results are influenced by the H% appearing in a minority of questions, another RM-MANOVA was run on the data excluding the questions which end in H% and their corresponding declaratives (see Appendix III). The results show that the H% does not influence the general effect of sentence type on the dependent variables, including the slope of the regression line of the complete sentence.

Referring to the positive high correlation between the duration of the complete sentence and the pre-wh part and between the F0 onset and the F0 mean of the complete sentence (see Section 2.4.2.3), we can argue that a shorter duration of the pre-wh part and a higher F0 onset characterize wh-questions.

Based on the positive high correlation between the F0 mean of the complete sentence and the F0 offset (see Section 2.4.2.3), a higher F0 offset can be claimed to be the prosodic characteristic of wh-questions in comparison to declaratives.

The statistics indicated that all of the pre-wh part acoustic features measured in this study mark the pre-wh part in wh-questions as opposed to declaratives. However, among the post-wh part acoustic features measured in this study only F0 offset characterizes the post-wh part in wh-questions.¹⁶ These results imply that a) in comparison with the post-wh part, the pre-wh part in a wh-question is prosodically more distinct from the pre-wh part in the corresponding declarative, and b) the acoustic correlates of the pre-wh part are mainly responsible for the prosodic characteristics of the complete sentence in wh-questions as opposed to declaratives.

2.5 Discussion and conclusion

The present study is concerned with the intonational properties of interrogativity in Persian wh-in-situ questions with the aim of investigating whether a) the prosodic correlates of interrogativity characterize the - syntactically ambiguous - pre-wh part, b) the presence of the interrogativity marker (wh-phrase) affects the distribution of intonational properties of interrogativity in the pre- and post-wh part, and c) the prosodic correlates of interrogativity set Persian wh-in-situ questions apart from their matching declaratives at a more global (complete sentence) level.

The results of the current study confirm the prediction that a) the prosodic correlates in the pre-wh part of the sentence, namely a higher F0 onset, a higher pitch mean, a steeper F0 upward trend and a greater excursion size of the pre-wh words differentiate wh-questions from declaratives, and b) wh-questions are made distinct from declaratives by a raised register level and a shorter duration at a more global level. In contrast with our prediction, except for F0 offset, the prosodic correlates in the post-wh part do not play a significant role in marking wh-questions. As opposed to the dominant role of prosody in the pre-wh part, where sentence type ambiguity is maintained, the role of prosody is diminished in the post-wh part. This

¹⁶ As mentioned in Section 2.3.1, the presence of the wh-phrase causes de-accentuation in the post-wh part in wh-questions, but there can be a pitch accent in the post-wh part in declaratives depending on the non-interrogative phrase, i.e. DWCs in this study (Sadat Tehrani, 2008). The absence of the pitch accent in the post-wh part in questions vs. the presence of the pitch accent in the post-wh part of some declaratives may influence the prosodic properties of the post-wh part. The results of the RM-MANOVA showed that the main effect of sentence type ($F(3,15) = 0.238, p > .05$; Wilks' Lambda = .955, $\eta_p^2 = .045$) and the interaction effect of the sentence type and DWCs on the prosodic properties of the post-wh part measured in this study ($F(9,9) = 1.491, p > .05$; Wilks' Lambda = .401, $\eta_p^2 = .599$) is not significant. Thus, it can be proposed that the variation in the DWCs does not influence the insignificant effect of the sentence type on the prosodic properties of the post-wh part.

limited role of prosody in the post-wh part can be attributed to the occurrence of the lexical cue (the wh-phrase) to the sentence type. As soon as the wh-phrase is uttered, the sentence type is revealed, so the role of prosody in indicating the sentence type is diminished. This is in line with the suggestion that the presence of an interrogative marker (lexical or syntactic) in Persian questions makes the role of prosody in marking questions less crucial (Mahjani, 2003; Sadat Tehrani, 2008, 2011). The difference in the role of prosody in characterizing the pre- and post-wh parts suggests that prosodic correlates of interrogativity are not equally distributed throughout wh-in-situ questions. Rather, they tend to localize in the pre-wh part, where prosodic information is most needed to resolve sentence type ambiguity.

The reported differences in the pitch height between wh-in-situ questions and declaratives agree with what has been reported earlier for yes-no questions without the question word “aya” in Persian as well as for interrogatives in different languages: (i) a higher F0 onset (Benkirane, 1998; Hermann, 1942; Haan, 2001; Iivonen, 1998), (ii) a higher F0 offset (Haan, 2001), (iii) a raised level of pitch register (Bolinger, 1982; Geluykens, 1986; Hermann, 1942; Haan, 2001; Haan et al., 1997; Inkelas & Leben, 1990; Lindsey, 1985; Sadat Tehrani, 2008, 2011), (iv) less declination (Haan, 2001; Haan et al., 1997; Thorsen, 1980; Vaissière, 1983), and (v) a greater excursion size of the pitch accents associated with the pre-wh words (Haan, 2001; Haan et al., 1997; Sadat Tehrani, 2008, 2011). These findings also support the general claim that a greater pitch height in questions can be regarded as a universal feature of languages (Hermann, 1942).

Another acoustic feature that distinguishes Persian wh-in-situ questions from statements is the shorter duration of the pre-wh part and the complete sentence in wh-questions.¹⁷ The decreased duration of questions in comparison with declaratives has already been demonstrated for Mandarin Chinese wh-in-situ questions (Jorrisen, 2014; Yang et al., 2016a September), for yes-no and declarative questions in Dutch, yes-no questions in Manado Malay and Orkney English (Van Heuven & Van Zanten, 2005), and in Hausa (Lindsey, 1985).

Ohala (1984) explained the use of high pitch in questions in ethological terms as a phonologized remnant of animal behavior. The questioner needs to be polite or subservient to his/her respondent when posing a question. In comparison to large (dangerous) creatures, small (harmless) creatures have higher pitches, and make faster movements. Hence, the speaker raises his/her pitch to pretend to be smaller than he/she is physically. Pitch raising by the speaker is interpreted as a sign of submission or politeness to the respondent. Van Heuven and Van Zanten (2005) interpreted faster movements made by small creatures as fast speech rate. They suggested that fast speech rate may express the same message conveyed by high F0 to the respondent. This proposal ties high pitch with fast rate in signaling questions. Previous studies provided evidence for the positive correlation between speech rate and pitch (Rietveld & Gussenhoven, 1987) and the association of high pitch

¹⁷ As to the knowledge of the researchers, no study has been conducted on the prosodic properties of words (constituents) which present given information vs. words (constituents) which present new information in Persian. However, inspection of Taheri Ardali and Xu's (2012) results suggests that information structure does not affect the duration of the words. Therefore, we suggest that the shorter duration of the pre-wh part is not a consequence of the given-information status of the constituents in wh-questions.

(Bolinger, 1964, 1978) and fast rate with questions (Bolinger, 1989). In a speech perception study, Rietveld and Gussenhoven (1987) asked listeners to make judgements about the speech rate of original and pitch manipulated versions of utterances. They found that listeners judged higher-pitched versions of temporally unaltered sentences as being faster. Bolinger (1989), in his dichotomy theory of relaxation and tension, argued that statements and questions are universally characterized by a dichotomy between relaxation (low, falling pitch) and tension (high, rising pitch), respectively. According to Van Heuven and Van Zanten (2005), Bolinger's dichotomy theory might be naturally extended to argue that slow rate or deceleration would be a correlate of relaxation, and that high rate or acceleration is associated with tension.

Further support for the possible association (we are not considering this association as universal, it can be language-specific) of high F0 values and questions is suggested by Bartels (1997) and lately by Gunlogson (2008). The purpose of asking a question is to elicit a verbal response from the respondent and to fill in the questioner's knowledge gap. This purpose implies that the questioner is dependent on the respondent's cooperation to provide him or her with an answer. Bartels (1997) claimed that rising intonation in questions shows dependency whereas falling intonation shows assertiveness of the speaker. Similar to Ohala (1983, 1984) and Bolinger (1978), Gunlogson (2008, p.29) suggested that "rising intonation is the defining property essential for imparting a questioning flavor to a declarative". He suggested that rising intonation implies that the speaker is dependent on the addressee's information and the addressee is the sole source of information. In other words, rising intonation declares that the addressee is the authority and has superiority over the speaker as a source. Therefore, we can propose that questioners use high F0 (possibly in combination with shorter duration) to show their dependence on the respondent. Likewise, respondents use low F0 (possibly in combination with longer duration) to show their authority. This association between acoustic correlates and sentence type may be more prominent in interrogatives where the syntactic or lexical markers of interrogativity are absent or appear later in the sentence. It can possibly explain the dominant role of prosody in the pre-wh part, in contrast with its diminished role in marking the post-wh part in wh-questions. Indeed, appearance of the lexical cue to the sentence type (wh-word) obviates or diminishes reliance on the role of prosody in cueing the sentence type.

Snedeker and Trueswell (2003) proposed that language users possess some implicit knowledge about the relationship between prosody and syntax, and are capable of using this knowledge to guide linguistic choices. Prosodic marking of the pre-wh part in wh-in-situ questions is in line with Snedeker and Trueswell's (2003) proposal. Speakers were capable of using prosody to encode sentence type (wh-questions vs. statements) before the cue (wh-phrase) to the sentence type was uttered. Language users' awareness of the association between prosody and syntax and their capability to use it has also been demonstrated in other studies (e.g. Cooper & Paccia-Cooper, 1980).

According to Rialland (2004), polarity questions in 18 languages (in a database of 80 African languages) are not characterized by high tones and a rising intonation. This suggests that the association between high pitch and questions is not universal. Our study indicates that the linking of high pitch and fast rate (Rietveld &

Gussenhoven, 1987; Van Heuven & Van Zanten, 2005) is applicable to Persian wh-in-situ questions. This result is in line with the suggestion that linking high pitch with interrogativity is the preferred, but not universal, association in the languages of the world (Gussenhoven, 2004).

The present study accounts for the role of prosody in characterizing a cross-linguistically understudied type of interrogative (wh-in-situ questions) in a prosodically understudied language (Persian). The results suggest that the presence of a lexical or syntactic cue to sentence type may affect the role of prosody in encoding different parts of the same sentence. The role of prosody in encoding the part of the sentence which is marked by a lexical or syntactic cue to the sentence type is decreased in comparison to the role of prosody in marking another part of the same sentence which is not marked by a lexical or syntactic cue to the sentence type. This study also provides evidence for the applicability of the hyper- and hypo-theory of speech (Lindblom, 1990) to interrogatives where lexico-syntactic features of interrogativity appear later in the sentence.

Speakers' use of prosody to encode sentence type can help listeners to predict the sentence type before the actual production of the syntactic cue to the sentence type. Prediction of the sentence type can facilitate processing and decoding of sentence type and response preparation on the part of the listener (Grosjean, 1983, 1996). Perception studies are required, therefore, to investigate whether listeners rely on prosody to decode the sentence type before hearing the wh-phrase.

Chapter three

The role of prosody in the identification of Persian sentence types: Declarative or wh-question?¹⁸

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 & Colas, 2006). 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, Caspers, & Schiller, 2018). 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 the wh-phrase type. The results of the experiment corroborate our hypothesis. The outcome is discussed in terms of Ohala's (1984) frequency code, and Bolinger's (1989) claim about the universal dichotomous association between relaxation and declarativity on the one hand and tension and interrogativity on the other hand.

¹⁸ This chapter is based on 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: <https://doi.org/10.1515/lingvan-2016-0085>.

3.1 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 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, Moinszadeh, & Gharaei, 2012; Adli, 2010; Gorjian, Naghizadeh, & Shahramiri, 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdooimian & Ganjavi, 2000; Mirsaeeadi, 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 2).

- (1) a. Mary carries a book.
 b. What does Mary carry?
- (2) a mæɾjæm diruz ketab xærid.
 Maryam yesterday book buy.PAST.3SG.
 “Maryam bought a book yesterday.”
- b mæɾjæm diruz tʃi xærid?
 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 lexical cue to the 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 (Brazil, 1981). This suggests that when a lexico-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.

3.1.1 Background

3.1.1.1 Prosodic correlates of Persian wh-in-situ questions

In a previous study, Shiamizadeh et al. (2018) compared prosodic correlates of wh-in-situ questions and their declarative counterparts. They found that a higher pitch mean, 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. Figures 3.1 and 3.2 present an example of the pitch track of the data.

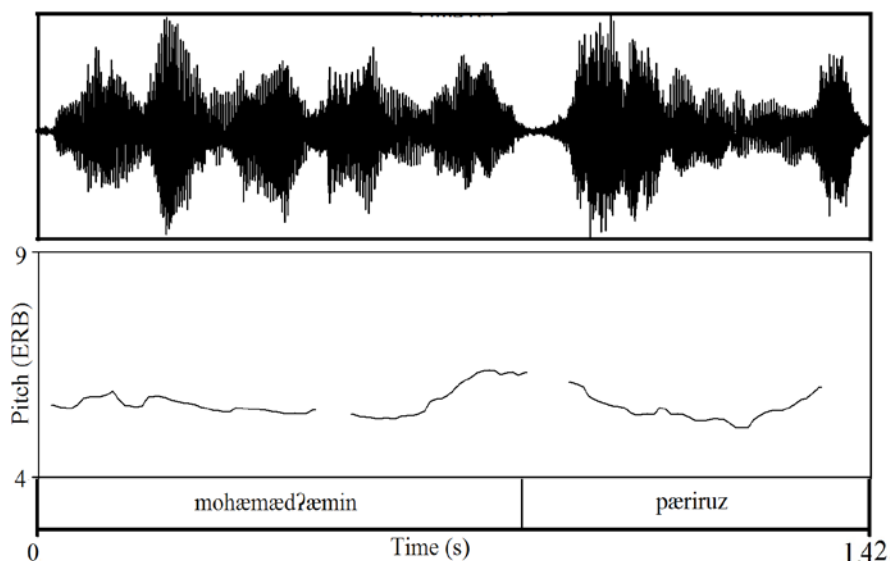


Figure 3.1. The pitch track of the pre-wh part of a declarative sentence. The solid line is the pitch contour. The tier represents the word boundaries. The first word is the subject and the second word is the adverb of time.

3.1.1.2 Theoretical background

Hermann (1942) claimed that the 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.¹⁹ 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/she 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

¹⁹ Polarity 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.

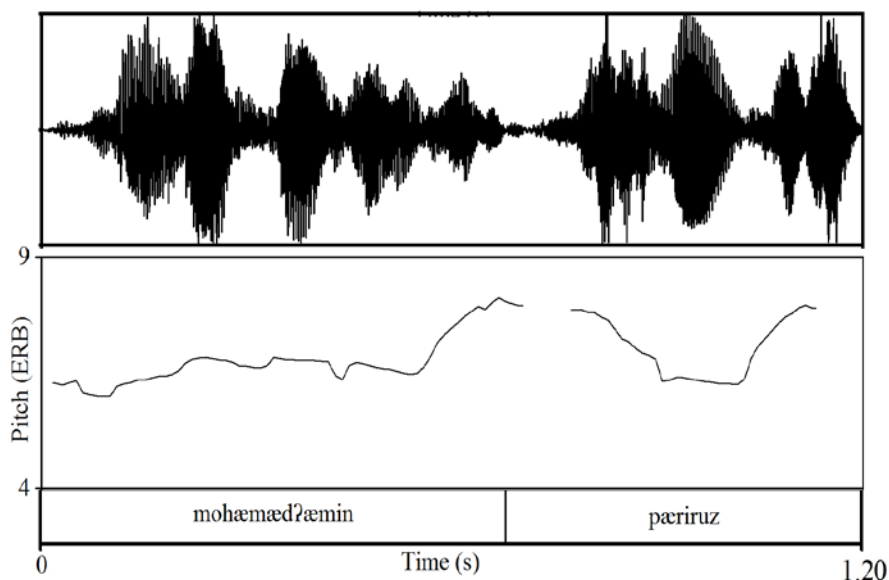


Figure 3.2. The pitch track of the pre-wh part of a wh-question. The solid line is the pitch contour. The tier represents the word boundaries. The first word is the subject and the second word is the adverb of time.

same message conveyed by high F0 to the hearer. It suggests that high pitch is tied with fast rate in signaling questions (Van Heuven & 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, Katz, & Skowronski, 1996; Carlson, Clifton, & Frazier, 2001), Snedeker and Trueswell (2003) proposed that native speakers possess 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.

3.1.1.3 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, Kainada, Lengeris, & Nicolaidis, 2015; Sensui, 1995; Truckenbrodt, Sandalo, & Abaurre, 2009). Plain yes-no questions and statements in Greek are the

same, differing only in intonation (Baltazani et al., 2015).²⁰ A higher level and an earlier alignment of the peak of the pitch accent, steeper slope of rise of the pitch accents, down-stepping of consecutive pitch accents (Baltzanai 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 the sentence type on the part of Spanish native speakers based on the prosodic information available in the stimuli.

Since neither syntax nor morphology regularly marks 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. The final tonic is the last and the main²¹ pitch accent in the intonational phrase. It is called nuclear by Truckenbrodt et al. (2009). Focusing on the nuclear contour, Truckenbrodt et al. (2009) transcribed 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 (Grosjean, 1980) to investigate whether and how prosody guides identification of yes-no and declarative questions as opposed to statements in Spanish, Neapolitan Italian, Northern Standard German, Dutch and French (Face, 2005; Petrone & D'Imperio, 2011; Petrone & Niebuhr, 2014; Van Heuven & Haan, 2000; Vion & Colas, 2006). Gryllia, Yang, Pablos, Doetjes and Cheng (2016 September) and Yang, Gryllia, Pablos, Doetjes and Cheng (2016b September) investigate the role of prosody in the perception of wh-in-situ questions vs. statements in Mandarin Chinese. The remainder of 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 a gating paradigm to investigate whether the acoustic cues of prosody allow for correct perception of the sentence type. The results of his

²⁰ Plain yes-no questions are the questions that lack the question particle “mipos”.

²¹ The main pitch accent in BP falls on the last word of the sentence by default (Gebara, 1976).

experiment showed that native speakers can correctly distinguish declaratives from yes-no questions in 95% of the cases when 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). According to Petrone and D'Imperio (2008), the F0 peak of the NPA is aligned later in questions than in statements (L + H* in questions vs. L* + H in statements). The 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) investigate 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 presentation of the complete sentence.

German questions can be signaled lexically, syntactically and intonationally (Petrone & 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 using the gating paradigm, Petrone and Niebuhr (2014) found that the 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 the sentence-final rise characterize yes-no questions containing more than two stress groups as opposed to 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. The 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 & 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 (almost 90%). The accuracy was almost at a perfect level (100%) when participants were exposed to the final rise.²²

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), wh-in-situ questions in Mandarin Chinese differ from declaratives in terms of F0, duration and intensity. They conducted a gating experiment to investigate whether the clause type (declarative vs. wh-in-situ questions) can be identified based on prosody 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, Gryllia, Doetjes and Cheng (2016a September) reported that Mandarin Chinese wh-in-situ questions in which the wh-phrase is preceded by “dianr” can have an interrogative and non-interrogative interpretation. They found 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 et al. (2016b September) conducted two perception experiments to investigate if prosody cues identification of sentence type (the first experiment) and when the correct sentence type is perceived (the second experiment). In the first perception experiment the complete sentence was presented at once. This experiment showed that prosody guides perception of the intended sentence type at a high level of accuracy (declaratives 95.0% and questions 93.9% correct). The second perception experiment was a gating experiment. 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 above chance level at the first gate, i.e. response accuracy is 59% to declaratives and 54.6% to 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).

The studies reported in this section converge in the conclusion that prosody cues correct perception a) of declarative questions as opposed to declaratives even before the end of the sentence and b) of wh-in-situ questions as opposed to declaratives before the appearance of the wh-phrase. An implication of this conclusion 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.

3.2 Research questions, approach and hypotheses

It has been shown that prosody can cue identification of questions lacking lexico-syntactic features of interrogativity in different languages (c.f. Section 3.1.1.3). A

²² The values of accuracy percentage reported here are based on Figure 9 in Van Heuven & Haan (2000).

few studies investigated the role of prosody in the perception of questions in which lexico-syntactic interrogativity features appear later in the sentence, e.g. wh-in-situ questions (c.f. Section 3.1.1.3). However, it is not known yet if prosody can contribute to identification of wh-in-situ questions in Persian. 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. A possible explanation for this is a higher frequency of occurrence of declaratives than questions in daily conversation (c.f. Van Heuven and Haan (2000) and Vion and Colas (2006)).

According to Ohala's frequency code, Bolinger's dichotomy theory of relaxation and tension, Snedeker and Trueswell's (2003) proposal and based on earlier experimental results (Face, 2005; Gryllia et al, 2016 September; Petrone & D'Imperio, 2011; Petrone & Niebuhr, 2014; Sensui, 1995; Truckenbrodt et al., 2009; Van Heuven & Haan, 2000; Vion & Colas, 2006; Yang et al., 2016b September) we predict that listeners use their implicit knowledge that a high pitch and a fast speech rate represent an interrogative while low pitch and slow speech rate denote a declarative.²³ 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.

3.3 Methodology

3.3.1 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. All of the participants were right handed except two. None of them reported any hearing impairment.

²³ We 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., 2018) 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.

3.3.2 Material

3.3.2.1 Selection of the speakers

Sentences produced in an earlier production experiment (Shiamizadeh et al., 2018) were used as the material 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 selected speakers who are the best representatives of all participants of the production experiment separately for the male and female speakers. A male and a female speaker whose mean value of the acoustic measurements (cf. Section 3.1.1.1) were closest to the mean values in the production of all speakers were chosen.

3.3.2.2 Selection of the stimuli

Twenty pairs of sentences elicited from a male and a female speaker in the production experiment by Shiamizadeh et al. (2018) 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/AdjP	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 chapter. An example of a declarative and a matching wh-question is given in (5a) and (5b).

- (5) a. mohæmædʔæmin pæriruz ʔæsr fɛnq-kærd.
 Mohamadamin two days ago afternoon swim- do.PAST.3SG.
 'Mohamadamin swam in the afternoon two days ago.'
- b. mohæmædʔæmin pæriruz keɟ fɛnq-kærdʔ
 Mohamadamin 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 verbs²⁴ 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 Appendix I. The 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 $\Lambda = .970$, $\eta_p^2 = .030$) and between the words used as the adverb and the sentence type ($F(6,12) = 0.432$, $p > .05$; Wilk's $\Lambda = .968$, $\eta_p^2 = .032$) on the acoustic features described in Section 3.1.1.1 (as dependent variables) 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 constant.

The total number of stimuli in this experiment equals 40 (1 subject \times 1 adverb \times 2 DWCs \times 5 wh-phrases or matching verb \times 2 conditions \times 2 speakers).²⁵ The pre-wh part of the sentences was separated from the remainder of the sentence in Praat version 6.0.04 (Boersma & Weenink, 2014) and was used as the stimuli of the current experiment. Appendix IV presents descriptive statistics of the prosodic correlates (cf. Section 3.1.1.1) of the stimulus material used in this experiment.

3.3.3 Procedure

A forced-choice categorization 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 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 was 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

²⁴ As Appendix I shows, the verb varies along the wh-phrase.

²⁵ Though 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.

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.

3.3.4 Data analysis

The responses and the reaction time (RT) data were transferred from E-prime 2.0.10 (Psychology Software Tools, 2012) 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 & Todorov, 1999). RT was calculated as the time between response and stimulus offset. Two separate repeated measures ANOVAs 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.

3.4 Results

3.4.1 Accuracy

Table 3.1 presents the accuracy of sentence type perception per condition. As Table 3.1 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 = .947, $p > .1$; $\eta_p^2 = .053$), speaker ($F(1,25) = 4.089$, Wilks' Lambda = .859, $p > .05$; $\eta_p^2 = .141$) and wh-phrase type ($F(4,22) = 1.229$, Wilks' Lambda = .817, $p > .1$; $\eta_p^2 = .183$) on response accuracy. The interaction between speakers and sentence type ($F(1,25) = 1.183$, Wilks' Lambda = .955, $p > .1$; $\eta_p^2 = .045$) and between wh-phrase type and sentence type ($F(4,22) = 0.541$, Wilks' Lambda = .910, $p > .1$; $\eta_p^2 = .090$) on response accuracy is not significant, either.

In addition to calculating accuracy percentages, A' score was calculated to correct for a possible response bias (Stanislaw & Todorov, 1999). 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.

Table 3.1. Perception of intended sentence type per condition.

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

3.4.2 Reaction time

RT 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,²⁶ the reaction time value was negative.

The mean RT to declaratives was shorter than the mean RT to questions (see Table 3.2). A RM-ANOVA revealed a main effect of sentence type ($F(1,25) = 14.740$, Wilks' Lambda = .629, $p < .01$; $\eta_p^2 = .371$). In contrast, the effects of speaker ($F(1,25) = 0.036$, Wilks' Lambda = .999, $p > .1$; $\eta_p^2 = .001$), wh-phrase type ($F(4,22) = 1.456$, Wilks' Lambda = .791, $p > .1$; $\eta_p^2 = .209$), the interaction between speaker and sentence type ($F(1,25) = 0.860$, Wilks' Lambda = .967, $p > .1$; $\eta_p^2 = .033$) and the interaction between sentence type and wh-phrase type ($F(4,22) = 1.546$, Wilks' Lambda = .781, $p > .1$; $\eta_p^2 = .219$) were not significant.²⁷

Table 3.2. Mean reaction time (and standard deviation) (in sec) for declaratives and wh-questions.

	Mean (SD)
Declarative	0.282 (0.312)
Wh-question	0.430 (0.330)

3.5 Discussion and conclusion

The current study was designed to address the question 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. As it turns out, wh-questions can be correctly distinguished from

²⁶ 12.9% ($f = 134$) of declaratives and 6.1% ($f = 63$) of wh-questions were responded to before the stimulus offset.

²⁷ To check whether including incorrect decisions (9.5%) in the RT analysis influences the results, a separate RM-ANOVA was conducted with sentence type, speaker and wh-phrase type as independent and only RTs for correct decisions as the dependent variables. Similar to the results reported in Section 3.4.2, the main effect of sentence type ($F(1,18) = 12.780$, Wilks' Lambda = .585, $p < .01$; $\eta_p^2 = .415$) was significant. However, the effects of speaker ($F(1,18) = 0.112$, Wilks' Lambda = .994, $p > .1$; $\eta_p^2 = .006$), wh-phrase type ($F(4,15) = 0.820$, Wilks' Lambda = .821, $p > .1$; $\eta_p^2 = .179$), the interaction between speaker and sentence type ($F(1,18) = 0.365$, Wilks' Lambda = .980, $p > .1$; $\eta_p^2 = .365$) and the interaction between sentence type and wh-phrase type ($F(4,15) = 1.055$, Wilks' Lambda = .780, $p > .1$; $\eta_p^2 = .220$) were not significant. We suggest that including incorrect decisions does not influence the results.

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 the 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 the previous perception experiments on the important role of prosody in the identification of interrogatives as opposed to declaratives (Face, 2005; Gryllia et al., 2016 September; Petrone & D'Imperio, 2011; Petrone & Niebuhr, 2014; Sensui, 1995; Truckenbrodt et al., 2009; Van Heuven & Haan, 2000; Vion & Colas, 2006; Yang et al., 2016b September).

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 (as suggested by Van Heuven & Haan (2000) and Vion & Colas (2006)). Speaker variation and wh-phrase type influence 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's (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 the sentence type and response preparation on the part of the listener (Grosjean, 1983, 1996). From the view of efficient communication, it is logical that sentence mode is designated as early as possible in an utterance (Brazil, 1981; Petrone & 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.

Chapter four

When is a wh-in-situ question identified in

Persian?²⁸

Abstract

Previous literature provides evidence for the influential role of prediction in processing speech (Brazil, 1981; Grosjean, 1983, 1996; Snedeker & Trueswell et al., 2003), as well as for the role of prosody in predicting the eventual syntactic structure of ambiguous sentences (e.g. Snedeker & Trueswell, 2003). Wh-in-situ questions are typical of structures containing temporary syntactic ambiguity. One of the languages characterized by wh-in-situ questions is Persian (e.g. Karimi, 2005). The current research adopted the gating paradigm (Grosjean, 1980) 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. The output was in line with our expectation that correct identification responses were given from the beginning of the sentence. The result is discussed in the context of several proposals regarding the need to integrate prosody and prediction into models of language and speech processing (Beach, 1991; Grosjean, 1983, 1996).

²⁸ This chapter is based on Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (under review). When is a wh-in-situ question identified in Persian? *Language, Cognition and Neuroscience*.

4.1 Introduction

Processing conversational speech is part of language processing. According to Grosjean (1983, 1996), 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, 1996). One source of information in speech processing and prediction of upcoming events is prosody. According to Grosjean (1983, 1996), 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 speech processing (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 utilize 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, 1996; Carlson, Clifton, & Frazier, 2001; Kjelgaard & 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 "You are going shopping?". The syntactic ambiguity is local if the information in the early parts of the sentence does not reveal which of the several possible structures completes the sentence, but the information in a later portion of the sentence assigns only one possible grammatical interpretation to the sentence (Beach, 1991). Wh-in-situ questions typically have local syntactic ambiguity since the syntactic feature relating to the clause type, namely the wh-phrase, occurs later in the sentence. In fronted wh-questions the wh-phrase moves to the beginning of the sentence to form a wh-question (see 1), whereas in wh-in-situ questions the wh-phrase does not move to sentence-initial position (Carnie, 2007; Chomsky, 1977) (see 2). One of the languages which is characterized by wh-in-situ questions is Persian (Abedi, Moinzadeh, & Gharaei, 2012; Adli, 2010; Gorjian, Naghizadeh, & Shahramiri, 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdooimian & Ganjavi, 2000; Mirsaedi, 2006; Toosarvandani, 2008). In Persian, wh-questions are in-situ by default; the wh-phrase

needs not move to the beginning of the sentence, rather it occurs at the same site where its declarative counterpart is expected to occur (see 2b).²⁹

- (1) a. Mary carries a book.
 b. What does Mary carry?
- (2) a. mæɾjæm diruz ketab xæɾid.
 Maryam yesterday book buy.PAST.3SG.
 “Maryam bought a book yesterday.”
- b. mæɾjæm diruz tʃi xæɾid?
 Maryam yesterday what buy.PAST.3SG.
 “What did Maryam buy yesterday?”

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, Enfield, Brown, Englert, Hayashi, Heinemann, Hoymann, Rossano, De Ruiter, Yoon, & Levinson, 2009). Combining the proposal of minimizing 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, & Schiller (2017a) suggest that the prosody of the pre-wh part of a sentence can help predict sentence type in Persian. The result of that perception study raises a new question:

²⁹ The wh-phrase can optionally move to the earlier parts, including the beginning of the sentence (Abedi et al., 2012; Adli, 2010; Gorjian et al., 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdooimian & 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 (2010), Kahnemuyipour (2001), Karimi (2005), Karimi & 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. The declarative and wh-in-situ question counterparts of it are given in (2a) and (2b) within the text.

1. tʃi mæɾjæm diruz xæɾid?
 what Maryam yesterday buy.PAST.3SG.
 “What did Maryam buy yesterday?”

where in the pre-wh part does the relevant distinctive prosodic information become available to feed the process of sentence type prediction?

4.1.1 Background

4.1.1.1 Prosodic correlates of Persian wh-in-situ questions

In a related study, Shiamizadeh, Caspers, & 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 F0 contour and a greater excursion size of the pre-wh words are the two additional features that give rise to the prosodic markedness of the pre-wh part in wh-questions.

4.1.1.2 Empirical background

Gating studies try to determine the amount of acoustic-phonetic information required to identify a stimulus, for example a sentence type (Grosjean, 1996). 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, Gryllia, Pablos, Doetjes & Cheng, 2016b September). The studies by Gryllia et al. (2016 September) and Yang et al. (2016b September) are on wh-in-situ questions and the other studies focused on yes-no questions or declarative questions. 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 recognizable prosodic characteristics, namely a raised F0 peak height in pitch accents and a 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 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 categorization 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 signaled lexically, syntactically and intonationally (Petrone & Niebuhr, 2014). According to Petrone and Niebuhr (2014), questions are not necessarily marked by a H% boundary tone 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 Cristo and Hirst (1993), in French a final F0 rising movement and a sequence of lowered pitches preceding the sentence-final rise characterize 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 in 61% of the cases. 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 & 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 (almost 90%). The accuracy was nearly 100% when the participants were exposed to the final rise.³⁰

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, Gryllia, Doetjes and Cheng (2016a September) 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 et al. (2016b September) conducted two perception experiments to investigate whether prosody cues identification of sentence type (the first experiment) and when the correct sentence type is perceived (the second experiment). In the first perception experiment the complete sentence was presented at once. This experiment showed that prosody enables perception of the intended sentence type at a high level of accuracy (declaratives 95.0% and questions 93.9% correct). In the second perception experiment, which was a gating experiment, only the part of the sentence preceding the wh-phrase was presented. The results showed that listeners can identify the intended sentence type above chance level at the first gate, i.e. response accuracy is 59.0% to declaratives and 54.6% to 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 typical of locally ambiguous sentences. The results of the studies by Face (2005), Gryllia et al. (2016 September), Petrone & D’Imperio (2011), Petrone & Niebuhr (2014), Van Heuven and Haan (2000), Vion and Colas (2006) and Yang et al. (2016b September) 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. This finding implies that prosodic correlates of the pre-wh part in wh-in-situ questions (as locally syntactic ambiguous sentences) could also cue prediction of wh-in-situ questions as opposed to declaratives in Persian.

³⁰ The values of accuracy percentage reported here are based on Figure 9 in Van Heuven & Haan (2000).

4.2 Research questions, approach and hypotheses

4.2.1 Research questions and approach

This study was conducted to answer the following research question: where can Persian native speakers use prosodic correlates to predict wh-in-situ questions before the wh-phrase is made audible? The answer to this question can improve current understanding of how prosody guides syntactic interpretation, in particular temporary syntactic ambiguity resolution. “Fundamental information about this processing mechanism is necessary in order to determine whether and how prosody might be incorporated into a model of spoken sentence processing, and in particular, whether speculation about the (online) use of prosody for relatively immediate, local syntactic disambiguation is worthwhile” (Beach, 1991: 646). Answers to the research question could also contribute to the evaluation of the proposal of integrating prediction into language processing models (Grosjean, 1983; 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, 1996).

To answer the research question, a forced-choice sentence identification task was designed, which also applied the gating method of stimuli presentation. The gating technique was adopted because it allows us to limit the amount of information input by controlling for the temporal presentation of the acoustic signal. This property helps to determine when in the signal the discriminant acoustic information is accessible to feed the process of comparing competitors³¹ 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 signal (Grosjean, 1983, 1996).

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 on a scale from one to five.

4.2.2 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 4.1.1.1). We hypothesize 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

³¹ In this chapter, the competitors are statements and wh-questions in Persian.

onsets characterize 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.

4.3 Methodology

4.3.1 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 the last two years³² to continue their education at the Technology University of Delft. Their age range was between 26 and 40. All of the participants were right handed. None of them reported any hearing impairment.

4.3.2 Material

4.3.2.1 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 4.1.1.1) were used as the material 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 generalizable, we chose speakers who are the best representatives of all participants of the production experiment separately for male and female speakers. 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 4.1.1.1) in the production of all speakers were chosen.

4.3.2.2 The stimuli

4.3.2.2.1 Selection of the Stimuli

Part of the sentences elicited in the production experiment by Shiamizadeh et al. (2018) (see Section 4.1.1.1) 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

³² The data were collected in February 2016.

experiment, (3) and (4) represent the structure of wh-question and declarative stimuli of the current experiment as well.

(3) Subj	Adv	Wh-phrase	Verb
(4) Subj	Adv	ADO/IDO/AdjT/AdjM/AdjP	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). 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 pæriɾuz ʔæsr ʃenq-kærd.
 Mohamadamin two days ago afternoon swim- do.PAST.3SG.
 ‘Mohamadamin swam in the afternoon two days ago.’
- b. mohæmædʔæmin pæriɾuz kej ʃenq-kærdʔ?
 Mohamadamin 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 verbs were used as sentence constituents of the original stimuli in the production experiment. The word constituents of the declaratives and wh-questions are presented in Appendix I.

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 4.4.1, 4.1.1.1 & 4.3.2.3). 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 4.1.1.1). 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 this 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 4.3.2.3. The complete version of each stimulus was played to the participants at the end of the experiment. The complete versions are syntactically unambiguous.

4.3.2.2 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 x 1 adverb x 2 DWCs x 5 wh-phrases and the matching verbs x 2 sentence types x 2 speakers x 8 gates). Although only the pre-wh part of the sentences forms the stimuli of the current experiment, variation in the DWCs, the wh-phrases and their matching verbs are included in the formula to clarify how we arrived at 320 stimuli. The number of wh-questions and their matching declaratives was the same across wh-phrases.

4.3.2.3 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 4.1 and 4.2). 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 4.1 and 4.2 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 chapter.

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

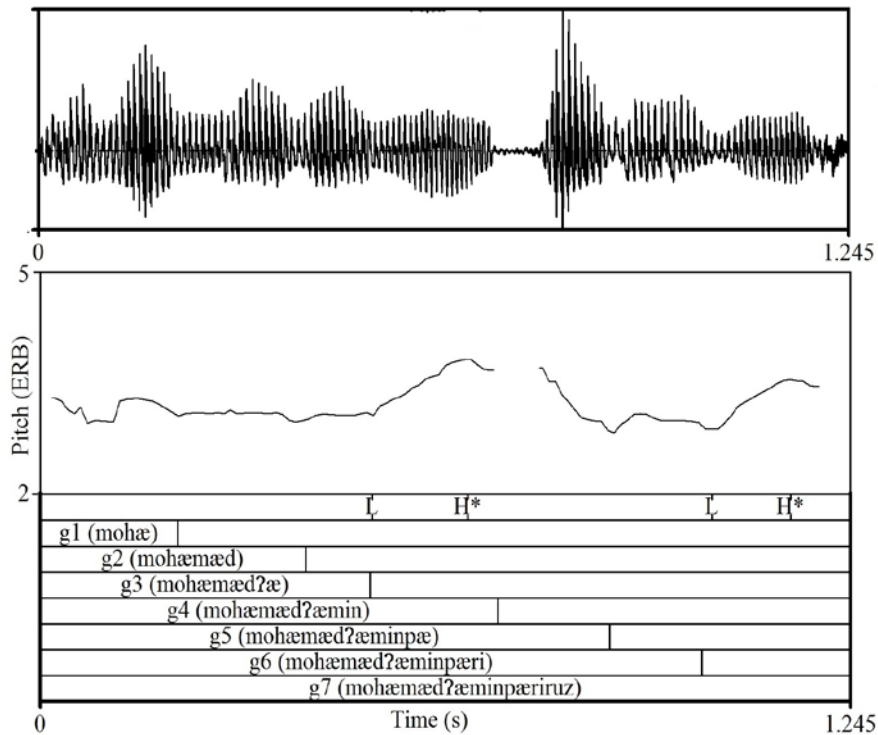


Figure 4.1. The seven gates of a declarative stimulus. The “L” and “H*” represent the valleys and the peaks of the realized pitch accents. The other tiers represent the gate boundaries. The letter g represents the word gate and the number designates the gate number.

The truncation point of the gates corresponds with syllable boundaries. Using Praat, syllable boundaries were indicated manually, then each gate was extracted from the original sound file by running a script. As (6b) and Figures 4.1 and 4.2 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.

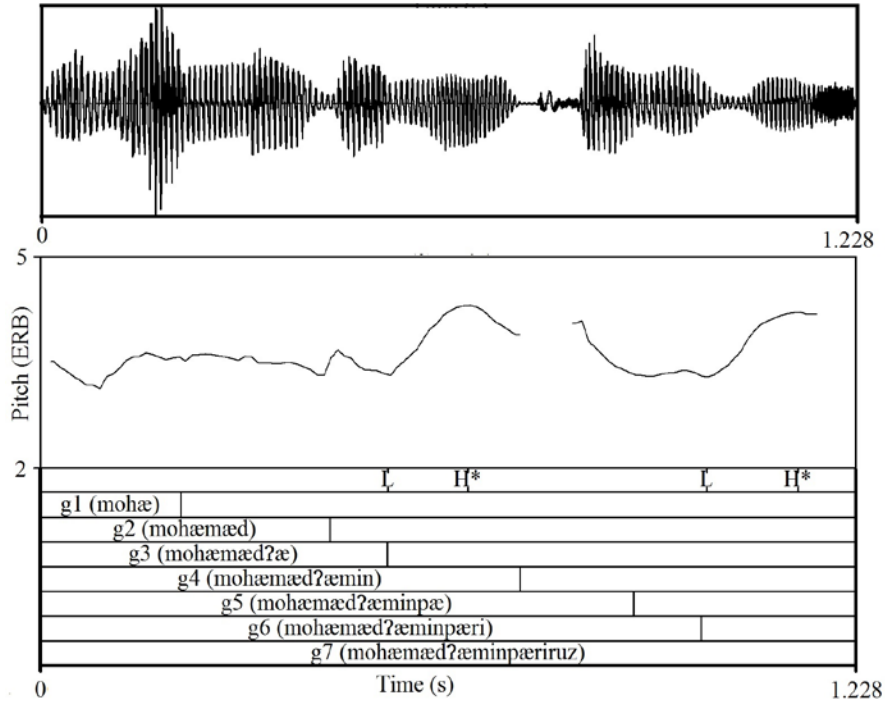


Figure 4.2. The seven gates of a question stimulus. The “L” and “H*” represent the valleys and the peaks of the realized pitch accents. The other tiers represent the gate boundaries. The letter g represents the word gate and the number designates the gate number.

4.3.3 Procedure

A forced-choice sentence categorization 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 familiarized 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 4.3.2.3. The items were a declarative and a question read by one of the speakers from the production task (Shiamizadeh et al., 2018). The stimuli were played to the participants through Sennheiser PC 141 Headset headphones. 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 block. 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 randomized per participant. However, the fifth block was always presented at the end of the experiment to avoid a practice effect on sentence modality identification, as indicated above. The presentation order of the items within all blocks was randomized per participant. The procedure of the main session was identical to that of the practice session. The experiment took about 40 minutes to complete.

4.3.4 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 these RM-ANOVAs were met.

4.4 Results

4.4.1 Response accuracy

Table 4.1 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 is above chance level (75.5%). 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 4.3.

Table 4.1. Perception of intended sentence type across gates and sentence type.

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

Note. CUV = complete unambiguous version of the stimuli; Decl = declaratives; Wh-q = wh-in-situ questions.

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, $\eta_p^2 = .865$) and sentence type ($F(1,19) = 7.577, p < .02$; Wilks' Lambda = .715, $\eta_p^2 = .285$) 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.³³

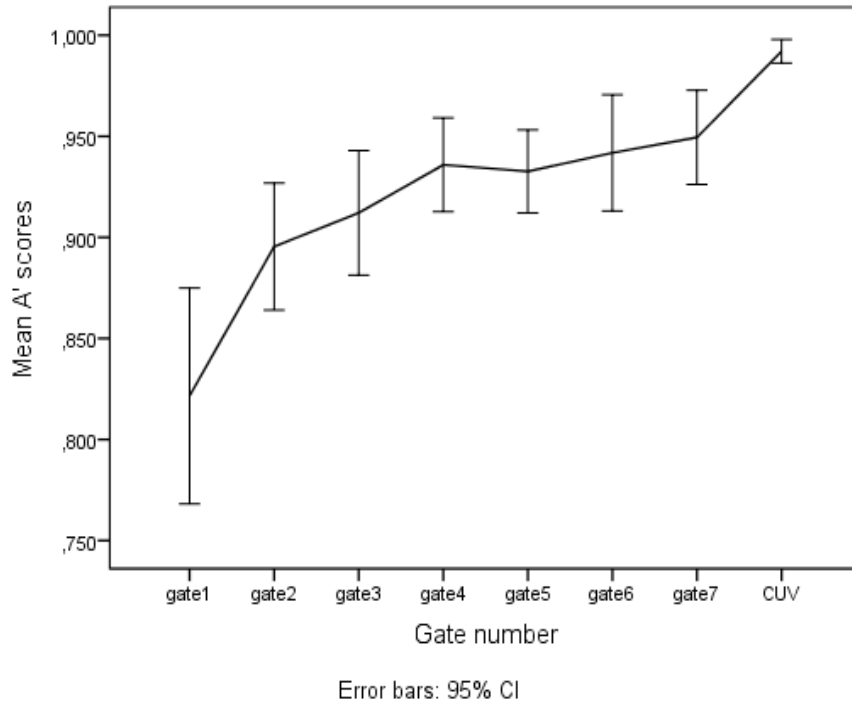


Figure 4.3. Mean A' scores across gates. CUV stands for complete unambiguous version of the stimuli.

Pairwise comparison tests using Bonferroni correction (see Table 4.2) 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 ($p < .01$).³⁴

³³ Response accuracy to declaratives is higher than questions at all gates except at gates 3, 4 and 5.

³⁴ 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 ($p < .01$).

Table 4.2. Results of pairwise comparison tests for response accuracy differences between gates (the result is based on Bonferroni correction test).

Gates	p 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		1.00	1.00	1.00	.001*
Gate 5	.000*	.005*	.607	1.00		1.00	.092	.000*
Gate 6	.000*	.001*	.013*	1.00	1.00		1.00	.030*
Gate 7	.000*	.000*	.006*	1.00	.092	1.00		.036*
CUV	.000*	.000*	.000*	.001*	.000*	.030*	.036*	

Note. CUV = complete unambiguous version of the stimuli.

As we can observe in Figures 4.1 and 4.2, 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 4.1 and 4.2). The pitch accent associated with the subject is presented at this gate. At gate 7, the adverb is entirely presented and the pitch accent realized on it is made audible (see also Figures 4.1 and 4.2). The larger excursion size of the pitch accents realized on the subject and the adverb are the other prosodic features that characterize 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.

According to the result of RM-ANOVA, the difference between the response accuracy to gates 3 and 4, 3 and 5 is not significant. This 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, 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, 3 and 7.

The insignificant increase in response accuracy from gate 4 to gate 7 (2.3%, $p > 0.5$; (see Table 4.2) suggests that the prosodic information until gate 4 provides a strong cue to sentence type identification.

4.4.2 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.³⁵

As Table 4.3 illustrates, the RT to declaratives was shorter than the RT to wh-questions within each gate. According to Figure 4.4, the RT to stimuli decreases as the gate number increases, likely reflecting the increased availability of prosodic information as the gate number increases.

Table 4.3. Mean reaction time (and standard deviation) (in sec) for declaratives and wh-questions across gates.

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

Note. CUV = Complete unambiguous version of the stimuli.

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 ($F(1,19) =$

³⁵ 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.

11.583, $p < .01$; Wilks' Lambda = .621, $\eta_p^2 = .379$), gate ($F(7,13) = 38.080$, $p < .001$; Wilks' Lambda = .047, $\eta_p^2 = .953$) and the interaction of sentence type and gate ($F(7,13) = 4.512$, $p < .01$; Wilks' Lambda = .292, $\eta_p^2 = .708$)³⁶ significantly affected RT.³⁷ 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.

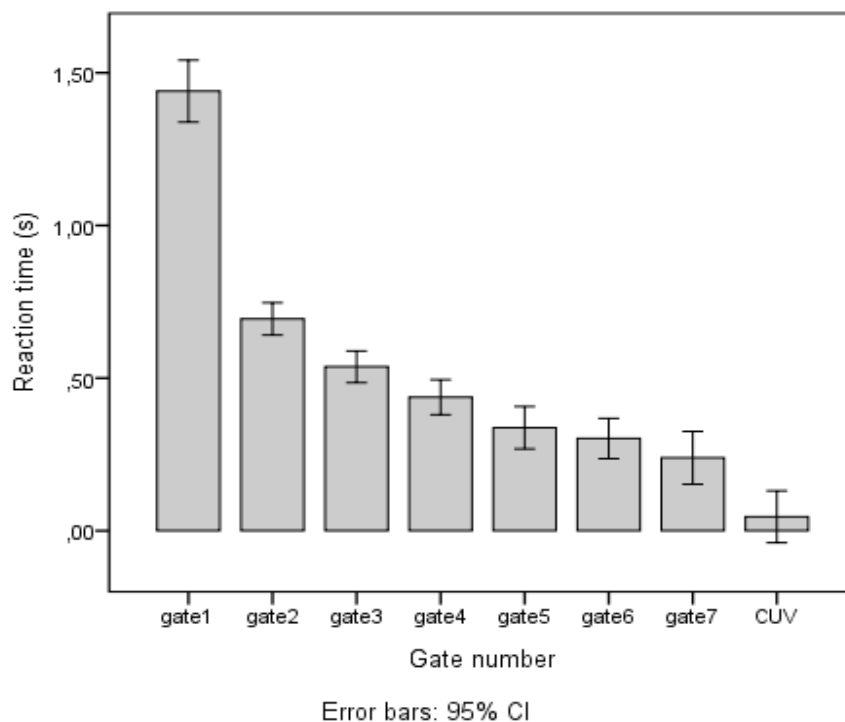


Figure 4.4. Mean reaction time (in seconds) across gates. CUV stands for the complete unambiguous version of the stimuli.

³⁶ At gates 5 and 7, RT to declaratives was significantly shorter than questions.

³⁷ To check whether including incorrect decisions (see Table 4.1) in the RT analysis influences the results, a separate two-way RM-ANOVA was run with sentence type and gate as independent variables and only RT for correct decisions as the dependent variable. Similar to the results reported in Section 4.4.2, effects of sentence type ($F(1,19) = 4.482$, $p < .05$; Wilks' Lambda = .809, $\eta_p^2 = .191$), gate ($F(7,13) = 56.130$, $p < .001$; Wilks' Lambda = .032, $\eta_p^2 = .968$) and the interaction of sentence type and gate ($F(7,13) = 2.915$, $p < .05$; Wilks' Lambda = .389, $\eta_p^2 = .611$) significantly affected RT. We suggest that including incorrect decisions does not influence the results.

The pitch movement on the first two syllables of the adverb “pæriruz” (two days ago) can account for the insignificant decrease in RT from gate 5 to gate 6. Gate 6 presents the pre-wh part of the sentence until the end of the second syllable of the adverb (see Figures 4.1 and 4.2). 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 realized on “-ri” (see Figures 4.1 and 4.2). 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.

4.4.3 Confidence rating

As observable in Figure 4.5, participants’ confidence in their responses increased as the gate number also increased. This is in line with the results regarding response accuracy and the RT to different gates, namely that response accuracy increased and RT decreased as the gate number increased.

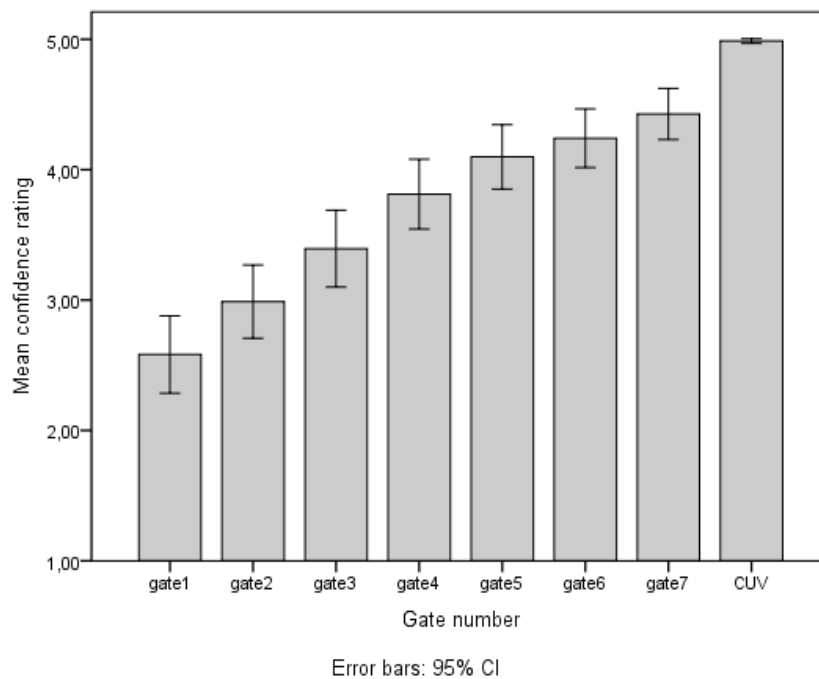


Figure 4.5. Mean confidence rating across gates. CUV stands for complete unambiguous version of the stimuli.

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

The response accuracy to declaratives is higher (see Table 4.1) and the RT to declaratives is shorter (see Table 4.3) in comparison to questions at all gates (though not always significantly). However, the confidence rating to declarative is higher than the confidence rating to questions only at gates 1 and 8 (Figure 4.6).³⁸

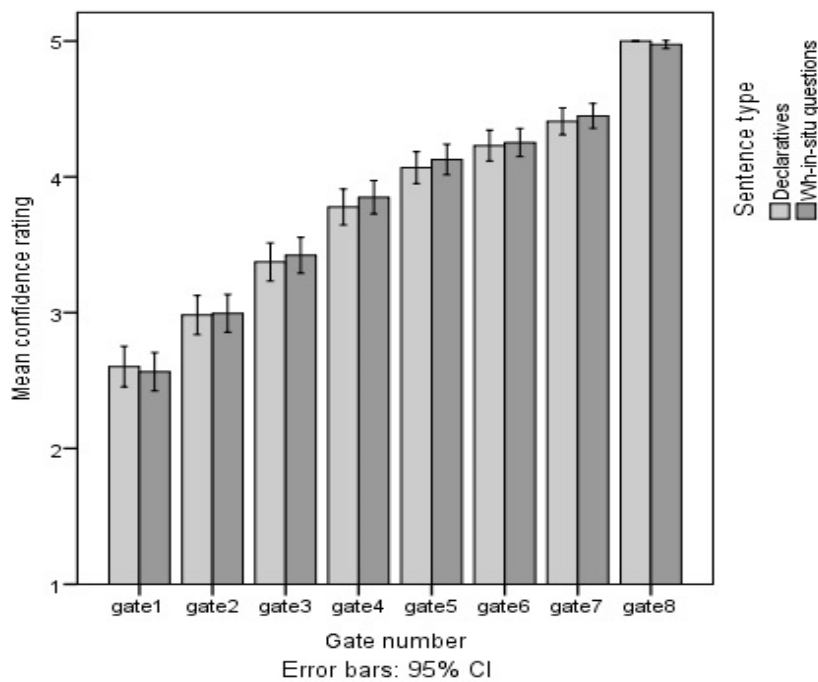


Figure 4.6. Mean confidence rating across sentence type and gates. CUV stands for complete unambiguous version of the stimuli.

³⁸ As mentioned earlier, at gate 8 the complete unambiguous stimulus is presented. The confidence rating to all declaratives ($N = 400$) and to 395 (out of 400) questions is 5 at this gate. An inspection of the five question stimuli which were not scored 5 on the confidence rating scale suggests that a) four items are wrongly replied to, b) the four items with a wrong reply are not scored 5 or have a missing confidence rating because listeners understood that they pressed the wrong response button and c) one of the listeners missed giving a confidence rating to an item which was correctly identified.

At gates 2, 3, 4, 5, 6 and 7 questions have a higher confidence rating than declaratives. The stimuli presented from gate 1 to 7 are syntactically ambiguous and the participants have to rely on prosody to decide on the sentence type. The prosodic information at gate 1 is limited. At this gate, listeners have a tendency to give more declarative responses than questions responses (the difference between response accuracy to questions and statements at this gate is 11.5%, cf. Table 4.1). Therefore, when giving a declarative response they are more confident than when giving a question response. From gates 2 to 7, the amount of discriminating prosodic information is increasing. Though the tendency to give more declarative responses is decreased, listeners still have the inclination to give a declarative rather than a question response. This means that listeners do not give a question response unless there is compelling evidence that a stimulus is a question. Since listeners give a question response after hearing compelling evidence, they are more confident when giving question responses than when giving a declarative response.

4.5 Discussion and conclusion

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.

The first pre-wh word on which the first pitch accent of the sentence is realized was presented at gate 4 in our stimuli. 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 the last gate (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. Language processing models may need to account for this possibility (Grosjean, 1983). 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; Vion & Colas, 2006), declaratives also have shorter reaction times in comparison to questions. A possible reason for the decreased RT and the higher response accuracy to declaratives could be easier identification of declaratives. Easier identification of declaratives might stem from

the higher frequency of declaratives in comparison to questions in daily conversation (as suggested by Van Heuven & Haan (2000) and Vion & Colas (2006)).

The general result of this research corroborates several proposals suggested in the literature. First, prosody plays a prominent role in processing syntactically ambiguous sentences (e.g. Beach, 1991; Beach et al., 1996; Carlson et al., 2001; Kjelgaard & Speer, 1999; Nagel et al., 1994; Snedeker & Trueswell, 2003; Warren et al., 1995). 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 and the role of prosody in prediction may need to be incorporated into models of language processing (Grosjean, 1983). Fourth, prediction can be reset as more prosodic information is provided to the listener and language-processing models may also need to account for this (Grosjean, 1983). Finally, models of spoken sentence processing may need to integrate the (online) use of prosody in interpreting constructions which have temporary syntactic ambiguity (Beach, 1991). There are, however, also 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. EEG experiments with matching and mismatching syntax-prosody could help to clarify how prosody is utilized in the online processing of language.

Chapter five

The role of F0 and duration in the identification of wh-in-situ questions in Persian³⁹

Abstract

Prosody plays an influential role in the recognition of Persian wh-in-situ questions (Shiamizadeh, Caspers, & Schiller, 2017a). Perception of sentence modality contrast is supported by several prosodic correlates (e.g. Cangemi & D’Imperio, 2013; Gryllia, Yang, Pablos, Doetjes & Cheng, 2016 September; Peng, Chatterjee, & Lu, 2012). Perception studies demonstrate that prosodic correlates do not contribute equally to the identification of sentence modality contrast, and F0 contributes more prominently to the categorization of this contrast (e.g. Cangemi & D’Imperio, 2013; Peng et al., 2012). The current study manipulates a number of different prosodic correlates to investigate their relative contribution to sentence type categorization in Persian. For this purpose, a perception experiment was designed in which native Persian speakers were presented with manipulated stimuli and decided if they perceived a wh-question or a declarative sentence. Based on the literature, we expect a primary role for F0 and a secondary role for duration in the perception of sentence modality contrasts. We also predict improved identification of sentence type in matching cue conditions in comparison to mismatching cue conditions. The results of the study confirm our predictions, i.e. a primary role of F0, and an increased identification in the matching cue condition is applicable to the perception of the contrast between Persian wh-in-situ questions and declaratives.

¹ This chapter is based on 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.

5.1 Introduction

One linguistic function that is frequently used in daily conversation is asking a question. Wh-questions are interrogative expressions that use a wh-phrase to enquire about desired information. Wh-questions 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 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 that is characterized by wh-in-situ is Persian (Abedi, Moinzadeh, & Gharaei, 2012; Adli, 2010; Gorjian, Naghizadeh, & Shahramiri, 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdoomian & Ganjavi, 2000; Mirsaedi, 2006; Toosarvandani, 2008). In Persian, wh-questions are in-situ by default. The wh-phrase does not move to the beginning of the sentence, rather it occurs at the same site where its declarative counterpart is expected to occur (see 2b).⁴⁰

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

- (2) a. mæɾjæm diruz ketab xærid.
Maryam yesterday book buy.PAST.3SG.
“Maryam bought a book yesterday.”
- b. mæɾjæm diruz tʃi xærid?
Maryam yesterday what buy.PAST.3SG.
“What did Maryam buy yesterday?”

Prosody plays an influential role in production (e.g. Baltazani, Kainada, Lengeris, & Nicolaidis, 2015; D’Imperio, 2000; Gryllia, Cheng & Doetjes, 2016; Haan, 2001; Petrone & D’Imperio, 2008; Petrone & Niebuhr, 2014; Shiamizadeh, Caspers & Schiller, 2018; Van Heuven & Van Zanten, 2005; Vion & Colas, 2006; Yang,

⁴⁰ The wh-phrase can optionally move to the earlier parts, including the beginning of the sentence (Abedi et al., 2012; Adli, 2010; Gorjian et al., 2012; Kahnemuyipour, 2009; Karimi, 2005; Karimi & Taleghani, 2007; Lotfi, 2003; Megerdoomian & 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 (2010), Kahnemuyipour (2001), Karimi (2005), Karimi & 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. The declarative and wh-in-situ question counterparts of it are given in (2a) and (2b) within the text.

1. tʃi mæɾjæm diruz xærid?
what Maryam yesterday buy.PAST.3SG.
“What did Maryam buy yesterday?”

Gryllia, Doetjes & Cheng, 2016a September) and perception (Baltazani et al., 2015; Cangemi & D'Imperio, 2013; Gryllia et al., 2016 September; Petrone & D'Imperio, 2011; Petrone & Niebuhr, 2014; Shiamizadeh, Caspers & Schiller, 2017a; Truckenbrodt, Sandalo & Abaurre, 2009; Vion & Colas, 2006; Yang, Gryllia, Pablos, Doetjes & Cheng, 2016b September) of interrogatives. It is reported that the perception of the prosodic contrast between interrogatives and declaratives is mainly achieved by relying on F0 as the primary cue (Cangemi & D'Imperio, 2013; Peng et al., 2012), while duration is considered the secondary cue (Cangemi & D'Imperio, 2013).

Our previous studies demonstrate that in the production (Shiamizadeh et al., 2018) and perception (Shiamizadeh et al., 2017a) of the pre-wh part in wh-in-situ questions vs. declaratives in Persian, prosodic correlates play an important role in marking the sentence modality contrast. The result of the studies by Shiamizadeh et al. (2018, 2017a) brings the focus to the relative contribution of F0 and duration in the perception of wh-in-situ questions as opposed to statements in Persian. A perception experiment in which relevant prosodic correlates are manipulated can investigate the relative contribution of F0 and duration to sentence modality contrast identification in Persian.

5.1.1 Background

5.1.1.1 Production and perception of prosodic correlates of Persian wh-in-situ questions

To our knowledge, Shiamizadeh et al. (2018, 2017a) is the first systematic and experimental study investigating the role of acoustic correlates of prosody in the production and perception of wh-in-situ questions vs. declaratives in Persian. In a production study, Shiamizadeh et al. (2018) compared the prosodic correlates of wh-in-situ questions and their declarative counterparts to investigate 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. They found that a higher level of pitch mean, a higher F0 onset and a shorter duration contributed to the prosodic distinction of the pre-wh part in wh-questions as opposed to 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.

The production study (Shiamizadeh et al., 2018) was followed by a perception study (Shiamizadeh et al., 2017a), the purpose of which was to investigate whether the prosody of the pre-wh part can cue the identification of wh-in-situ questions even before the disambiguating wh-word is heard midway through the sentence. The results revealed that prosody indeed cues perception of wh-in-situ questions before the wh-phrase is uttered by the speaker.

5.1.1.2 Empirical background

Tapping into the relative contribution of melodic cues and durational cues for the identification of interrogatives vs. declaratives is a relatively untouched territory. To our knowledge, Cangemi and D'Imperio (2013) is the only study investigating the relative contribution of duration in the perception of declaratives as opposed to interrogatives. According to the production experiment by Cangemi and D'Imperio (2013), the contrast between statements and questions in the Neapolitan regional variety of Italian is intonationally specified by different tunes and different durational patterns. Durational differences between these two sentence modalities are localized at specific portions of the sentence: the first sound segment is longer in statements, while the last sound segment is longer in questions. In the same study, Cangemi and D'Imperio (2013) investigated the relevance of the durational differences between statements and questions in the perception of the correct sentence type. In their perception study, Cangemi and D'Imperio (2013) used resynthesized sentences with three different F0 contours: question F0 contour, ambiguous F0 contour and statement F0 contour. Then, they implemented three different durational patterns (question duration, ambiguous duration and statement duration) on stimuli in each group of F0 contour. Therefore, three different durational patterns were transplanted on a stimulus with the same F0 contour. For example, there were three versions of a stimulus with a question F0 contour: a) with question duration, b) with ambiguous duration, c) with statement duration. They predicted that if duration influences sentence type identification, items with different durational patterns but the same F0 contour would receive different responses. The result of the perception experiment showed that a) F0 cues are relevant in the perception of sentence modality contrasts, whereas durational cues do not affect identification of the correct sentence type, and b) the reaction time (RT) to the stimuli with ambiguous F0 is longer than the RT to other conditions.

This result implies that, in comparison to durational cues, F0 cues contribute more to the identification of sentence modality contrast. Furthermore, the cooperation of F0 and durational cues has a stronger effect on the perception of correct sentence type than conflicting F0 and duration. Augmented identification of speech contrast in the cooperating cue condition in comparison to the conflicting cue condition has been reported in other studies which investigated the role of acoustic cues to the identification of minimal pairs (e.g. Fry, 1955; Hazan & Rosen, 1991).

5.2 Research questions, approach and hypotheses

5.2.1 Research questions and approach

To investigate the relative contribution of F0 and durational cues, F0 contour and durational pattern must be manipulated. This study addresses the following questions:

- 1) Does F0 or duration or their cooperation contribute more to the distinction of wh-in-situ questions from declaratives in Persian?

2) Is there a difference between RT to the congruent and incongruent cue conditions?

The answer to the first research question potentially sheds light on the applicability of two proposals to sentence type identification in Persian: a) the primary role of F0 (Cangemi & D'Imperio, 2013; Peng et al., 2012) and the secondary role of duration in the recognition of prosodic contrast (Cangemi & D'Imperio, 2013), b) increased categorization of contrasts in the congruent cue condition in comparison to the conflicting cue condition (Fry, 1955; Hazan & Rosen, 1991).

To answer the two research questions above, F0 and durational patterns of the pre-wh part of declaratives and their matching wh-questions were manipulated. Then, a forced-choice identification task was designed in which Persian native speakers listened to the manipulated pre-wh part of wh-in-situ questions and their matching declaratives. They were required to decide whether what they heard was going to be a wh-question or a declarative. The participants were asked to opt for either a declarative or a question by pressing a key as quickly as possible. Their reaction times were recorded.

5.2.2 Hypotheses

The primary role of F0 in the perception of prosodic contrast is reported in the literature (Cangemi & D'Imperio, 2013; Peng, et al., 2012). Other studies demonstrated an increased effect of congruent cue conditions on the identification of speech contrasts (Fry, 1955; Hazan & Rosen, 1991) in comparison to the conflicting cue conditions. The above-mentioned findings can be logically extended to the perception of wh-in-situ question in Persian. Therefore, we hypothesize that:

- a) F0 is the dominant contributor to the categorization of sentence modality (in-situ wh-questions versus statements) in Persian. In other words, listeners will draw on F0 cues even in duration-manipulated items to decide on the sentence type. This means F0 manipulated items will receive more correct (expected) responses⁴¹ than duration manipulated items.
- b) Identification of the sentence type in congruent cue conditions (stimuli where both F0 and duration are manipulated) is augmented in comparison to conflicting cue conditions (stimuli in which either F0 or duration is manipulated). Therefore, more identification responses will be given to the items with congruent F0 and duration in comparison to the items with conflicting F0 and duration.
- c) The RT to congruent cue conditions is shorter than the RT to conflicting cue conditions.

⁴¹ Expected and correct responses are defined in Section 5.3.4.

5.3 Methodology

5.3.1 Participants

Twenty-four native speakers of Persian, twelve males and twelve females, took part in this experiment. All of the participants were brought up in Tehran. Fifteen participants lived in Delft. They came to the Netherlands in 2014 or 2015⁴² to continue their education at the Technical University of Delft. Nine participants were students at the University of Qom, Iran. Their age range was between 26 and 40. All of the participants were right handed except one. None of them reported any hearing impairment.

5.3.2 Material

5.3.2.1 Speaker selection

Part of the sentences elicited from the native Persian speakers in the production experiment on the prosodic correlates of Persian wh-in-situ questions (Shiamizadeh et al., 2018) were used as the basic material for the current experiment. To control for the effect of gender on the listeners' performance in the perception experiment, we chose 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 broaden the generalizability of the results of the current experiment, we picked out those speakers who are most representative of all participants of the production experiment separately for the male and female speakers. A male and a female speaker whose mean value of the acoustic measurements was closest to the mean value of the acoustic measurements in the production of all speakers (cf. Section 5.1.1.1) were selected.

5.3.2.2 Selection of the stimuli

Twenty pairs of sentences elicited from a male and a female speaker in the production experiment by Shiamizadeh et al. (2018) 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/AdjP	Verb

⁴² The data were collected in October, 2015.

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 chapter. An example of a declarative and a matching wh-question is given in (5a) and (5b).

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

Five different wh-phrases, two different nouns as the subject, two words as the adverb, two words in each category of DWC and five verbs were used as sentence constituents of the original stimuli in the production experiment. The word constituents of the declaratives and wh-questions are presented in Appendix I. As Appendix I shows, the verb varies along the wh-phrase.

A repeated measures multivariate analysis of variance 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 described in Section 5.1.1.1 (dependent variables) were 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 pre-wh part of sentences was separated from the remaining part of the sentence in Praat version 6.0.04 (Boersma & Weenink, 2014) and were used as the stimuli of the current experiment. To prepare the stimuli for this experiment, acoustic correlates of the prosody of the pre-wh part must be manipulated. The manipulation conditions are F0 manipulation (F0M), duration manipulation (DM) and F0 and duration manipulation (F0DM) and resynthesized version of original sentences (RSO).

The total number of stimuli equals $160 (1 \text{ subject} \times 1 \text{ adverb} \times 2 \text{ DWCs}^{43} \times 5 \text{ wh-phrases and matching verbs} \times 2 \text{ sentence types} \times 4 \text{ manipulation conditions} \times$

⁴³ As mentioned earlier, there are five categories of DWCs which correspond with the number of wh-phrases. DWCs replace the wh-phrases in declaratives. Therefore, the use of the categories of the DWCs in declaratives varies along with the use of the wh-phrases in wh-questions, i.e. if a wh-phrase asking about ADO is used in the wh-question, a category of the DWC referring to ADO is used in the declarative. This means that the coefficient of wh-phrases in the formula accounts for the number of categories of DWCs as well. However, since two words are used in each category of DWC, the coefficient of 2 is defined for DWC in the formula.

2 speakers). Although only the pre-wh parts of the sentences comprise 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 160 stimuli.

Sections 5.3.2.3 to 5.3.2.7 explain the manipulation process. In these sections, the stimuli with manipulated acoustic features are called target sound (TS) and the stimuli from which the acoustic features are transplanted to the target sound are called the source sound (SS). For example, if the F0 contour of a declarative is replaced by the F0 contour of a question, the declarative is the TS and the question is the SS. It must be noted that all steps of the manipulation process were automatically accomplished by running scripts in Praat.

5.3.2.3 Stylizing the pitch contour

The preliminary step in manipulating F0 is stylizing the pitch contour. Before stylizing the pitch contour, the erroneous F0 values e.g. octave errors were corrected in the original F0 contour using Praat functions. Then, the original pitch contour of the utterance was stylized using a one-semitone frequency resolution in Praat. Thus, two F0 contours were available, the original or non-stylized pitch contour (the second panel in Figures 5.1 and 5.2) and the stylized pitch contour (the third panel in Figures 5.1 and 5.2).

The stylized contour produces points which can be manipulated. All pitch points in the stylized pitch contour of the pre-wh part of the sentence were removed except for 6 points: the F0 onset, the valley (L) and peak (H*) of the subject pitch accent, the L and H* of the adverb pitch accent and the end of the fall of the subject pitch accent (ESPA) (see Figures 5.1 & 5.2). The F0 onset and the L and H* of the subject and adverb pitch accents are kept because these are the F0 points that make the prosody of wh-in-situ questions distinct from that of declaratives (Shiamizadeh et al., 2018). ESPA was also kept so that no audible difference existed between the original and the stylized pitch curves. The end of the fall of the adverb pitch accent is not considered here because the relevant pitch point is not produced by the stylized pitch. In addition, adding a pitch point manually as the end of the fall of the adverb pitch accent does not affect the audible similarity between the original and the stylized pitch curves. The pitch points were interconnected with straight lines.

After stylizing the pitch contour, two steps were taken to accurately determine the location and the F0 value of the valleys and the peaks of subject and adverb pitch accents. At the first step, L and H* were indicated in the stylized pitch. At the second step, the valley and the peak points (L and H*) produced in the stylized pitch contour were automatically moved to the maximum and minimum F0 of the original F0 contour using Praat functions (this method was adopted from Sadat Tehrani, 2009). The place and the value of the minimum and maximum F0 of the original F0 contour determined by Praat functions (the second step) either coincide with or closely correspond to the L and H* produced by the stylized pitch (the first step). There were no F0 plateaus.

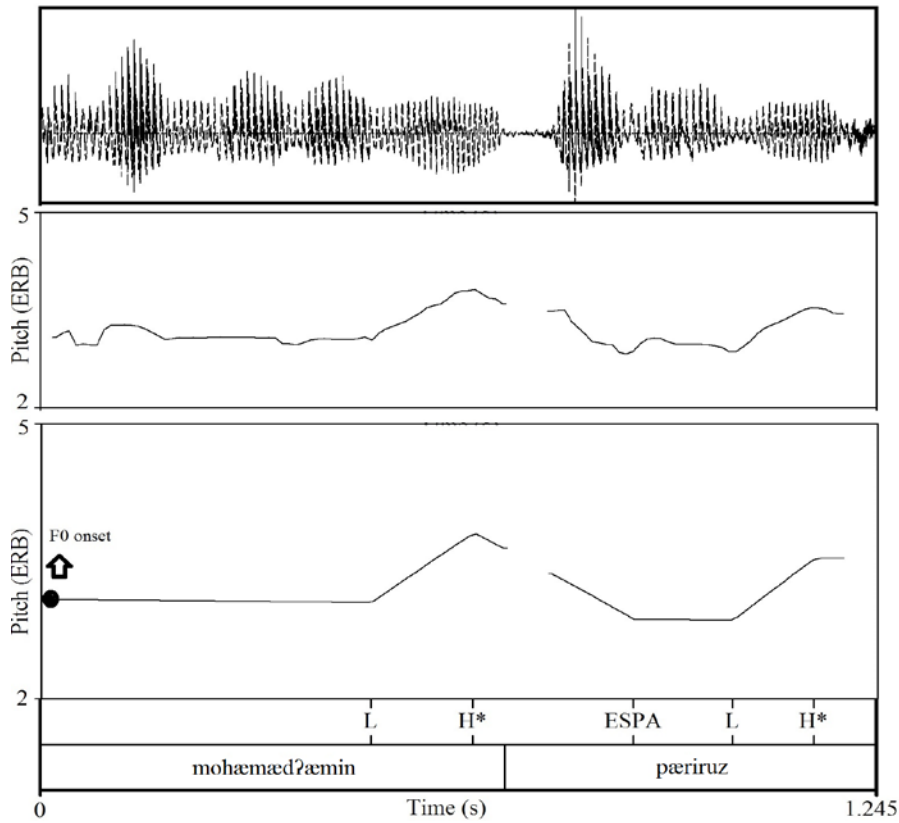


Figure 5.1. The waveform, original (non-stylized) pitch contour and stylized pitch contour of the pre-wh part of a declarative sentence. The first panel presents the waveform, the second panel displays the non-stylized pitch contour and the last panel contains the stylized pitch contour. In the stylized contour, only the points designating F0 onset, L and H* and the end of the fall of the subject pitch accent (ESPA) are kept. In the first tier of the third panel, the “L” and “H*” which correspond with the word “mohæmædʔæmin” in the second tier represent the valley and the peak of the subject pitch accent, ESPA is the end of the fall of the subject pitch accent and the “L” and “H*” which correspond with the word “pæriruz” in the second tier demonstrate the valley and the peak of the adverb pitch accent. The second tier in the third panel represents word boundaries.

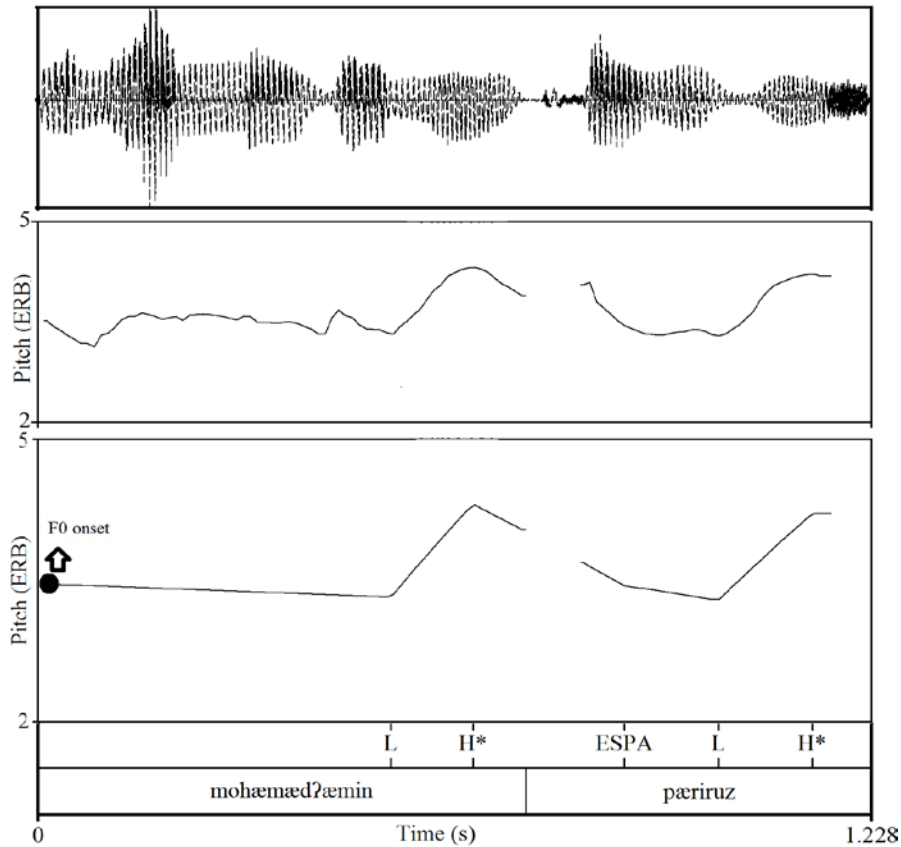


Figure 5.2. The waveform, original (non-stylized) pitch contour and stylized pitch contour of the pre-wh part of a wh-question. The first panel presents the waveform, the second panel displays the non-stylized pitch contour and the last panel contains the stylized pitch contour. In the stylized contour, only the points designating F0 onset, F0 onset, L and H* and the end of the fall of the subject pitch accent (ESPA) are kept. In the first tier of the third panel, the “L” and “H*” which correspond with the word “mohæmædʔæmin” in the second tier represent the valley and the peak of the subject pitch accent, ESPA is the end of the fall of the subject pitch accent and the “L” and “H*” which correspond with the word “pæriruz” in the second tier demonstrate the valley and the peak of the adverb pitch accent. The second tier in the third panel represents word boundaries.

5.3.2.4 F0 manipulation

To investigate the role of F0 in the identification of wh-questions, the F0 contour was manipulated. After stylizing the pitch contour, F0 manipulation was performed. The pitch points which were kept during the stylization process were manipulated: the F0 onset, the L and H* of the subject and adverb pitch accents, and the end of

the fall of the subject pitch accent. These pitch points in the stylized pitch of the TS were manipulated to create the intonational characteristics of their corresponding pitch points in the stylized pitch of the SS.

In the F0 manipulation process, F0 values were manipulated, but the time coordinates and duration were left unaltered. Therefore, in a question where F0 was manipulated, the question keeps its original durational pattern but its F0 contour is changed to the F0 pattern of its declarative counterpart. The correct (expected) response to this item is “declarative”. Similarly, in a declarative where F0 was manipulated, the declarative keeps its original durational pattern but its F0 contour is changed to the F0 pattern of its corresponding question. The correct (expected) response to this item is “question”.

5.3.2.5 Duration manipulation

Duration was manipulated to inspect the role of duration in the identification of declaratives as opposed to wh-questions. Since the pitch contour was stylized in the F0 manipulated stimuli, the pitch contour in duration manipulated items has to be stylized as well. The pitch was stylized in the same manner described in Section 5.3.2.3.

To manipulate duration, the duration tier of the SS must be copied to the TS. Therefore, duration of the duration tier of the TS was multiplied by the quotient of the duration of the SS divided by the duration of the TS.

In duration manipulated items, only the duration was manipulated and the F0 was left unaltered. Therefore, in a question where the duration was manipulated, the question keeps its original F0 contour but its duration is switched to the durational pattern of its declarative counterpart.⁴⁴ The correct (expected) response to this item is “declarative”. Similarly, in a declarative where duration was manipulated, the declarative keeps its original F0 contour, but its duration is changed to the duration of its corresponding question. The correct (expected) response to this item is “question”.

5.3.2.6 F0 and duration manipulation

Manipulating the duration tier of F0 manipulated items provides us with items in which both F0 and duration are manipulated. The duration was manipulated as explained in the Section 5.3.2.5. For instance, a statement gets both the F0 contour and durational pattern of its corresponding question and a question receives the F0 contour and durational pattern of its corresponding statement. The alignment of the valleys and peaks of the pitch accents were not altered. The correct (expected) answer to a question which gets the F0 and durational properties of its corresponding declarative is “declarative”. And the correct (expected) answer to a

⁴⁴ Though our purpose was to keep the F0 contour unaltered in the DM items, changing the duration inevitably influences the steepness of the rises and falls.

declarative to which the F0 and duration of its corresponding question is transplanted is “question”.

5.3.2.7 Resynthesized original sounds

To manipulate the stimuli of this experiment, the pitch contour of each sound was stylized as a first step and the resynthesized version of the sound was extracted from the manipulation object of each sound as the last step. Therefore, the resynthesized version of each sound with a stylized pitch contour is required as the baseline condition of the current experiment. Pitch stylization procedure is explained in Section 5.3.2.3. After stylizing the pitch contour of each sound, the resynthesized version of the sound was automatically extracted from manipulation object of the sound file in Praat. Resynthesized stimuli are not manipulated. Response accuracy of resynthesized stimuli and the reaction time to these items form the baseline for the conclusions of this study.

5.3.3 Procedure

A forced-choice sentence categorization task was designed in E-prime 2.0.10 (Psychology Software Tools, 2012). The subjects were seated in front of a computer screen in a quiet room. The experiment started with the presentation of written instructions on the computer screen. Participants could take as much time as they wanted to read the instructions. They were allowed to ask questions about the instructions if they had any. Then, they were introduced to the task by performing a practice session. The practice session included 16 non-experimental items (1 item \times 2 sentence types \times 4 manipulation conditions \times 2 speakers). The pre-wh part of one declarative and one wh-question from each manipulation condition read by two different speakers were played to them over headphones (Sennheiser PC 141 Headset). Participants were instructed to decide whether what they heard was 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 the “M” key or “Z” key on the keyboard. If participants did not give a response within four seconds, the experiment automatically proceeded to the next trial. To help participants not confuse which key they needed to press for which option (declarative or wh-question), a full stop and a question mark along with the letters M and Z appeared on two opposite sides of the screen, at the same time as a stimulus was played to them. Three seconds after pressing the key, the next stimulus was played. Three seconds were entered as the inter-item interval (ITI). During the ITI the Persian equivalent of the word “next item” was shown on the screen as an indication that the next stimulus was coming. The presentation order of the items of the practice session was the same for all participants. They 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.

The stimuli were divided into four blocks. Each block contained 40 items. The participants took at least a three-minute break between blocks. After the break,

they were asked to press the space bar to continue to the next block when they felt ready to start again. Every block started with a warming-up which consisted of two items other than the main items of the block. The purpose of including warming-up items was to make participants ready for the new block after the break. The sequence in which the blocks and the sentences within each block were presented was randomized for each participant. The order in which the full stop and the question mark and the corresponding letters (M or Z) was displayed on the screen was fixed for individual participants, whereas it was counterbalanced for different subjects. The procedure of the main session was identical to that of the practice session. The experiment lasted about 35 minutes for each participant.

5.3.4 Data analysis

As mentioned previously, the F0 contour and durational pattern were manipulated to investigate their relative contribution to sentence type categorization. Therefore, an expected response was defined for each stimulus based on the manipulated prosodic feature in that stimulus. To be more specific, F0 contour of an F0M stimulus, durational pattern of a DM stimulus and F0 and durational pattern of an F0DM stimulus determine the expected response to that stimulus. Table 5.1 presents the expected response to the stimuli in different manipulation conditions.

Table 5.1. Expected response to stimuli across manipulation conditions.

Manipulation condition	Stimulus type	Expected response
F0 manipulation	F0qDs ^a	Question
	F0sDq	Declarative
Duration manipulation	F0sDq	Question
	F0qDs	Declarative
F0 and duration manipulation	F0qDq	Question
	F0sDs	Declarative
Resynthesized original	Declarative ^b	Declarative
	Wh-question ^c	Question

a. F0q means that a stimulus has the F0 contour of a question and Ds means that a stimulus has the durational pattern of a declarative.

b & c. Sentence type can be defined for resynthesized original items since they are not manipulated. However, sentence type of the manipulated stimuli cannot be determined because the prosodic correlates of the stimuli are manipulated in manipulation conditions. Hence, abbreviations for prosodic correlates were used to define the stimulus type in manipulation conditions.

The responses given by participants were transferred from E-prime to SPSS version 22 (IBM SPSS, 2012). If the given response is compatible with the expected response, it is coded as correct and if it is not compatible with the expected

response, it is coded as incorrect. Thus, response is reported in terms of percentage of correct answers. To correct for possible response bias, responses were transformed to A' scores (Stanislaw & Todorov, 1999). In addition to percentage of correct answers, A' scores of the responses to each manipulation condition are reported as well. A repeated-measures ANOVA (RM-ANOVA) was run to investigate the effect of manipulation condition on response accuracy.

RT data were also transferred from E-prime to SPSS. RT was calculated as the time between response and stimulus offset. A separate RM-ANOVA was performed to inspect the influence of manipulation condition on RT.

5.4 Results

5.4.1 Response Accuracy

Table 5.2 presents the percentage and number of declarative and question responses given by participants across manipulation conditions.

Table 5.2. Percentage and number of declarative and question responses across manipulation conditions.

		Question responses		Declarative responses		Missing cases		Total	
		%	N	%	N	%	N	%	N
F0 manipulation	F0qDs ^a	53.8	258	45.8	220	0.4	2	100	480
	F0sDq	40.4	194	59.6	286	0	0	100	480
Duration manipulation	F0sDq	26.5	127	73.1	351	0.4	2	100	480
	F0qDs	61.3	294	38.3	184	0.4	2	100	480
F0 & duration manipulation	F0qDq	80.4	386	19.6	94	0	0	100	480
	F0sDs	11.1	53	88.5	425	0.4	2	100	480
Resynthesized original	Declarative ^b	4.4	21	95.6	459	0	0	100	480
	Wh-question ^c	87.5	420	12.5	60	0	0	100	480

a. F0q means that a stimulus has the F0 contour of a question and Ds means that a stimulus has the durational pattern of a statement.

b & c. Sentence type can be defined for resynthesized original items since they are not manipulated. However, sentence type of the manipulated stimuli cannot be determined because the prosodic correlates of the stimuli are manipulated in manipulation conditions. Hence, abbreviations for prosodic correlates were used to define the stimulus type in manipulation conditions.

Table 5.3 presents the response accuracy across manipulation conditions. As Table 5.3 illustrates, the response accuracy increases in the following order: DM < FOM < FODM < RSO.

Table 5.3. Perception of sentence type per condition.

	Correct ^a		Incorrect ^a		Missing cases		Total	
	%	N	%	N	%	N	%	N
F0 manipulation	56.7	544	43.1	414	0.2	2	100	960
Duration manipulation	32.4	311	67.2	645	0.4	4	100	960
F0 & duration manipulation	84.5	811	15.3	147	0.2	2	100	960
Resynthesized original	91.6	879	8.4	81	0	0	100	960

a. As mentioned in Section 5.3.4, the given response is coded as correct if it is compatible with the expected response and it is coded as incorrect if it is not compatible with the expected response.

As mentioned in Section 5.3.4, the responses were transformed to A' to correct for the possible response bias (Stanislaw & Todorov, 1999). 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. Mean A' score for each manipulation condition is presented in Figure 5.3.

To inspect the effect of the manipulation condition on response accuracy, a RM-ANOVA was run with aggregated responses as the dependent variable and the manipulation condition as the independent variable. The multivariate test demonstrated a main effect of the manipulation condition on the response accuracy ($F(3,21) = 227.161$, $p < .01$; Wilks' Lambda = .030, $\eta_p^2 = .970$). Pairwise comparison tests using the Bonferroni correction delineated that the difference between all manipulation conditions is significant $p < .01$ except for the difference between the FOM and DM items $p > .05$.⁴⁵

⁴⁵ A separate RM-ANOVA was conducted with the manipulation condition as the independent variable and the A' data as the dependent variable. The multivariate test showed a significant effect of the manipulation condition on the A' , ($F(3,21) = 316.007$, $p < .01$; Wilks' Lambda = .022, $\eta_p^2 = .978$). Pairwise comparison tests using the Bonferroni correction delineated that the difference between all manipulation conditions is significant $p < .01$ except for the difference between the FOM and DM items $p > .05$.

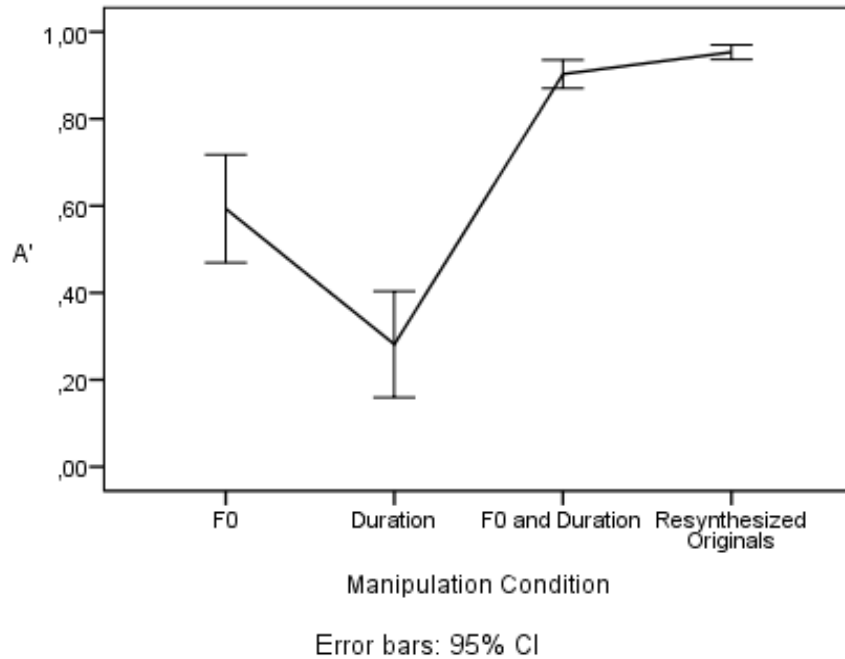


Figure 5.3. Mean A' scores across manipulation conditions.

5.4.2 Reaction time analysis

Reaction times were calculated from the stimulus offset. Table 5.4 provides mean and standard deviation of RT across manipulation conditions. All RT data is reported in seconds.

Table 5.4. Mean reaction time and standard deviation (in sec) across manipulation conditions

	Mean	SD
F0 manipulation	0.577	0.323
Duration manipulation	0.567	0.307
F0 and duration manipulation	0.537	0.329
Resynthesized original sound	0.490	0.315

According to Table 5.4, the RT to different manipulation conditions increases as follows RSO < F0DM < DM < F0M. RT data was submitted to a RM-ANOVA with

manipulation conditions as independent variable.⁴⁶ A multivariate test showed a main effect of the manipulation condition ($F(3,21) = 16.231, p < .001$; Wilks' Lambda = .301, $\eta_p^2 = .699$).⁴⁷ Pairwise comparison tests revealed that the difference between all manipulation conditions is significant ($p < .03$), except for the difference between DM and F0M and between DM and F0DM. The p-value was adjusted for multiple comparisons using the Bonferroni correction test.

5.5 Discussion and conclusion

The present study was designed to weigh prosodic cues to the identification of wh-in-situ questions as opposed to declaratives in Persian. The result is consistent with the prediction that F0 variation is the more important contributor to the sentence modality contrast in Persian and that sentence type identification is improved in the congruent cue condition in comparison to the incongruent cue condition. Sentence type categorization across manipulation conditions is improved in the following order: DM < F0M < F0DM < RSO. Sentence type identification in the F0M items is higher in comparison to the sentence type identification in the DM items. A possible justification is that listeners draw on F0 cues even in the duration-manipulated condition to decide on the sentence type. This finding corroborates what has been proposed in the literature: F0 performs a prominent role (Cangemi & D'Imperio, 2013; Peng et al., 2012) and duration a secondary role in the recognition of sentence modality contrast (Cangemi & D'Imperio, 2013). Categorization of sentence type in the RSO and F0DM items (congruent cue condition) was augmented in comparison to the F0M and DM items (incongruent cue condition). This result suggests that the proposal of the increased identification of speech contrast in the congruent cue condition in comparison to the conflicting cue condition (e.g. Fry, 1955; Hazan & Rosen, 1991) is applicable to wh-question vs. statement identification in Persian.

RT to different manipulation conditions increases as follows: RSO < F0DM < DM < F0M. In line with our predictions, the RT to congruent cue conditions is shorter than the RT to the incongruent cue conditions.

To our knowledge, no study has investigated whether or not dialectal differences affect the role of prosody in the production and perception of Persian interrogatives, including wh-in-situ questions vs. statements. Therefore, we suggest that the results of this study might be limited to standard Persian.

The general conclusion of this study sets the ground for further research: whether the dominant role of F0 and the improved perception of the sentence type in

⁴⁶ In accordance with the ANOVA model, the RT data were transformed by log10 to normalize the distribution. The missing cases ($f = 10, 0.16\%$) were discarded. They are reported in Table 5.2.

⁴⁷ To check whether including incorrect decisions (see Table 5.3) in the RT analysis influences the results, a separate RM-ANOVA was conducted with manipulation conditions as independent variable and the RT for only correct responses as the dependent variable. Similar to the result of RM-ANOVA on the RT for all responses, a multivariate test showed a main effect of the manipulation condition ($F(3,21) = 28.058, p < .001$; Wilks' Lambda = .200, $\eta_p^2 = .800$). Pairwise comparison tests revealed that the difference between all manipulation conditions is significant ($p < .05$), except for the difference between DM and F0M and between DM and F0DM. The p-value was adjusted for multiple comparisons using the Bonferroni correction test. We suggest that analyzing only RTs for correct decisions does not influence the results.

84 Prosody and processing of wh-in-situ questions in standard Persian

congruent cue conditions in comparison to conflicting cue conditions is extendable to other Persian structures that lack syntactic markers and are prosodically marked.

Chapter six

Discussion and conclusion

6.1. Introduction

Previous studies on the intonation system of Persian, such as Mahjani (2003) and Sadat Tehrani (2008), have largely concentrated on the phonological description of the accentual structure of different sentence types, including wh-questions and declaratives. The role of prosodic correlates in marking different parts of wh-in-situ questions (the complete utterance, including the pre-wh part and the post wh-part) as opposed to declaratives is, however, an understudied area in research in Persian prosody. Further understudied topics include the role of prosodic correlates in the identification of wh-in-situ questions before the occurrence of the wh-phrase in Persian.

This thesis investigated a) whether prosodic correlates characterize the pre-wh part, the post-wh part and the complete utterance of Persian wh-in-situ questions in the absence of the wh-phrase in sentence-initial position, b) whether the prosody of the pre-wh part of the sentence cues the identification of wh-in-situ questions in Persian, c) if the F0, the duration or their cooperation contributes more to the correct recognition of sentence type, and d) where in the pre-wh part Persian native speakers are able to identify wh-in-situ questions. To the best of our knowledge this is the first study to investigate the role of prosody in marking wh-in-situ questions as opposed to declaratives and in the identification of wh-in-situ questions before the wh-phrase is uttered by the speaker in Persian.

Chapter 1 comprised an introduction to the dissertation. Chapter 2 presented the production experiment conducted to tap into the first research question. A repeated-measures ANOVA (RM-ANOVA) revealed that the pre-wh part in wh-questions was marked by a shorter duration, a higher F0 onset, a raised pitch register level, a steeper upward F0 trend and a greater excursion size of the pitch accents associated with the pre-wh words. It was also shown that the complete sentence in wh-questions is made distinct from a declarative by the global acoustic correlates of raised pitch register level and shorter duration. Except for the F0 offset, which was higher in questions, other prosodic correlates in the post-wh part do not play a significant role in marking wh-questions.

Chapter 3 investigated whether the prosody of the pre-wh part can cue the identification of wh-in-situ questions in the absence of a wh-phrase at the beginning of the sentence. A perception experiment was conducted to investigate this question. The results showed that wh-questions were correctly distinguished from declaratives in 90.30% of cases. Accordingly, it was suggested that the prosodic characteristics of the pre-wh part of the sentence can cue identification of wh-in-situ questions as

opposed to declaratives in Persian. A RM-ANOVA showed that the reaction time (RT) to declaratives was shorter than the RT to questions.

The perception experiment in Chapter 3 is based on one-time stimulus presentation. This method cannot delineate how the decision whether a sentence is going to be a statement or a question develops over time and it cannot determine the amount of acoustic-phonetic information needed to identify a stimulus (Grosjean, 1996). Therefore, a gating paradigm perception experiment (cf. Grosjean, 1980) was designed (Chapter 4). The gating experiment tackled the question as to where in the pre-wh part Persian native speakers can identify the sentence type. In this experiment response accuracy, RT and confidence rating were recorded. The results showed that already upon hearing the first gate, where the F0 onset is introduced, listeners are able to recognize the sentence type above chance level (75.5%). Correct recognition of the sentence type and confidence rating reach the highest level of 91.5% and 4.4 (on a scale of 5) respectively at the last gate, where the pre-wh part was presented in its entirety. A RM-ANOVA indicated that identification improves, confidence rating increases and RT decreases as the amount of discriminating prosodic information increases.

Sentence type can be identified based on the prosody of the pre-wh part, that is before the occurrence of the lexical cue to sentence type (i.e., the wh phrase). This raises the question as to whether F0, durational cues in the pre-wh part or their cooperation contributes more to the correct recognition of the sentence type. A perception experiment with four conditions was conducted to examine the contribution of these variables and their cooperation, to the identification of wh-questions as opposed to declaratives (Chapter 5). The four conditions were: a) F0 manipulation condition, where only the F0 was manipulated (F0M), b) duration manipulation condition, where only the duration was manipulated (DM), c) F0 and duration manipulation condition, where both F0 and duration were manipulated (F0DM), and d) resynthesized condition, where a resynthesized version of the original sound was used (RSO). In the RSO condition, the F0 contour of the stimuli was stylized but none of the cues were manipulated. The F0M and DM conditions investigated the contribution of F0 and durational cues to sentence type identification respectively. The F0DM condition examined the contribution of both F0 and duration (their cooperation) to sentence type recognition. The RSO condition was used as the baseline for the experiment. The F0M and DM are called mismatching cue conditions, while the F0DM and RSO are called matching cue conditions. The results revealed a primary role for F0 and a secondary role for duration in the perception of the prosodic contrast. According to the results of the RM-ANOVA, identification increased and RT decreased in the matching cue conditions as compared with the mismatching cue conditions.

6.2 Theoretical implications

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 or are fewer in the sentence (Van Heuven & Haan, 2000). Prosody is the defining feature of questions in the absence of a syntactic interrogativity marker (e.g. Bolinger, 1978; Ohala, 1983, 1984). Generally, the results of the production experiment (Chapter 2) indicate that Lindblom's (1990) proposal is applicable to Persian *wh*-in-situ questions, as an interrogative sentence type where the lexico-syntactic interrogativity feature occurs later in the sentence.

The production experiment (Chapter 2) provides evidence for the association of high pitch (Bolinger, 1964, 1978, 1989; Ohala, 1983, 1984; Bartels, 1997; Gunlogson, 2008) and faster speech rate⁴⁸ (Bolinger, 1989; Van Heuven & Van Zanten, 2005) with questions. Ohala posited that the questioner needs to be polite or subservient to the respondent. Hence, the speaker raises his/her pitch as a sign of submission or politeness to the hearer. Bartels (1997) claimed that rising intonation shows a dependency of the speaker on the addressee. Ohala (1983, 1984) and lately Gunlogson (2008) suggested that rising intonation implies that the speaker is dependent on the addressee's information and that the addressee is the sole source of information. Therefore, the addressee is the authority and has superiority over the speaker as a source. The purpose of asking a question is to elicit a verbal response from the respondent and to fill in the knowledge gap that the questioner has. This purpose implies that the questioner is uncertain and is dependent on the respondent's cooperation to provide him or her with an answer. Therefore, we can propose that questioners use high F0 to show their dependence on the respondent. On the other hand, the respondents use low F0 to show their authority. Van Heuven & Van Zanten (2005) proposed that a faster speech rate may express the same message conveyed by a high F0 to the hearer. Therefore, it can be suggested that high pitch is tied to a fast speech rate in signaling questions. Bolinger's (1989) theory on the dichotomous association between relaxation and declarativity on the one hand, and tension and interrogativity on the other, can be the reason for the association of high pitch and fast rate with questions. Indeed he claimed that statements and questions are characterized by a dichotomy between relaxation (low, falling pitch) and tension (high, rising pitch), respectively (Bolinger, 1989). It would seem a natural extension of this principle to argue that slow speech rate or deceleration would be another correlate of relaxation, and that high speech rate and acceleration are associated with tension (Van Heuven & Van Zanten, 2005).

Lindblom's (1990) hyper- and hypo-theory of speech could also possibly explain the characteristic role of prosody in the pre-*wh* part in contrast with its insignificant role in marking the post-*wh* part in *wh*-questions (Chapter 2). The prosodic interrogativity cues will be stronger in the pre-*wh* part since the syntactic cue (the *wh*-phrase) to the sentence type is absent. On the other hand, the role of prosodic interrogativity marking becomes weaker in the post-*wh* part since the syntactic cue to the sentence type is uttered before the post-*wh* part. Indeed, the appearance of the syntactic cue to the sentence type obviates or diminishes reliance on the role of prosody in cueing the sentence type in the post-*wh* part.

⁴⁸ This equates with a shorter duration of utterances.

Based on earlier studies on the cuing of syntactic structure by prosody (e.g. Cooper & Paccia-Cooper, 1980),⁴⁹ Snedeker and Trueswell (2003) proposed that language users possess some implicit knowledge about the relationship between prosody and syntax, and are capable of using this knowledge to guide linguistic choices. Prosodic marking of Persian wh-in-situ questions (Chapter 2) and identification of sentence type before the occurrence of the syntactic cue to sentence type (Chapters 3, 4, 5) is in line with Snedeker and Trueswell's (2003) proposal. Speakers in these studies were capable of using prosody to encode (Chapter 2) and decode (Chapters 3, 4, 5) the sentence type before the cue to the sentence type was uttered. Listeners use their language-specific implicit knowledge that high pitch and fast speech rate represent an interrogative while low pitch and slow speech rate denote a declarative in order to identify the sentence type (Chapter 3, 4).

Listeners' use of prosody to predict the sentence type (Chapters 3, 4) before the occurrence of the syntactic cue can be discussed from the point of view of processing. Prediction can facilitate the processing and decoding of the sentence type and response preparation on the part of the listener (Grosjean, 1983, 1996). Therefore, prediction and how prosody guides prediction must be integrated into language processing models (Grosjean, 1983, 1996). Processing models also need to account for the fact that the listener's prediction can be reset as more prosodic information becomes available (Grosjean, 1983 & 1996).

Wh-in-situ questions can be considered to be sentences with temporary syntactic ambiguity. Chapters 3 and 4 demonstrated that prosody can guide the resolution of this temporary syntactic ambiguity before the occurrence of the syntactic cue to sentence type. In line with Beach (1991), this result suggests that models of spoken sentence processing need to account for the online use of prosody in temporary syntactic ambiguity resolution.

From the view of efficient communication, it is logical that sentence mode is recognized by the listener as early as possible in an utterance (Chapter 4; Brazil, 1981; Petrone & Niebuhr, 2014). Identifying the sentence type in the absence of syntactic or lexical cues and based on prosody serves this communicative need.

6.3 Suggestions for further research

This dissertation has found that prosody has an influential role not only in the characterization but also in the identification of wh-in-situ questions as opposed to declaratives in Persian. However, more studies are required to tap into the cognitive and neural mechanisms of prosody-driven processing of wh-in-situ questions as sentences with temporary syntactic ambiguity. Neurolinguistic research techniques, such as electroencephalography (EEG), are needed to examine how prosody guides identification and processing of wh-in-situ questions. Due to its fine-grained time resolution, EEG can give additional insights into the time course of prosody processing.

⁴⁹ There are even earlier studies that show the cuing of syntax by prosody, e.g. Blesser (1969) and Lindblom and Svensson (1973).

While Persian *wh-in-situ* questions have temporary syntactic ambiguity, Persian yes-no questions without the question word “*aya*” have global syntactic ambiguity. Following Lindblom (1990), and based on the results of the current dissertation, it can be proposed that prosody also contributes to the characterization and the identification of yes-no questions. Investigating the role of prosody in the marking and processing of Persian yes-no questions is therefore another suggestion for further research.

References

- Abedi, F., Moinzadeh, A., & Gharaei, Z. (2012). WH-movement in English and Persian within the framework of government and binding theory. *International Journal of Linguistics*, 4, 419-432.
- Abolhasanizadeh, V., Bijankhan, M., & Gussenhoven, C. (2012). The Persian pitch accent and its retention after focus. *Lingua*, 122, 1380-1394.
- Adli, A. (2010). 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 C. Gussenhoven & T. Riad (Eds.), *Tones and tunes: Experimental studies in word and sentence prosody* (pp. 387-405). Berlin: Mouton de Gruyter.
- Baltazani, M., Kainada, E., Lengeris, A., & Nicolaidis, K. (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*. Glasgow: University of Glasgow. Paper number 1041.1-5 retrieved from <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0941.pdf>.
- Bartels, C. (1997). *Toward a compositional interpretation of English statement and question intonation* (Doctoral dissertation, University of Massachusetts Amherst). Retrieved from <http://search.proquest.com/docview/304352313>.
- 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. *Journal of the Acoustical Society of America*, 99, 1148-1160.
- Benkirane, T. (1998). Intonation in Western Arabic (Morocco). In D. Hirst, & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 345-359). Cambridge, UK: Cambridge University Press.
- Beyssade, C., Delais-Roussarie, E., & Marandin, J. M. (2007). The prosody of interrogatives in French. *Nouveaux Cahiers de Linguistique Française*, 28, 163-175.
- Boersma, P. & D. Weenink. (2014). Praat: Doing phonetics by computer (Version 6.0.04) [Computer program]. <http://www.praat.org/>.
- Blessner, B. A. (1969). *Perception of spectrally rotated speech*. (Doctoral dissertation, Massachusetts Institute of Technology). Retrieved from <http://www.blessner.net/downloads/Perception%20of%20Spectrally%20Rotated%20Speech.pdf>.
- Bolinger, D. (1964). Intonation as a universal. In H. G. Lunt (Ed.), *Proceedings of*

- the Ninth International Congress of Linguists* (pp. 833-848). The Hague: Mouton & Co.
- Bolinger, D. (1978). Intonation across languages. In J. H. Greenberg (Ed.), *Universals of human language* (Vol. 2, pp. 471-524). Stanford: Stanford University Press.
- Bolinger, D. (1982). Nondeclaratives from an intonational standpoint. In R. Schneider, K. Tuite & R. Chametzky (Eds.), *Papers from the parasession on nondeclaratives* (pp. 1-22). Chicago: Chicago Linguistic Society.
- Bolinger, D. (1989). *Intonation and its uses*. Palo Alto: Stanford University Press.
- 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, UK: Routledge & Kegan Paul.
- Cangemi, F., & D'Imperio, M. (2013). Tempo and the perception of sentence modality. *Labarotory Phonology*, 4, 191-219.
- 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, UK: Blackwell.
- Cheng, L. L. S. (1991). *On the typology of wh-questions*. (Doctoral dissertation, Massachusetts Institute of Technology). Retrieved from https://www.researchgate.net/publication/38004328_On_the_Typology_of_Wh-Questions.
- Cheng, L. L. S., & Rooryck, J. (2000). Licensing wh-in-situ. *Syntax*, 3, 1-19.
- Chisholm, W. S., Milic, L. T., & Greppin, J. A. C. (Eds.). (1984). *Interrogativity: a colloquium on the grammar, typology and pragmatics of questions in seven diverse languages*. Amsterdam: John Benjamins Publishing.
- Chomsky, N. (1977). On wh-movement. In P. W. Culicover, T. Wasow & A. Akmajian (Eds.), *Formal syntax* (pp. 71-132). New York: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum Associates.
- Cohen, A., & 't Hart, J. (1967). On the anatomy of intonation. *Lingua*, 19, 177-192.
- Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.
- Cooper, W. E., & Sorensen, J. M. (1981). *Fundamental frequency in sentence production*. New York: Springer-Verlag.
- Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Dabir-Moghaddam, M. (1982). *Syntax and semantics of causative constructions in Persian* (Doctoral dissertation, University of Illinois). Retrieved from ProQuest Dissertations and Theses database.
- de Moraes, J. A. (1998). Intonation in Brazilian Portuguese. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 179-194). Cambridge, UK: 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* (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.
- Esposito, C., & Barjam, P. (2007). The intonation of questions in Farsi: Wh-questions, yes/no questions, and echo questions. *ULCA Working Papers in Phonetics*, 105, 1-18.
- 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.
- Ferguson, C. A. (1957). Word stress in Persian. *Language*, 33, 123-135.
- Field, A. (2009). *Discovering statistics using SPSS*. London, UK: Sage Publications.
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of Acoustical Society of America*, 27, 765-768.
- Gebara, E. S. (1976). *Alguns aspectos da intonação no Português* (Unpublished master's thesis). Universidade Estadual de Campinas.
- Geluykens, R. (1986). Questioning intonation: An empirical study into the prosodic feature 'rising intonation' and its relevance for the production and recognition of questions. *Antwerp Papers in Linguistics*, 48, 1-103.
- 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. (1996). Using prosody to predict the end of sentences in English and French: Normal and brain damaged subjects. *Language and Cognitive Processes*, 11, 107-134.
- Gryllia, S., Cheng, L. L. S., & Doetjes, J. (2016). On the intonation of French wh-in-situ questions: What happens before the wh-word is reached? *Proceedings of Speech Prosody*, Boston, 611-614. doi: 10.21437/SpeechProsody.2016-125.
- 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.
- Gunlogson, C. (2008). A question of commitment. *Belgian Journal of Linguistics*, 22, 101-136.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge, UK: 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.
- Haan, J., Van Heuven, V. J., Pacilly, J. & Van Bezooijen, R. (1997). An anatomy of Dutch question intonation. *Linguistics in the Netherlands*, 14, 97-108.

- Hadding-Koch, K. (1961). *Acoustic-phonetic studies in the intonation of Southern Swedish*. Lund: Gleerup.
- Hadding-Koch, K., & Studdert-Kennedy, M. (1964). An experimental study of some intonation contours. *Phonetica*, 11, 175-185.
- Hamlaoui, F. (2008). *On the role of discourse and phonology in French wh-questions* (Unpublished master's thesis). University of Ottawa.
- Hamlaoui, F. (2011). On the role of phonology and discourse in Francilian French wh-questions. *Journal of Linguistics*, 47, 129-162.
- Hazan, V., & Rosen, S. (1991). Individual variability in the perception of cues to place contrasts in initial stops. *Perception & Psychophysics*, 49, 187-200.
- Hermann, E. (1942). *Probleme der Frage*. Göttingen: Vandenhoeck & Ruprecht.
- Hermes, D. J., & Van Gestel, J. C. (1991). The frequency scale of speech intonation. *Journal of the Acoustical Society of America*, 90, 97-102.
- 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)*, USA, 1957-1960.
- IBM SPSS. (2012). IBM SPSS. *IBM Software Business Analytics*.
- Iivonen, A. (1998). Intonation in Finnish. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 311-327). Cambridge, UK: Cambridge University Press.
- Inkelas, S., & Leben, W. (1990). Where phonology and phonetics intersect: the case of Hausa intonation. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 17-34). Cambridge, UK: Cambridge University Press.
- Ishihara, S. (2005). Prosody-scope match and mismatch in Tokyo Japanese Wh-Questions. *English Linguistics*, 22, 347-379.
- Jorrisen, C. (2014). *Ne, ne? The prosodic properties of the sentence final particle ne in Mandarin Chinese* (Unpublished master's thesis). Leiden University.
- Jun, S. A., & Oh, M. (1996). A prosodic analysis of three types of wh-phrases in Korean. *Language and Speech*, 39, 37-61.
- Kahnemuyipour, A. (2001). On wh-questions in Persian. *Canadian Journal of Linguistics*, 46, 41-61.
- Kahnemuyipour, A. (2003). Syntactic categories and Persian stress. *Natural Language and Linguistic Theory*, 21, 333-379.
- Kahnemuyipour, A. (2009). *The syntax of sentential stress*. Oxford, UK: Oxford University Press.
- Karimi, S. (1999). Is scrambling as strange as we think it is? *MIT Working Papers in Linguistics*, 33, 153-190.
- Karimi, S. (2003). On object position, specificity and scrambling in Persian. In S. Karimi (Ed.), *Word order and scrambling* (pp. 91-124). Oxford, UK: Blackwell Publishing.
- 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. Samiian & W.K. Wilkins (Eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation* (pp. 167-187). Amsterdam: Benjamins.

- Kitagawa, Y. (2005). Prosody, syntax and pragmatics of wh-questions in Japanese. *English Linguistics*, 22, 302-346.
- 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. (1996). *Intonational phonology* (1st ed.). Cambridge, UK: Cambridge University Press.
- Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Lazard, G. (1992). *A grammar of contemporary Persian*. (S. Lyons, Trans.). Costa Mesa: Mazda Publishers. (Original work published 1957).
- Lee, H. Y. (1990). *The structure of Korean prosody* (Doctoral dissertation, University of London). Retrieved from <http://discovery.ucl.ac.uk/1382398/1/395201.pdf>.
- Lee, O. J. (2005). *The prosody of questions in Beijing Mandarin* (Doctoral dissertation, Ohio State University). Retrieved from https://etd.ohiolink.edu/rws_etd/document/get/osu1122332580/inline.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge: MIT Press.
- Lehiste, I. (1976). Suprasegmental features of speech. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 225-239). New York, NY: Academic Publishers.
- Lieberman, P., Katz, W., Jongman, A., Zimmerman, R., & Miller, M. (1985). Measures of the sentence intonation of read and spontaneous speech in American English. *Journal of the Acoustical Society of America*, 77, 649-657.
- 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.
- Lindblom, B., & Svensson, S. (1973). Interaction between segmental and nonsegmental factors in speech recognition. *IEEE Transactions on Audio and Electroacoustics*, 21, 536 - 545.
- Lindsey, G. A. (1985). *Intonation and interrogation of tonal structure and the expression of a pragmatic function in English and other languages* (Doctoral dissertation, University of California). Retrieved from ProQuest Dissertations and Theses database.
- 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.
- Mahootian, S. (1997). *Persian descriptive grammars*. London, UK: Routledge.
- Marslen-Wilson, W. D., Tyler, L. K., Warren, P., Grenier, P., & Lee, C. S. (1992). Prosodic effects in minimal attachment. *Quarterly Journal of Experimental Psychology*, 45A, 73-87.
- 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 Press.
- 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.
- Ohala, J. (1983). Cross-language use of pitch: An ethological view. *Phonetica*, 40, 1-18.
- Ohala, J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica*, 41, 1-16.
- Pallant, J. (2013). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (5th ed.). Boston: MacGraw-Hill.
- Peng, S. C., Chatterjee, M., & Lu, N. (2012). Acoustic cue integration in speech intonation recognition with cochlear implants. *Trends in Amplification*, 16, 67-82.
- Petrone, C. (2008). *Le rôle de la variabilité phonétique dans la représentation des contours intonatifs et de leur sens* (Doctoral dissertation, Université de Provence). Retrieved from <https://tipa.revues.org/410>.
- 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.
- Rahmani, H., Rietveld, T., & Gussenhoven, C. (2018). Post-focal and factive deaccentuation in Persian. *Glossa: a journal of general linguistics*. doi: <http://doi.org/10.5334/gjgl.328>.
- Rialland, A. (2004). *A typology of question prosody in African languages*. Paper presented at the International Conference on Tone and Intonation, Santorini, Greece.
- Rietveld, A. C. M., & Gussenhoven, C. (1987). Perceived speech rate and intonation. *Journal of Phonetics*, 15, 273-285.
- Rietveld, T., & Vermillion, P. (2003). Cues for perceived pitch register. *Phonetica*, 60, 261-272.
- Rohany Rahbar, E. (2012). *Aspects of Persian phonology and morpho-phonology* (Doctoral dissertation, Ohio State University). Retrieved from ProQuest Dissertations and Theses database.
- 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: 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
- Sadat Tehrani, N. (2009). The alignment of L+H* pitch accents in Persian Intonation. *Journal of the International Phonetic Association*, 39, 205-230.
- Sadat Tehrani, N. (2011). The intonation patterns of interrogatives in Persian. *Linguistic Discovery Journal*, 9, 105-136.
- Sadat Tehrani, N. (2011). The intonation patterns of interrogatives in Persian. *Linguistic Discovery Journal*, 9, 105-136.
- Sadock, J. M., & Zwicky, A. M. (1985). Speech acts distinction in syntax. In T. Shopen (Ed.), *Language typology and syntactic description* (pp. 159-196). Cambridge, UK: Cambridge University Press.
- 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, UK: Berg Publishers.
- Sensui H. (1995). Percepción de la entonación interrogativa del español: Un estudio experimental. *Sophia Lingüística*, 38, 1-23.
- Shen, X. S. (1990). *The prosody of Mandarin Chinese* (Doctoral dissertation, University of California). Retrieved from ProQuest Dissertations and Theses database.
- 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: <https://doi.org/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: <https://doi.org/10.1177/0023830917753237>.
- Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (under review). When is a wh-in-situ question identified in Persian? *Language, Cognition and Neuroscience*.
- Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert J., & Hirschberg, J. (1992). TOBI: A standard for labeling English prosody. In J. J. Ohala, T. M. Nearey & B. L. Berwing (Eds.), *Proceedings of the Second International Conference on Spoken Language Processing* (pp. 867-870). Edmonton: University of Alberta.
- 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., and 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., Hoymann, G., Rossano, F., De Ruiter, J.P., Yoon, K.E., & Levinson, S.C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106, 10587-10592.
- Swerts, M., Strangert, E., & Heldnert, M. (1996). F0 declination in read-aloud and spontaneous speech. In H. T. Bunnell & W. Idsardi (Eds.), *Proceedings of the 4th International Conference on Spoken Language Processing* (pp. 1501-1504). Piscataway: IEEE.
- Taheri Ardali, M., & Xu, Y. (2012). Phonetic realization of prosodic focus in Persian. In Q. Ma., H. Ding & D. Hirst (Eds.), *Proceedings of the 6th International Conference on Speech Prosody* (pp. 326-329). Shanghai: Tongji University Press.
- Thorsen, N. (1980). A study of the perception of sentence intonation: Evidence from Danish. *Journal of the Acoustical Society of America*, 67, 1014-1030.
- Toosarvandani, M. (2004). Vowel length in modern Farsi. *Journal of the Royal Asiatic Society*, 14, 241-251.
- Toosarvandani, M. (2008). Wh-movement and syntax of sluicing. *Journal of Linguistics*, 44, 677-722.
- Towhidi, J. (1974). *Studies in the phonetics and phonology of modern Persian: Intonation and related features*. Hamburg: Helmut Buske Verlag.
- Truckenbrodt, H., Sandalo, F., & Abaurre, M. B. (2009). Elements of Brazilian Portuguese intonation. *Journal of Portuguese Linguistics*, 8, 75-114.
- Uldall, E. T. (1962). Ambiguity: Question or statement?, or, "Are you asking me or telling me?". In A. Rigault (Ed.), *Proceedings of the 4th International Congress of Phonetic Sciences* (pp. 779-783). The Hague: Mouton.
- Vaissière, J. (1983). Language-independent prosodic features. In A. Cutler & D. R. Ladd (Eds.), *Prosody: Models and measurements* (pp. 53-66). Berlin: Springer Verlag.
- Van Heuven, V. J., & Haan, J. (2000). Phonetic correlates of statement versus question intonation in Dutch. In A. Botinis (Ed.), *Intonation, analysis, modeling and technology* (pp. 119- 143). Dordrecht: Kluwer Academic Publishers.
- Van Heuven, V., & Van Zanten, E. (2005). Speech rate as a secondary prosodic characteristic of polarity questions in three languages. *Speech Communication*, 47, 87-99.
- Vion, M., & Colas, A. (2006). Pitch cues for the recognition of yes-no questions in French. *Journal of Psycholinguistics 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.
- Windfuhr, G. L. (1991). Persian. In B. Comrie (Ed.), *The world's major languages* (pp.523-546). London, UK: Routledge.
- Yang, Y., Gryllia, S., Doetjes, J., & Cheng, L. L. S. (2016a, September). *The role*

- prosody plays in disambiguation: a study on Mandarin*. Poster presented at the Conference on Tone and Intonation in Europe (TIE), Kent.
- Yang Y., Gryllia, S., Pablos, L., Doetjes, J., & Cheng, L. L. S. (2016b, September). *The role prosody plays in disambiguation: earlier than you thought*. Poster presented at the Conference on Architectures and Mechanisms for Language Processing (AMLaP), Bilbao.
- Zubizarreta, M. L. (1998). *Prosody, focus and word order*. Cambridge, MA: MIT Press.
- Zubizarreta, M. L. (2003). Intervention effects in the French wh-in-situ construction: Syntax or interpretation? In R. Núñez-Cedeño, L. López & R. Cameron (Eds.), *A Romance perspective on language, knowledge and use: Selected papers from the 31st symposium on Romance languages* (pp. 359-379). Amsterdam: John Benjamins.

Appendices

Appendix I.

Structure and words used as the constituents of sentence stimuli in declaratives and wh-questions.

Subject	Adverb	3rd constituent of the sentence		Verb
		Wh-phrase	DWC	
Inanimate DO ^a				
1) mohmædʔæ <u>min</u> (Mohamadamin)	1) seʃæn <u>be</u> (Tuesday)	1) <u>tʃi</u> (what)	1) <u>kæfʃ</u> (shoe) 2) <u>kif</u> (bag)	tæʔ <u>mir</u> -kærdæn (repair)
Animate DO				
2) mohæmædʔæ <u>li</u> (Mohamadali)	2) pæri <u>ruz</u> (two days ago)	2) <u>ki</u> (who)	1) <u>jas</u> (Yas) 2) <u>naʒ</u> (Naz)	dæʔ <u>yæt</u> -kærdæn (invite)
Adjunct of Place ^b				
		3) <u>kodʒa</u> (where)	1) <u>dʒæŋgæɭ</u> (jungle) 2) <u>kuʃe</u> (street)	bæʒi-kærdæn (play)
Adjunct of Time				
		4) <u>kei</u> (when)	1) <u>zohr</u> (noon) 2) <u>ʔæsr</u> (afternoon)	ʃenɑ-kærdæn (swim)
Adjunct of Manner				
		5) <u>tʃetori</u> (how)	1) bɑʔæ <u>nduh</u> (sadly) 2) bade <u>ggæt</u> (carefully)	soh <u>bæt</u> -kærdæn (talk)

a. DO refers to direct object. The object marker “ra” occurs after direct object in declaratives and wh-in-situ questions. The stressed syllable in each word is underlined.

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

Appendix II.

We ran a script in Praat to compute a higher regression line and a lower regression line for the pre-wh part, post-wh part and the complete sentence separately. The Praat script splits the pitch contour in an upper and lower half and computes a separate regression line for each half. A RM-MANOVA was run, with sentence type as the independent variable and the slope of the higher and lower regression line, the pitch mean and the duration of the complete sentence, the excursion size of the pre-wh words, the pitch mean and the slope of the higher and lower regression line of the pre-wh part and the slope of the higher and lower regression line, the pitch mean and the duration of the post-wh part as the dependent variables. Duration of the pre-wh part, the F0 onset and the F0 offset were not included in the RM-MANOVA (see Section 2.4.2.3). Similar to the results of the RM-ANOVA presented in Section 2.4.3, the multivariate test revealed a significant effect of the sentence type on the dependent variables, ($F(13,5) = 21.507, p < .003$; Wilks' Lambda = .018, $\eta_p^2 = .982$). See Table II.1 for the results.

Table II.1. Results of univariate tests and descriptive statistics for the acoustic correlates of the pre-wh part, the post-wh part and the complete sentence for the two sentence types (lower and higher regression lines are used to compute the declination of the stimuli).

Variable	F	df	p	η_p^2	M SD (Decl)	M SD (Wh-q)	N
Slope of higher regression line of the pre-wh part (ERB/sec)	28.593	(1,17) ^a	.000**	.627	0.180 0.320	0.662 1.125	1440
Slope of lower regression line of the pre-wh part (ERB/sec)	16.175	(1,17)	.001**	.488	-0.114 0.179	0.096 0.371	1440
Slope of higher regression line of the post-wh part (ERB/sec)	18.167	(1,17)	.001**	.517	-1.169 1.310	-3.145 2.134	1440
Slope of lower regression line of the post-wh part (ERB/sec)	4.581	(1,17)	.061	.016	-1.172 1.224	-2.293 2.243	1440
Slope of higher regression line of the Sentence (ERB/sec)	4.083	(1,17)	.059	.194	-0.241 0.297	-0.074 0.574	1440
Slope of lower regression line of the Sentence (ERB/sec)	4.662	(1,17)	.061	.204	-0.243 0.208	-0.459 0.386	1440

Note. Decl = declaratives; Wh-q = wh-in-situ questions.

^aThe values in brackets give the degrees of freedom for condition and participants.

** $p < .01$.

Similar to what has been reported about declination based on fitting a single regression in Section 2.4.3, the slope of the lower regression line shows that the post-wh part and the complete sentence in the both questions and declaratives has declination and the pre-wh part in questions has inclination. The lower regression line of the pre-wh part in declaratives has declination.

Appendix III.

To test the effect of the presence of high final boundary tones in questions, a reanalysis of the data was done excluding questions ending in H% and the corresponding declaratives. The independent and dependent variables are the same as the variables of the RM-MANOVA reported in Section 2.4.3. The variables which were highly correlated, namely the duration of the pre-wh part and that of the complete sentence ($r = 0.857$, $p < .001$), the pitch mean of the complete sentence and the F0 onset ($r = 0.915$, $p < .001$), the pitch mean of the complete sentence and the F0 offset ($r = 0.809$, $p < .001$) were not included in the RM-MANOVA (see Section 2.4.2.3). Similar to the results presented in Section 2.4.3, the multivariate test revealed a significant effect of the sentence type on the acoustic features of the pre-wh part, the post-wh part and the complete sentence, ($F(10,8) = 27.465$, $p < .001$; Wilks' Lambda = .028, $\eta_p^2 = .972$), see Table III.1. As these results demonstrate, the effect of sentence type (significant or non-significant) on individual variables is similar to the effects resulting from the RM-MANOVA conducted on the data including the questions with a H% (see Table 2.1).

Table III.1. Results of univariate tests and descriptive statistics for the acoustic correlates of the pre-wh part, the post-wh part and the complete sentence for the two sentence types excluding the wh-questions that end in a high boundary tone and their corresponding declaratives.

Variable	F	df	p	η_p^2	M SD (Decl)	M SD (Wh-q)	N
Subject pitch excursion (ERB)	28.536	(1,17) ^a	.000**	.627	1.008 0.396	1.509 0.606	1246
Adverb pitch excursion (ERB)	60.471	(1,17)	.000**	.781	0.704 0.329	1.349 0.420	1246
SLR of the pre-wh part (ERB/sec)	28.593	(1,17)	.000**	.627	0.058 0.202	0.589 0.457	1246
SLR of the post-wh part (ERB/sec)	18.167	(1,17)	.001**	.517	-1.128 0.913	-2.872 2.225	1246
SLR of the Sentence (ERB/sec)	4.083	(1,17)	.059	.194	-0.180 0.173	-0.353 0.356	1246
F0 mean of the pre-wh part (ERB)	88.966	(1,17)	.000**	.840	4.953 1.362	5.873 1.484	1246
F0 mean of the post-wh part (ERB)	0.000	(1,17)	.999	.000	4.572 1.183	5.306 1.298	1246
F0 mean of the complete sentence (ERB)	80.883	(1,17)	.000**	.826	4.852 1.297	5.722 1.401	1246
Post-wh part duration (sec)	0.635	(1,17)	.437	.036	0.841 0.134	0.818 0.131	1246
Duration of the sentence (sec)	143.486	(1,17)	.000**	.894	2.655 0.257	2.204 0.228	1246
Pre-wh part duration (sec) ^b	-	-	-	-	1.368 0.136	1.098 0.130	1246
F0 onset (ERB) ^b	-	-	-	-	5.079 1.450	6.103 1.626	1246
F0 offset (ERB) ^b	-	-	-	-	4.578 1.148	5.278 1.300	1246

Note. Decl = declaratives; Wh-q = wh-questions; SLR = slope of regression line.

^a The values in brackets give the degrees of freedom for condition and participants.

^b Since the duration of the pre-wh part, the F0 onset and the F0 offset were not included in the RM-MANOVA, only the descriptive statistics of these variables are reported.

**p < .01.

Appendix IV.

Descriptive statistics of the prosodic correlates of the stimuli across sentence types.

Variable	M SD (Decl)	M SD (Wh-q)	N
F0 onset (ERB)	4.999 1.209	5.828 0.988	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 the 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 (sec)	1.360 0.072	1.043 0.069	40

List of figures and tables

List of figures

Figure 2.1 Prosodic structure of a simple sentence in Persian	15
Figure 2.2 The acoustic correlates measured in the pre-wh part of a declarative	23
Figure 2.3 The acoustic correlates measured in the pre-wh part of a question	24
Figure 3.1 The pitch track of the pre-wh part of a declarative	33
Figure 3.2 The pitch track of the pre-wh part of a wh-question	34
Figure 4.1 The seven gates of a declarative stimulus	55
Figure 4.2 The seven gates of a question stimulus	56
Figure 4.3 Mean A' scores across gates	59
Figure 4.4 Mean reaction time (in seconds) across gates	62
Figure 4.5 Mean confidence rating across gates	63
Figure 4.6 Mean confidence rating across sentence type and gates	64
Figure 5.1 The waveform, original (non-stylized) pitch contour and stylized pitch contour of the pre-wh part of a declarative	75
Figure 5.2 The waveform, original (non-stylized) pitch contour and stylized pitch contour of the pre-wh part of a wh-question	76
Figure 5.3 Mean A' scores across manipulation conditions	82

List of tables

Table 2.1 Results of the univariate tests and descriptive statistics for the acoustic correlates of the pre-wh part, the post-wh part and the complete sentence for the two sentence types	26
Table 3.1 Perception of intended sentence type per condition	42
Table 3.2 Mean reaction time (and standard deviation) (in sec) for declaratives and wh-questions	42
Table 4.1 Perception of intended sentence type across gates and sentence type	58
Table 4.2 Results of pairwise comparison tests for response accuracy differences between gates	60
Table 4.3 Mean reaction time (and standard deviation) (in sec) for declaratives and wh-questions across gates	61
Table 5.1 Expected response to stimuli across manipulation conditions	79
Table 5.2 Percentage and number of declarative and question responses across manipulation conditions	80
Table 5.3 Perception of intended sentence type per condition	81
Table 5.4 Mean reaction time and standard deviation (in sec) across manipulation conditions	82

English Summary

The contrast between statements and questions is functionally prominent in everyday communication. Utterances require no sign to be recognized as statements. On the other hand, questions may be signaled lexically, for example by the insertion of *wh*-phrases and/or other syntactic mechanisms, such as inversion of the subject and finite verb. However, lexico-syntactic features are not the sole indicators of interrogativity (Haan, 2001). The literature on question intonation specifies that intonation also plays a diagnostic role in characterizing questions (e.g. Hermann, 1942; Haan, 2001). It has been demonstrated that the marking of prosodic interrogativity is stronger when lexico-syntactic features of interrogativity are absent or are fewer⁵⁰ in the sentence (e.g. Lindblom, 1990). This finding can be extended to interrogatives where the lexico-syntactic interrogativity feature occurs later in the sentence, as in *wh*-in-situ questions. Persian is one of the languages characterized by *wh*-in-situ questions.

This dissertation investigates the role of prosody in characterizing the production and cueing the perception of *wh*-in-situ questions in the absence of the syntactic interrogativity marker (the *wh*-phrase) in sentence-initial position. Following a production experiment, three perception experiments were conducted to examine the contribution of prosody to the identification of *wh*-in-situ questions before the *wh*-phrase is uttered by the speaker in Persian.

Chapter 1 provides an introduction to the dissertation, a summary of the theoretical background, the research questions and associated hypotheses. Chapter 2 investigates whether specific prosody characterizes the pre-*wh* part, the post-*wh* part and the complete sentence in *wh*-questions as opposed to declaratives. It does so using a question-response task, which was designed to elicit declaratives and *wh*-questions from Persian native speakers. The results reveal that a higher F0 onset, a raised level of pitch register, a steeper F0 uptrend, a greater F0 excursion size of the pre-*wh* words and a shorter duration mark the pre-*wh* part in *wh*-questions. Regarding the complete sentence, a decreased duration and a raised register level are characteristic prosodic correlates of *wh*-questions as opposed to declaratives. The prosodic correlates in the post-*wh* part do not play a significant role in marking *wh*-questions, except for the F0 offset which is higher in questions.

Following the production experiment, three perception experiments were run to explore the role of prosody in identifying *wh*-questions as opposed to statements. Chapter 3 presents the results of a perception experiment conducted to answer the question as to whether the prosody of the pre-*wh* part guides the identification of *wh*-in-situ questions. Part of the sentences elicited in the production experiment forms the base of the material for the perception experiment. A forced-

⁵⁰ There is one lexico-syntactic feature of interrogativity, namely the auxiliary “did” in the English yes-no question “Did he study last night?”. In contrast, the English *wh*-question “When did he study last night?” has two lexico-syntactic features of interrogativity: the *wh*-phrase “when” and the auxiliary verb “did”. Thus, the number of lexico-syntactic features of interrogativity is smaller in the yes-no question in comparison to the *wh*-question.

choice sentence identification task was designed in which the pre-wh part of declaratives and wh-in-situ questions was played to Persian native speakers. Listeners were required to decide whether what they heard is going to be a wh-question or a declarative as quickly as possible. The results showed that wh-questions can be correctly distinguished from declaratives in 90.3% of cases. It may be concluded that the prosodic characteristics of the pre-wh part of the sentence can cue the identification of wh-in-situ questions as opposed to declaratives. In line with the results of Vion and Colas (2006), a RM-ANOVA showed that the reaction time (RT) to declaratives was shorter than the RT to questions.

As a follow-up to the previous perception experiment in Chapter 3, Chapter 4 examines where in the pre-wh part of a sentence Persian native speakers are able to perceive the correct sentence type. To this end, a perception experiment employing the gating paradigm was run (Grosjean, 1980). Part of the sentences elicited in the production experiment formed the base of the material of the perception experiment. The pre-wh part of the sentences was divided into seven gates. The gated stimuli were played to Persian native speakers in a forced-choice sentence identification task. After hearing each gate, they were instructed to opt for either declarative or wh-question as the sentence type as quickly as possible, and also report how confident they were about their decision on a 5-point scale. The results showed that upon hearing the first gate, where the F0 onset is introduced, listeners can already recognize the sentence type above chance level (75.6%). A RM-ANOVA revealed that identification improves, the confidence rating increases and the RT decreases as more discriminating prosodic information is presented. At the last gate, where the pre-wh part was completely presented, the peak of the correct recognition of the sentence type (91.5%) and the confidence rating (4.4 on a scale of 5) was achieved. The RT was also shortest (0.23 sec) at gate 7. Similar to the results of the previous perception experiment (Chapter 3) and Vion and Colas' (2006) study, declaratives have a decreased RT in comparison to questions.

Chapters 3 and 4 show that sentence type can be identified on the basis of the prosody of the pre-wh part, before the occurrence of the lexical cue to sentence type. The findings in Chapters 3 and 4 lead us to investigate the relative contribution of the F0 and durational cues to the correct recognition of the sentence type. Therefore, another perception experiment was run, as presented in Chapter 5. Similar to the previous perception experiments, a sentence identification task was designed and conducted. Part of the sentences elicited in the production experiment forms the basis of the material of this perception experiment. The F0 onset, the excursion size of the pre-wh words and the duration of the pre-wh part of wh-questions as well as statements were manipulated. Persian native speakers listened to the pre-wh part of declaratives and wh-in-situ questions. Upon hearing each stimulus, they were asked to decide as quickly as possible whether what they heard is going to be a wh-question or a declarative. The results revealed a primary role for the F0 and a secondary role for duration in prosodic contrast perception. A RM-ANOVA showed that sentence type identification was augmented and RT was decreased in the matching cue condition in comparison to the mismatching cue condition. In the matching cue condition, the F0 and the durational cues match while in the mismatching cue condition the F0 and the durational cues do not.

Chapter 6 revisits the research questions and draws conclusions from the main results of the dissertation. This chapter also provides suggestions for further research.

Samenvatting in het Nederlands

Het onderscheid tussen vraagzinnen en mededelingen heeft in de dagelijkse communicatie een prominente functie. Daar waar mededelingen niet op een specifieke manier worden gemarkeerd, kunnen vraagzinnen lexicaal worden gemarkeerd, bijvoorbeeld door toevoeging van *wh*-frases en/of andere syntactische middelen zoals inversie van de posities van het onderwerp en de persoonsvorm. Buiten deze lexico-syntactische mechanismen bestaan er echter nog andere types markeerders van vraagzinnen (Haan, 2001). Onderzoek naar vraagintonatie heeft uitgewezen dat intonatie ook een diagnostische rol speelt in de herkenning van vragen (bijv. Hermann, 1942; Haan, 2001). Men zou verwachten dat dit soort prosodische markering van vraagzinnen prominenter zal zijn wanneer lexico-syntactische kenmerken minder aanwezig⁵¹ of geheel afwezig zijn (bijv. Lindblom, 1990). Deze voorspelling kan worden gegeneraliseerd naar vraagzinnen waarin de lexico-syntactische markering van vraagzinnen later in de zin optreedt, zoals in *wh*-in-situ-vragen. Het Perzisch is een van de talen waarin *wh*-in-situ-vragen voorkomen.

In deze dissertatie wordt de rol van prosodie in de productie van *wh*-in-situ-vragen bij afwezigheid van syntactische markering van vraagzinnen (de *wh*-frase) op de zinsinitiële positie onderzocht. Na een productie-experiment zijn drie perceptie-experimenten uitgevoerd om te onderzoeken hoe prosodie bijdraagt aan de identificatie van *wh*-in-situ-vragen in het gedeelte van de zin voordat de *wh*-frase is uitgesproken door de spreker.

Hoofdstuk 1 leidt de dissertatie in in de vorm van een samenvatting van de theoretische achtergrond, de onderzoeksvragen en de hypothesen van het onderzoek. In Hoofdstuk 2 is onderzocht of en hoe prosodie het gedeelte van de zin voorafgaand aan de *wh*-frase (voortaan ‘pre-*wh*-deel’), het gedeelte volgend op de *wh*-frase (‘post-*wh*-deel’) en de zin als geheel de zin als vraagzin markeert. Dit is gedaan door middel van een vraag-antwoordtaak waarbij mededelingen en *wh*-vragen werden geëliciteerd bij moedertaalsprekers van het Perzisch. Uit dit onderzoek is gebleken dat een hogere F0-onset, een verhoogd toonhoogteregister, een steilere F0-stijging, een grotere F0-uitslag van de woorden in het pre-*wh*-deel en een kortere duur het pre-*wh*-deel markeren als vraagzin. In de zin als geheel corresponderen een verkorte duur en een verhoogd register met *wh*-vragen. Behalve een hogere F0-offset spelen de prosodische correlaten in het post-*wh*-deel spelen geen significante rol in de markering van *wh*-vragen.

Na het productie-experiment zijn drie perceptie-experimenten uitgevoerd om de rol van prosodie in de identificatie van *wh*-vragen vs. mededelingen uit te diepen. Hoofdstuk 3 is het verslag van een onderzoek dat tot doel had om te

⁵¹ In de ja/nee-vraag “Did he study last night?” is één lexico-syntactisch kenmerk van vraagzinnen aanwezig, het hulpwerkwoord “did”. Echter, de *wh*-vraag “When did he study last night?” heeft twee lexico-syntactische kenmerken, namelijk de *wh*-frase “when” en het hulpwerkwoord “did”. Het aantal lexico-syntactische kenmerken van bevraging is daarom lager in de ja/nee-vraag dan in de *wh*-vraag.

achterhalen of de prosodie van het pre-*wh*-deel de identificatie van *wh*-in-situ-vragen mogelijk maakt. De basis van het materiaal werd gevormd door een deel van de zinnen die geëliciteerd waren in het productie-experiment. Het perceptie-experiment betrof een meerkeuze-identificatietoets op basis van het pre-*wh*-deel van mededelingen en *wh*-in-situ-vragen en werd uitgevoerd door moedertaalsprekers van het Perzisch. De luisteraars werden verzocht zo snel mogelijk aan te geven of het gedeelte van de zin dat ze gehoord hadden onderdeel was van een *wh*-vraag of van een mededeling. De resultaten lieten zien dat *wh*-vragen in 90,3% van de gevallen correct konden worden onderscheiden van mededelingen. Hieruit kan worden geconcludeerd dat prosodische eigenschappen van het pre-*wh*-deel identificatie van het zinstype mogelijk maken. Uit een RM-ANOVA bleek, net als in Vion en Colas (2006), dat de reactietijden voor de identificatie van mededelingen korter waren dan die voor vraagzinnen.

Als vervolg op het experiment uit Hoofdstuk 3 is in het experiment beschreven in Hoofdstuk 4 middels een gating-paradigma (Grosjean, 1980) onderzocht welk gedeelte van het pre-*wh*-deel moedertaalsprekers van het Perzisch gebruiken om het zinstype te herkennen. De basis van het materiaal werd ook hier gevormd door een deel van de zinnen die geëliciteerd waren in het productie-experiment. Het pre-*wh*-deel werd in zeven gates verdeeld. Moedertaalsprekers van het Perzisch gaven in een meerkeuze-identificatietoets van elk van de gates aan a) (zo snel mogelijk) of ze dachten dat het onderdeel was van een mededeling of van een *wh*-vraag en b) hoe zeker ze van hun antwoord waren, aangegeven op een vijfpuntsschaal. De resultaten lieten zien dat luisteraars reeds bij het horen van de eerste gate, waarbij de F0-onset waarneembaar wordt, het zinstype met een score boven kansniveau (75,6%) konden waarnemen. Volgens een RM-ANOVA namen de identificatiescore en de zekerheidsoordelen toe en namen de reactietijden bij toenemende beschikbaarheid van prosodische informatie (dat wil zeggen, met langere gates) af. De hoogste herkenningsscore (91,5%) en het hoogste zekerheidsoordeel (4,47) werden bereikt bij de laatste gate. De reactietijd was daar bovendien met 0,23 s. het kortste. In lijn met de resultaten uit Hoofdstuk 3 en met bevindingen van Vion en Colas (2006) zijn reactietijden bij mededelingen korter dan bij vragen.

Hoofdstukken 3 en 4 hebben aangetoond dat het zinstype kan worden bepaald op basis van de prosodie van het pre-*wh*-deel (dat wil zeggen, voordat de syntactische markerings van zinstype hoorbaar wordt). Naar aanleiding van deze bevindingen hebben wij ons afgevraagd wat de relatieve bijdragen van F0- en duurgues voor zinstype zijn. Dit is in een nieuw experiment onderzocht, dat net als in de voorgaande hoofdstukken een zinstype-identificatietoets betrof (Hoofdstuk 5). Opnieuw werd de basis van het materiaal gevormd door een deel van de zinnen uit het productie-experiment. Nu werden de F0-onset, de F0-uitslag van de pre-*wh*-woorden en de duur van het pre-*wh*-deel van *wh*-vragen en mededelingen gemanipuleerd. Moedertaalsprekers van het Perzisch gaven bij het horen van deze pre-*wh*-delen zo snel mogelijk aan tot welk zinstype ze dachten dat de zin waar dat fragment een onderdeel van was behoorde. De resultaten wezen op een primaire rol voor F0-informatie en een secundaire rol voor duurginformatie. Een RM-ANOVA toonde aan dat in de match-conditie de identificatiescore was toegenomen en de reactietijd was afgenomen ten opzichte van de mismatch-conditie. In de conditie met

corresponderende cues zijn de F0- en de duercues met elkaar in overeenstemming terwijl die twee cues in de conditie met niet-corresponderende cues niet met elkaar in overeenstemming zijn. Overeenstemming wil hierbij zeggen dat de cues behoren bij hetzelfde zinstype.

In Hoofdstuk 6 wordt in het licht van de resultaten van de verschillende experimenten teruggekomen op de onderzoeksvragen en worden de belangrijkste resultaten van de dissertatie op een rij gezet. In dit hoofdstuk komen ook voorstellen voor toekomstig onderzoek aan bod.

Summary in Persian

خلاصه به زبان فارسی

جملات پرسشی (interrogatives) و خبری (declaratives)، دو نوع جمله هستند که تضاد کاربردی مهمی دارند و کاربرد فراوانی در مکالمات روزمره دارند. جملات برای اینکه به عنوان جملات خبری شناخته شوند به هیچ نشانه نحوی یا لغوی نیاز ندارند. اما جملات پرسشی ممکن است با استفاده از ابزارهای لغوی، به عنوان مثال اضافه کردن کلمات پرسشی، یا با استفاده از ابزارهای نحوی، مانند جابجایی فاعل و فعل، نشانه گذاری شوند (Haan, 2003). بررسی های انجام شده در زمینه آهنگ گفتار (prosody) جملات پرسشی نشان میدهد آهنگ گفتار نیز نقش تشخیصی در شناسایی جملات پرسشی دارند (به عنوان مثال: Hermaan, 1942; Haan, 2003). نشانه های آهنگ گفتار جملات پرسشی (prosodic interrogativity marking) در غیاب نشانه های نحوی-لغوی (lexico-syntactic features) در جملات پرسشی نقش پررنگ تری در بازنمایی جملات پرسشی دارند (به عنوان مثال: Lindblom, 1990). این فرضیه را می توان به سوالی های پرسشواژه در جای اصلی (wh-in-situ questions) تعمیم داد که در آنها نشانه نحوی-لغوی جمله پرسشی (گروه پرسشواژه یا wh-phrase) در ابتدای جمله ظاهر نمی شود. زبان فارسی یکی از زبان هایی است که در جرگه زبان های پرسشواژه در جای اصلی (wh-in-situ languages) قرار دارد.

این پایان نامه نقش آهنگ گفتار در بازنمایی سوالی های پرسشواژه در جای اصلی در غیاب گروه پرسشواژه (wh-phrase) در ابتدای جمله را بررسی می کند. این پایان نامه هم چنین به بررسی نقش آهنگ گفتار در شناسایی سوالی های پرسشواژه در جای اصلی قبل از بیان گروه پرسشواژه توسط گوینده می پردازد. فصل اول پایان نامه، مقدمه، پیشینه نظری، سوالات تحقیقی و فرضیه های این پایان نامه را ارائه می دهد. فصل دو نحوه بازنمایی بخش پیش پرسشواژه (pre-wh part)، پس پرسشواژه (post-wh part) و کل جمله را در سوالی های پرسشواژه در جای اصلی در مقایسه با جملات خبری توسط نواخت بررسی می کند. بدین منظور یک آزمایش تولید گفتار انجام شد. در این آزمایش، یک کار پرسش و جواب (question-response task) طراحی شد که در طول آن هجده گویشور بومی فارسی چهل جمله خبری و چهل سوالی پرسشواژه در جای اصلی را بیان کردند. نتایج این آزمایش تولید گفتار نشان داد آغاز به سامدپایه بالاتر (higher F0 onset)، سیاق زیربومی بالاتر (higher pitch register)، شیب رو به بالای تندتر به سامد پایه (steeper F0 uptrend)، میزان نوسان تکیه زیربومی بیشتر (larger excursion size of the pitch accent) و دیرش کمتر (shorter duration) بخش پیش پرسشواژه در سوالی های پرسشواژه در جای اصلی را در مقایسه با جملات خبری بازنمایی می کنند. مانند بخش پیش پرسشواژه در سوالی های پرسشواژه در جای اصلی، دیرش کمتر (shorter duration) و سیاق زیربومی بالاتر (higher pitch register) کل جمله در سوالی های پرسشواژه در جای اصلی را بازنمایی کردند. اما در بخش پس پرسشواژه، فقط پایانه به سامدپایه بالاتر (higher F0 offset) سوالی های پرسشواژه در جای اصلی را بازنمایی می کنند و دیگر ویژگی های آکوستیکی (acoustic correlates) نقشی در بازنمایی بخش پس پرسشواژه سوالی های پرسشواژه در جای اصلی ندارند.

پس از انجام آزمایش تولید گفتار، سه آزمایش ادراکی انجام شد. هدف این سه آزمایش بررسی نقش آهنگ گفتار در شناسایی سوالی های پرسشواژه در جای اصلی در مقایسه با جملات خبری است. فصل سه نتایج آزمایش ادراکی اول را ارائه می دهد که هدفش یافتن پاسخ این سوال است: "آیا آهنگ گفتار بخش پیش پرسشواژه باعث شناسایی سوالی های پرسشواژه در جای اصلی در مقابل جملات خبری می شود؟". بخشی از جملاتی که شرکت کنندگان در آزمایش تولید گفتار بیان کردند به عنوان مواد مورد استفاده آزمون (experimental material) در آزمایش ادراکی استفاده شد. یک آزمون شناسایی جمله اجباری طراحی شد. در این آزمون گویشوران بومی فارسی به بخش پیش پرسشواژه جملات خبری و پرسشی های پرسشواژه در جای اصلی گوش می دادند. آنها باید در کوتاهترین زمان ممکن نوع جمله (خبری یا پرسشی) های پرسشواژه

در جای اصلی) را تشخیص می دادند. نتایج نشان داد گویشوران بومی فارسی می توانند در 90.3 درصد موارد نوع جمله را درست تشخیص دهند. می توان نتیجه گرفت آهنگ گفتار بخش پیش پرسشواژه می تواند به تشخیص پرسشی های پرسشواژه در جای اصلی در مقایسه با جملات خبری کمک کند. مشابه نتایج تحقیق (Vion & Colas, 2006)، در این تحقیق نیز زمان عکس العمل به جملات خبری در مقایسه با زمان عکس العمل به پرسشی های پرسشواژه در جای اصلی کمتر است.

پس از انجام آزمایش ادراکی اول در فصل سه، در فصل چهار به سوال تحقیقی "در کدام قسمت از بخش پیش پرسشواژه گویشوران بومی فارسی می توانند نوع جمله را تشخیص دهند؟" برای پاسخ دادن به این سوال، آزمایش ادراکی دیگری با استفاده از تکنیک گیتینگ (gating) برگرفته از (Grosjean, 1980) طراحی شد. بخشی از جملاتی که گویشوران بومی فارسی در آزمایش تولید گفتار بیان کردند به عنوان مواد مورد استفاده آزمون در این آزمایش ادراکی استفاده شد. بخش پیش پرسشواژه جملات به هفت گیت (gate) تقسیم شد. مانند آزمایش ادراکی قبلی، یک آزمون شناسایی جمله اجباری طراحی شد. در این آزمون گویشوران بومی فارسی به محرک های گیت بندی شده (gated stimuli) گوش می دادند. از گویشوران بومی فارسی خواسته شد پس از شنیدن هر گیت در کوتاهترین زمان ممکن نوع جمله را تشخیص دهند و میزان اطمینان خود نسبت به پاسخ خود را با انتخاب عددی از یک تا پنج اعلام کنند. نتایج نشان داد که بعد از شنیدن گیت اول، گویشوران بومی فارسی می توانند با درصدی بالاتر از سطح تصادفی (75.6%) نوع جمله را تشخیص دهند. آنوای تکرار سنجش (RM-ANOVA) نشان می دهد که با افزایش اطلاعات آهنگ گفتار متمایز کننده (discriminating prosodic information) در گیت ها درصد تشخیص نوع جمله و میزان اطمینان شرکت کنندگان نسبت به دادن پاسخ درست افزایش و زمان عکس العمل کاهش پیدا می کند. بالاترین درصد تشخیص جمله (91.5%)، بالاترین میزان اطمینان نسبت به دادن پاسخ درست (4.47) و کوتاه ترین زمان عکس العمل (0.23 میلی ثانیه) در گیت آخر است که بخش پیش پرسشواژه جمله بطور کامل شنیده می شود. مشابه نتایج آزمایش ادراکی قبلی در این پایان نامه (فصل سوم) و نتایج تحقیق (Vion & Colas, 2006)، زمان عکس العمل به جملات خبری نسبت به جملات پرسشی کوتاه تر است.

در فصل سوم و چهارم نشان داده شد که نوع جمله براساس آهنگ گفتار بخش پیش پرسشواژه جمله (قبل از شنیدن گروه پرسشواژه) قابل تشخیص است. براساس یافته های فصل سوم و چهارم، ما به بررسی نقش نسبی بسامد پایه (F0) و دیرش در تشخیص نوع جمله پرداختیم. بدین منظور آزمایش ادراکی دیگری (فصل پنج) طراحی و انجام شد. مشابه آزمایش های ادراکی قبلی، یک آزمون شناسایی جمله اجباری طراحی شد. بخشی از جملاتی که شرکت کنندگان در آزمایش تولید گفتار بیان کردند به عنوان مواد مورد استفاده در آزمون در این آزمایش ادراکی استفاده شد. آغاز به سامدپایه، میزان نوسان تکیه زیرویمی کلمات بخش پیش پرسشواژه و دیرش بخش پیش پرسشواژه در جملات خبری و پرسشی دستکاری شدند. گویشوران بومی فارسی به بخش پیش پرسشواژه جمله در جملات خبری و پرسشی گوش دادند. آنها باید در کوتاهترین زمان ممکن نوع جمله را تشخیص می دادند. نتایج نشان داد که در اولویت بندی نقش بسامد پایه و دیرش در تشخیص نوع جمله، بسامد پایه نقش مهمتری نسبت به دیرش ایفا می کند. به عبارت دیگر، بسامد پایه نقش اولیه و دیرش نقش ثانویه را در تشخیص نوع جمله ایفا دارند. آنوای تکرار سنجش نشان داد که در شرایطی که نشانگرهای بسامد پایه با نشانگرهای دیرش همخوانی دارند، در مقایسه با شرایطی که نشانگرهای بسامد پایه با نشانگرهای دیرش همخوانی ندارد، درصد تشخیص نوع جمله افزایش و زمان عکس العمل کاهش پیدا می کند.

فصل شش سوالات تحقیق را بازخوانی می کند و یافته های اصلی و نتیجه گیری این پایان نامه را ارائه می دهد. این فصل، پیشنهاداتی برای تحقیقات بیشتر نیز ارائه می دهد.

Curriculum vitae

Zohreh Shiamizadeh was born in 1984 in Ghom, Iran and was brought up in Abadan, Iran. From 1999 to 2003 she studied Math in Palayeshgah High School in Abadan. She attended Shahid Chamran Univeristy from 2003 until 2007 and obtained her bachelor degree in English Translation. In 2008 she started a master's programme in Teaching English as a Foreign Language at Yazd University. She obtained her master's degree (with distinction) in 2010. From 2010 until 2013 she was a lecturer at the University of Applied Sciences and an instructor at Iran Language Institute in Ghom. In 2013 she began her PhD research on prosody and processing of wh-in-situ questions in standard Persian at Leiden University Center for Linguistics. This thesis is the result of her research.