東京大学大学院
総合文化研究科言語情報科学専攻
平成 26 年度博士論文

The Phonology and Phonetics of Prosodic Prominence in Persian

（ペルシア語における韻律的プロミネンスの音韻論と音声学）

指導教員 田中伸一 准教授
学生証番号 31－107024
ホセイニ・セイエッド アヤット
Hosseini Seyed Ayat

The Phonology and Phonetics of Prosodic Prominence in Persian

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL OF
ARTS AND SCIENCES, THE UNIVERSITY OF TOKYO
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

IN
LANGUAGE AND INFORMATION SCIENCES

BY
Seyed Ayat Hosseini

JANUARY 2014

Dissertation Committee:
Chair: Shin-ichi Tanaka (The University of Tokyo)
Yuki Hirose (The University of Tokyo)
Shuichi Yatabe (The University of Tokyo)
Yosuke Igarashi (Hiroshima University)
Satoko Yoshie (Tokyo University of Foreign Studies)

Abstract

This dissertation explores the phonological representation and the phonetic realization of prosodic prominence in Persian. It comprises two related parts: the first part addresses prosodic phrasing in Persian sentences, while the second part deals with phonetic correlates of prosodic prominence by reporting conducted production and perception experiments. The phonological part is carried out within the framework of Prosodic Phonology, and aims at determining the prosodic structure of Persian from foot level, up to utterance level. By adopting Optimality Theory, it tries to explain how morphosyntax-phonology interface constraints together with prosodic markedness constraints form the prosodic structure of the language. It begins with foot level and suggests that in languages like Persian which have one non-iterative weight-insensitive edgemost stress per word, a single foot which is edge-aligned with the minimal Phonological Word best explains the prosodic pattern at word level. This part also focuses on prosodic differences between lexical words and weak function words (clitics) and shows that any attempt to describe the prosodic structure of Persian without addressing this crucial difference, will not be able to provide explanation for a wide range of phenomena. This study suggests that proclitics and enclitics behave asymmetrically in Persian: enclitics prosodize as affixal clitics, while proclitics are free clitics. Next, it addresses the problem of weak function words which are not a part of their preceding or following XPs, and demonstrates how the phonological well-formedness constraints determine the direction of cliticization in these XP-external function words. It also deals with the issue of clitic clusters in Persian which was not explored in the previous works. Another contribution of this study is reclassification of so-called exceptionally initial-stressed words.

This dissertation also deals with the longstanding problem of Ezafe constructions and by reviewing previous proposals on the prosodic structure of these constructions, based on phonological evidence and phonetic observations suggests that each lexical word in an Ezafe construction maps onto a Phonological Phrase, and the Ezafe morpheme phrases with its preceding material to satisfy the phonological well-formedness constraint ONSET. The prosodic structure of XP-external clitics such as the Ezafe morpheme is explained by adopting a syntax-prosody interface constraint namely MAP-XP, that bans two sister XPs inside a single Phonological Phrase. This study proposes a ranking of OT constraints by which the prosodic structure of Ezafe constructions and other syntactic phrases such as DPs and VPs can be predicted and explained uniformly. It also proposes that the interaction between morphosyntax-phonology interface constraints and prosodic markedness constraints determine prosodic constituents of all levels and their heads, and other constraints require the heads of phonological phrases to be associated with audible

accents. The rightmost Phonological Phrase in an Intonational Phrase is the head. This head associates with an accent which is perceived more prominently than the other accents. One further issue explored here is the fact that in lexicalized Ezafe constructions and also in the ones containing given/old information, some words may appear without audible accent.

The phonetic difference between final and non-final accents is the subject of the second half of this dissertation. Previous research on Persian has shown that the main acoustic correlate of prosodic prominence is f0. This study reports production and perception experiment results conducted in order to answer the question whether final (nuclear) accents are perceived more prominently than the other ones only because they are not followed by any other accent, or because they are phonetically different from the non-final (pre-nuclear) accents. The results of production experiments reveal that nuclear accented syllables have a lower f0 range, but a longer duration in comparison with pre-nuclear accented ones. Other parameters such as overall intensity, spectral tilt and vowel quality do not differ significantly in the two types of accents. Perception experiments reveal that native listeners can indeed distinguish the two types of accents without having access to the portion of the utterance that follows the final accent. This proves that the two types of accents are phonetically different. Perception tests also show that the difference between the shapes of f0 curves in the two types of accents is the main acoustic parameter that helps the listeners distinguish them from each other. In pre-nuclear accented words, the f0 peak is at the right edge of the metrically strong syllable, and the curve has a rising slope at this point. In these syllables, the peak may even occur on the initial syllable of the following word. However, in the syllables associated with nuclear accents, the f0 peak is located inside the syllable, and the curve has a falling slope at the right edge of the syllable.

If the f0 at the right edge of a nuclear accented syllable is manipulated and raised so that the f0 peak is moved to the right edge, the native listeners will perceive the word containing this syllable as a pre-nuclear accented word. This study also shows that duration alone cannot cue the difference between the two types of accents. However, when accompanied by f0 changes, it can help the listeners distinguish the two accents more easily and more efficiently.

Acknowledgments

The completion of this dissertation would never have happened without the support and the contribution of many people. I would like to take this opportunity to express my sincere gratitude to some of them.

First and foremost, I would like to express my deepest respect and gratitude to my supervisor Prof. **Shin-ichi Tanaka**, for being a constant source of support, encouragement, and constructive advice. His exceptional dedication, generosity and caring have constantly extended far beyond the duty of an academic supervisor. With him I learned to avoid prejudice and stay unbiased with respect to linguistic theories. He created an excellent atmosphere in which all the students could share their knowledge and learn from each other in a friendly environment. I consider myself very fortunate to have such a wonderful supervisor.

I would also like to thank all the members of the dissertation committee, for taking the time to read the manuscript and for their helpful comments and precious advices. Prof. **Yuki Hirose** guided me through my experiments and statistical analyses. She kindly offered me a job in the Sound Lab as her research assistant for three years, where I learned a lot about laboratory phonetics and psycholinguistic experiments. Prof. **Shuichi Yatabe** helpfully criticized and challenged the syntactic analyses adopted in this dissertation, and helped me improve them. Prof. **Yosuke Igarashi** helped me modify and improve my interpretations of my experiment results, in our fruitful meetings and also by his valuable comments via email. And Prof. **Satoko Yoshie** generously shared with me her encyclopedic knowledge of Persian language, and opened my eyes to many aspects I had overlooked.

I would also like to thank those people who are not on my committee, but nevertheless helped me in one way or another. My special appreciation goes to Prof. **Junko Ito** for her inspiring phonology course in International Christian University (ICU), Tokyo, in spring 2010, where I came up with the basic idea of the analysis presented in chapter 3 of this dissertation. My sincere thanks are due to Prof. **Carina Jahani**, Prof. **Geoffrey Haig** and Prof. **Agnes Korn,** for their insightful comments and suggestions on one of my papers which contributed much to this dissertation. I am also grateful to all participants in regular meetings of Tokyo Circle of Phonologists (**TCP**) where I had the opportunity to present my work many times and benefit from their invaluable expertise. I also had many helpful comments from or interesting conversations with scholars and graduate students in all conferences and seminars I attended during past four years.

Although this dissertation is thoroughly devoted to prosody of Persian, I had the chance to pursue my other academic interests, Japanese language and second language education during my Ph.D. program in The University of Tokyo (*Todai*). I learned about methods of teaching Japanese as a foreign language in classes and seminars held by Prof. **Atsuko Kondoh**. Prof. **Takashi Nomura** deepened my insights to Japanese grammar, traditional Japanese studies (*kokugo-gaku*), and historical changes of Japanese language. In the classes held by Prof. **Izabelle Grenon**, I learned a lot about theories and research methodologies in the field of language education, and Prof. **Tom Gally** introduced me to second language writing pedagogy. I sincerely acknowledge all their teachings.

I would also like to acknowledge the financial support I received throughout my Ph.D. program. I greatly appreciate the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) for awarding me a scholarship to carry my academic life peacefully throughout my doctoral course. I would also like to thank the department of Language and Information Sciences for awarding me a six-month grant for research. Part of the work presented in chapter 4 of this dissertation was financially supported by National Institute for Japanese Language and Linguistics (NINJAL). I would like to thank all professors, post-docs and staff in NINJAL, especially Prof. **Haruo Kubozono** who has always been supportive and caring.

A special word of thanks is also due to Prof. **Akiko Katayama** for trusting me and allowing me to be a part of Komaba Writers' Studio (**KWS**) in *Todai*, where I worked as an academic writing tutor for three years and learned many things about academic writing and working with undergraduate students. I also thank the Department of English Language for having me as a teaching assistant from the very beginning of my doctoral program to the last days of it. I am also grateful to Mr. **Walter Arnold** the representative director in AIT Foreign Language Center in Tokyo where I worked as a part-time language teacher, translator and interpreter for four years.

I should take this opportunity to thank my previous linguistics teachers. I would like to express my sincere gratitude to my master's thesis supervisor in the University of Tehran, Prof. **Mahmood Bijankhan**. His encouragements and positive attitude was highly influential in drawing me to the field of phonology and phonetics. I am also thankful to my master's thesis advisor Prof. **Ali Darzi,** who familiarized me with argumentation in linguistics. I am deeply indebted to my first linguistics teacher Prof. **Reza Moghaddam-kia,** whose interesting classes attracted me to the field of linguistics. I cannot go without mentioning Prof. **Yosuke Momiyama** my supervisor in Nagoya University where I was an exchange undergraduate student in 2003-2004. He introduced me to Cognitive Linguistics, and encouraged me to start reading linguistics books and papers written in Japanese.

My heartfelt thanks go to my fellow graduate students from whom I learned many valuable things through our discussions, debates and chats, in seminars, meetings and over dinners. Without them my time in *Todai* would have been much poorer and less enjoyable. **Bum-Ki Son** helped me a lot when I first entered *Todai*. He taught me many things about Serial OT and syllable structure in Korean. He also showed me nice restaurants and diners around the campus. **Kyoko Yamaguchi** (Takano) and **Kohei Nishimura** were always helpful and supportive. I learned about Japanese morphology and prosodic phonology from them. **Yuki Asahi** was always there when I needed help. He helped me learn much about the rules and regulations in *Todai*. My conversations with Yuki always involved phonological phenomena in languages along the Silk Road, like reduplication and vowel harmony in Turkish, Persian and Mongolian. **Clemens Poppe** has always been a truly good friend to me. He helped me have access to some publications which I could not find anywhere. Clemens and I talked about dialects of Japan, Dutch phonology, music and football. I will never forget **Daiki Hashimoto** for his talent for linguistics and also for his wonderful sense of humor. I learned a lot about Government Phonology and Gunma prefecture from him. **Marco Fonseca**'s fascination with Usage-based Phonology and frequency-based approaches to linguistics intrigued me. I learned about the prosody of Japanese and various dialects of Chinese from **Xiaoyan Xie** and **Yang Yang**, and about diachronic changes in Slavic languages, especially about palatalization in Russian from **Naoya Watabe**. **Yu Tanaka**, **Yasushi Otaki** and **Ryota Horikawa** were not official members of our phonology group, but I had the opportunity to study with them and learn many things from them. I will always remember what Yasushi taught me about loanword phonology, gemination in Japanese, trains and baseball. And during the short period she studied with our group, **Maho Morimoto** kindly proofread two chapters of this dissertation for me. Thank you all for your friendship, support and for sharing your knowledge and thoughts with me.

My heartfelt thanks to my Iranian friends in Tokyo **Kaveh Maghsoudi**, **Mostafa Kamali-Tabrizi**, **Fatemeh Gharahkhani**, **Zeinab Sadoughi**, **Saeid-Reza Ettehadi**, **Leyla Ghiaee**, **Reza Amani-Eilanlou** and **Farzaneh Moradi** for being wonderful friends, for sharing with me their intuition as native speakers, and for helping me conduct the pilot version of my production and perception experiments as participants.

I am forever indebted to my parents, and my brother who encouraged and helped me at every stage of my personal and academic life, and longed to see this achievement come true. I would also like to thank my parents and brothers in law for all their support and encouragement. I am particularly grateful to my wife **Sarah** for her love, understanding, constant support and great patience at all times. This dissertation is dedicated to her.

Table of contents

Chapter 4 Phonological Phrase and above

Chapter 5 phonetic differences between pre-nuclear and nuclear accents

Key to abbreviations

ACC	accusative
COMP	comparative
EZ	Ezafe
IND	indefinite
INDI	indicative
PAST	past tense
PFV	perfective
PL	plural
POSS	possessive
PR	proper noun
SG	singular
fnc	function word
lex	lexical word
IPhrase	Intonational Phrase
PPhrase	Phonological Phrase
PWord	Phonological Word
υ	Utterance
ι	Intonational Phrase
φ	Phonological Phrase
ω	Phonological Word
π	Foot
σ	Syllable
μ	Mora

Chapter 1

This chapter will briefly introduce the goals of the present study and its significant findings. It will also provide information about the theoretical basis of the study, and the language under investigation in this dissertation.

1.1 Main Goals and major findings

This dissertation comprises two related parts. The first part addresses phonological characteristics of prosodic phenomena in Persian, while the second part deals with phonetic correlates of prosodic prominence by conducting production and perception experiments. The first phonological part is carried out within the framework of Prosodic Phonology, and aims at determining the prosodic structure of Persian from foot level, up to utterance level. By adopting Optimality Theory, this study tries to explain how morphosyntax-phonology interface constraints together with prosodic markedness constraints form the prosodic structure of the language. It begins from foot level and proposes a non-exhaustive foot structure in which a single foot is right-aligned with each lexical word. This study focuses on prosodic differences between lexical words and weak function words (proclitics and enclitics) and shows that proclitics and enclitics behave asymmetrically in this language. That is, enclitics prosodize as affixal clitics, while proclitics are free clitics. Next, it addresses the problem caused by certain weak function words which are not a part of their preceding or following XPs, and demonstrates that the phonological well-formedness constraints determine the direction of cliticization in these XP-external function words. The present study also deals with the longstanding problem of Ezafe constructions and by reviewing previous proposals on the prosodic structure of Ezafe constructions, based on phonological evidence and phonetic observations suggests that each lexical word present in an Ezafe construction, together with its following Ezafe morpheme maps on a Phonological Phrase. In the next step, this dissertation proposes a ranking of OT constraint by which the

prosodic structure of Ezafe constructions and other syntactic phrases such as DPs and VPs can be predicted and explained uniformly. This study also proposes that morphosyntax-phonology interface constraints and prosodic markedness constraints determine prosodic constituents of all levels and their heads, and other constraints require the heads of Phonological Phrases to be associated with audible accents. The rightmost Phonological Phrase in an Intonational Phrase which is the head of the Intonational Phrase, associates with an accent which is perceived more prominently than the other accents.

The phonetic differences between final and non-final accents are the subject of the phonetic part of this dissertation. This part reports production and perception experiments conducted in order to answer the question whether final (nuclear) accents are perceived more prominently only because they are not followed by any other accents, or because they are phonetically different from the non-final (pre-nuclear) accents. The findings of both production and perception experiments show that these two types of accent are indeed phonetically different. The shape of pitch curve on a syllable associated with a nuclear accent is different from that on a syllable associated with a pre-nuclear accent. Also the former syllables have a lower pitch range and longer duration than the latter syllables. The perception experiments reveal that these differences can help the listeners distinguish the two types of accents even without having access to the post-accentual portion of the utterance.

1.2 Outline of the dissertation

Chapter 1 gives an overall insight into the goals and findings of this dissertation, introduces the theoretical framework within which this study has been carried out and provides preliminary information about the language under study.

Chapter 2 introduces and reviews some of previously done studies on Persian prosody. This chapter consists of two parts. The first part reviews some major phonological studies on the

prosodic structure of Persian which are also addressed in this dissertation, and the second part, introduces some recent instrumental studies on the phonetic characteristics of prosodic prominence in Persian.

Chapter 3 starts with foot structure and provides a theoretical argumentation in favor of a non-exhaustive foot structure. It also addresses the asymmetrical behavior of proclitics and enclitics and explains this difference by proposing a ranking of OT constraints. This chapter also deals with the direction of cliticization in XP-external function words and studies the case of clitic clusters in Persian. Finally, this chapter tries to shed more light on the class of words known as 'exceptionally initial stressed' words and explains them in terms of lexicalization and grammaticalization.

Chapter 4 deals with higher level prosodic categories namely Phonological Phrase and intonation phrase. A considerable portion of this chapter is devoted to the prosodic structure of Ezafe constructions in Persian. It introduces these constructions and the existing literature on their syntactic and prosodic structures, and offers a new proposal for their prosody which is consistent with both theoretical principles of Prosodic Phonology and empirical observations of the attested data. This chapter also proposes a ranking of OT universal constraints by which the prosodic structure of Ezafe constructions and all the other types of syntactic phrases can be explained.

Chapter 5 tries to provide an answer to the question raised at the end of chapter 4, that whether final and non-final accents differ phonetically or not. First, the production experiments are reported in this chapter, and then two perception tests are detailed. The conclusion drawn at the end of this chapter is that the two types of accents are phonetically different, and that native listeners actually use these phonetic cues to distinguish the two types of accents.

Chapter 6 contains a short review of the dissertation, conclusions of both phonological and phonetic parts and final remarks.

1.3 Theoretical framework of the study

The phonological theory applied in this study is known as Prosodic Phonology, Prosodic Structure Theory, Phrasal Phonology or Prosodic Hierarchy Theory, and is based on the work of such researchers as Selkirk (1978, 1981, 1995, 2011), Nespor and Vogel (1986), Selkirk and Tateishi (1988, 1991), Selkirk and Shen (1990), Prince and Smolensky (1993) McCarthy and Prince (1993b) Truckenbrodt (1995, 1999), Ito and Mester (2009, 2012) among others. Prosodic Phonology is a theory of the way in which the flow of speech is organized into a finite set of phonological units. It is also, however, a theory of interactions between phonology and the other components of the grammar. The main idea of Prosodic Phonology, as shown in (1), is that utterances are phrased in prosodic constituents which are themselves organized into a hierarchy. That is, phonological constituents occur in clusterings, and these clusterings occur in still larger clusterings and so on, up to the level of the whole utterance.

(1)	Utterance	υ
		|
	Intonational Phrase (IPhrase)	ι
		|
	Phonological Phrase (PPhrase)	φ
		|
	Phonological Word (PWord)	ω
		|
	Foot	π
		|
	Syllable	σ
		|
	Mora	μ

In the last three decades there has been a debate around a fundamental question in the literature: What is the nature of the linguistic representation in terms of which domain-sensitive phenomena of sentence phonology and phonetics are defined. In other words, do phonological rules/constraints take morpho-syntactic categories as their domains, or are there purely phonological categories that function as domains for these phenomena? Prosodic Phonology takes an indirect approach to the relation between morpho-syntax and phonology. As discussed in Selkirk (2011), Prosodic Phonology proposes a prosodic structure representation of phonological domains, a representation which is independent of syntactic constituency but related to it by a module of syntactic-prosodic constituency correspondence constraints.

In Prosodic Phonology it is assumed that morpho-syntactic words (X^0-level elements) are mapped onto PWords, syntactic XPs (maximal projections of lexical elements) are mapped onto PPhrases and syntactic clauses are mapped onto IPhrases.

The present study also uses Optimality Theory (Prince and Smolensky, 1993) as a tool to explain how attested output structures in a language arise as a result of the language-specific ranking of universal constraints on the input structures. Following McCarthy (2003), this study assumes that OT constraints are all categorical, because in gradient constraints, calculating the number of violations might be an ad hoc issue, and moreover, accepting gradient constraints often has undesirable typological consequences as is discussed in detail by McCarthy (2003).

So far in the literature, various constraints have been proposed that require mapping the morpho-syntactic categories onto phonological categories, among which the most widely adopted constraints are ALIGNMENT constraints that require syntactic categories to be edge-aligned (right or left) with phonological categories, or phonological categories to be edge-aligned with syntactic categories. ALIGNMENT constraints are based on Selkirk's theory of edge alignment (Selkirk, 1986) which was later formalized to the format of Generalized Alignment in McCarthy and Prince (1993b) in Optimality Theory. Selkirk (2009, 2011)

proposed Match Theory according to which, morpho-syntactic constituents are matched by corresponding prosodic constituents. Matching in this sense can be interpreted as simultaneous alignment of left and right edges. The constraint MATCHWORD requires all the words in syntactic representation to be matched by a PWord in the phonological representation. MATCHPHRASE or MATCH-XP requires all syntactic XPs to be match by a PPhrase, and MATCHCLAUSE requires all clauses in syntactic constituent structure to be matched by a corresponding IPhrase. Other syntax-prosody interface constraints have also been proposed in the framework of Prosodic Phonology. For instance, Truckenbrodt (1995) proposes WRAP-XP according to which, each XP must be contained in a PPhrase.

In Prosodic phonology, it is assumed that Constraints relating syntactic and prosodic categories apply to lexical syntactic elements and their projections, but not to functional elements and their projections, or to empty syntactic elements and their projections. This is called Lexical Category Condition (LCC) (Truckenbrodt, 1999).

In the earlier versions of Prosodic Phonology, Selkirk (1981, 1984) and Nespor and Vogel (1986) proposed that phonological categories are strictly layered. This proposal is known as Strict Layer Hypothesis and is composed of four universally inviolable constraints which are defined in (2) taken from Selkirk (1995), where C^n = some prosodic category.

(2) i. Layeredness No C^i dominates a C^j, j > i, e.g. "No σ dominates a π."
ii. Headedness Any C^i must dominate a C^{i-1}, e.g. "A ω must dominate a π."
iii. Exhaustivity No C^i immediately dominates C^j, j< i-1, e.g. "No ω dominates a σ.'
iv. Nonrecursivity No C^i dominates C^j, j = i, e.g. "No π dominates a π."

After Strict Layer Hypothesis was proposed, the inviolability of the two constraints *Exhaustivity* and *Nonrecursivity* has been challenged frequently and recursive and non-exhaustive prosodic structures have been shown to exist almost in all levels of hierarchy in different languages by researchers such as Ladd (1986, 1992), Inkelas (1989), Hayes (1991), McCarthy and Prince (1991, 1993ab), Ito and Mester (1992, 2008, 2009, 2010, 2012) and Féry (2010). In recent studies within the framework of Prosodic Phonology, EXHAUSTIVITY

and NONRECURSIVITY are assumed to be violable constraints which based on their rankings with other constraints, may ban or allow nonexhaustive and recursive structures in specific languages.

This study also makes a distinction between metrical strength and accent. Metrically strong elements are heads of phonological categories. According to *headedness* constraint, each prosodic constituent must have a head which must be an immediately lower-level constitute. Heads of prosodic constituents are metrically strong positions. Inside an iambic foot, the syllable on the right is the head of the foot and is a metrically strong constituent. Inside a PWord, the head (mostly the leftmost or the rightmost) foot is metrically strong. Similarly, inside a PPhrase, the head PWord, and inside an IPhrase, the head PPhrase are strong positions. Being located at a strong position does not guarantee a constituent to have an audible prominence. Metrically strong positions serve as docking sites to which audible accents may or may not be associated. Metrical strength is an abstract notion, while accent is regarded to be the actual audible prominence in the utterances. Accents may be associated with different levels of hierarchy in different languages. For instance, in one language audible accents may be associated with PPhrases, and in the other with PWords. However, since accent-bearing units are syllables, when a PPhrase is associated with an accent, the head PWords inside the PPhrase takes the accent, and inside the PWord, the accent associates with the strong syllable of the head foot. The notion 'stress' in this study is regarded to be the metrical strength at the PWord level, thus a 'stressed syllable' is metrically the strongest syllable of a PWord.

Since this dissertation deals only with simple sentences in which the utterance corresponds to an IPhrase, the terms 'IPhrase' and 'Utterance' practically bear the same meaning. It should also be noted that the primary subject of this study are 'informationally neutral' or 'all-new' sentences, however chapter 4 will explore some utterances containing focused elements, or elements conveying old/given information.

1.4 Persian Language

Persian, known by its speakers in Iran as *Fârsi* or *Pârsi*, is a Southwestern Iranian language, belonging to the Indo-Iranian branch of the Eastern Indo-European languages, and is the official language of Iran. Persian is also used in Afghanistan (officially known as Dari) and Tajikistan (officially known as Tajiki) and other countries which historically came under Persian influence. The Persian language spoken in Iran is known to have two different registers: a formal (non-colloquial) register, and a colloquial one. The variety studied in this dissertation is the colloquial register of the language, spoken by educated individuals in Tehran.

Modern colloquial Tehrani Persian, hereafter Persian, has six vowels and twenty-three consonants (Mahootian, 1997: 286-287) which are shown in figure 1 and table 1 below.

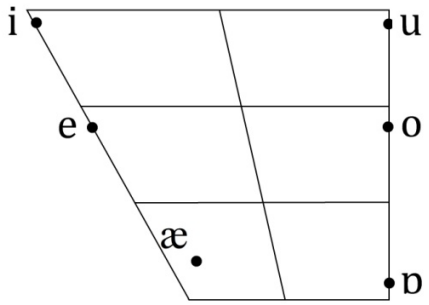


Figure 1 Persian vowels and their place of articulation

In this dissertation, the symbols /a/ and /â/ have been used instead of the IPA symbols /æ/ and /ɒ/ respectively. Persian also has one diphthong /ou/ which has been transcribed as /ow/ in this dissertation. Figure 1 and table 1 have been made based on Windfuhr (1987), Mahootian (1997) and Samareh (1999).

Table 1 Persian consonantal phonemes, with their place and manner of articulation

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Palatal	Velar	Uvular	Glottal
Plosive	p b		t d				k g	ɢ	ʔ
Nasal	m			n					
Trill				r					
Fricative		f v		s z	ʃ ʒ		x		h
Affricate					ʧ ʤ				
Approximant									
Lateral				l		j			

In the present study, the phonemes /ʃ/, /ʧ/, /ʤ/ and /j/ have been shown by /š/, /č/, /j/ and /y/ respectively, and the phoneme /ɢ/ and its fricative allophone [ʁ] have been shown by /q/.

Syllables in Persian must have a vowel in the nucleus position and a glide or a consonant in the onset positions. They also can optionally have up to two consonants or glides in the coda position. Samareh (1999) and Haghshenas (2001) discuss that seemingly word-initial vowels are actually preceded by a glottal stop. Jahani (2005) argues that word-initial glottal stops never contrast with ∅, thus they are not present at phonological level. Since word-initial glottal stops are always predictable in Persian, in this dissertation, following the tradition, they will not be shown in phonemic transcriptions, unless where syllable structure is concerned.

In Persian simple sentences with unmarked word order, the verb phrase tends to appear at the end of the sentence as can be seen in (3). That is why scholars such as Greenberg (1963), Dabir-Moghaddam (1982) and Karimi (1989) classify Persian as an SOV language.

(3) *mard* *nân-râ* *xord*
man bread-ACC eat-PAST
'The man ate the bread.'

However, Persian also displays some characteristics of SVO languages as well. For example, in the unmarked word order, phrasal complements of the verb follow the verb as is shown in (4).

(4) *mardom fekr mi-kard-and oranus setâre ast*

people think INDI-doPAST-3PL Uranus star is

'People used to think that Uranus is a star.'

Persian also uses prepositions rather than postpositions, a characteristic which is common to SVO languages according to Greenberg (1963).

As mentioned above, syllables in Persian can have at most one segment in their onset position, while there may be up to two segments in coda. Word-initial consonant clusters in loanwords are broken by inserting the epenthetic vowel /e/.

In nouns, adjectives, verb roots and most adverbs, the rightmost syllable is metrically strong (stressed). Every lexical word in Persian has one and only one stressed syllable and the place of stress does not depend on the weight or the number of syllables as is shown in (5) where prominent syllables are shown by an acute accent (′) on the nucleus, and syllable boundaries by a dot (.). As can be seen, even in long words containing heavy syllables, only the rightmost syllable is stressed.

(5) i. *pe.dár* 'father'

ii. *vâ.qeʔ.ge.râ.yâ.né* 'realistically'

iii. *sang.ne.veš.té* 'stone inscription'

Compound nouns and adjectives as well, have only one stress on the last syllable of the compound, regardless of the word length.

(6) i. *šotor-mórq* camel-bird 'ostrich'

ii. *now-javấn* new-young 'teenager'

iii. *pedar-bozórg* father-great 'grandfather'

Derivational affixation is very similar to compounding, in that the stress falls on the last syllable of the whole affixed word.

(7) i. *pe.dár* 'father'
ii. *pedar-âné* 'fatherly'
iii. *pedar-âne-gí* 'fatherliness'

Inflectional suffixes on the other hand, fall out of the stress domain and are always stressless.

(8) i. *pedár-am* ii. *mâdár-i* iii. *búd-and*
father-POSS1SG mother-IND isPAST-3PL
'my father' 'a mother 'they were'

Some semantically inflectional suffixes behave like derivational ones by attracting the stress. These suffixes are the plural markers *-hâ* and *-ân* on nouns and adjectives, the comparative and superlative markers *-tar* and *-tarin* on adjectives, and the perfective marker *-e*, on verbs.

(9) i. *pedar-hấ* father-PL 'fathers'
ii. *bad-tár* bad-COMP 'worse'
iii. *mord-é* diePAST-PFV 'dead'

In conjugated verbs, the stress tends to occur near the left edge, but as argued in Kahnemuyipour (2003), conjugated verbs should be regarded Phonological Phrases composed of one or more Phonological Words as shown in (10). The head of a PPhrase, is the leftmost PWord in it, thus the indicative marker *mi-* being the leftmost PWord is perceived more prominently.

(10) *mí-xând-am*
INDI-readPAST-1SG
'I was reading.'

Hyman (2006, 2009) in his Canonical Typology of prosodic systems outlines the properties of canonical tone and those of canonical stress. By adopting Prague School terms, he argues that the canonical function of tone is to distinguish morphemes, hence *paradigmatic*, while the canonical function of stress is *syntagmatic*: it should unambiguously identify and mark off major category words within utterances. Hyman (2006, 2009) postulates nine properties for canonical stress. These properties together with the properties of stress in Persian are given in table 1. As can be seen, Persian word stress meets all the criteria of canonical stress.

Table 2 Criteria developed in Hyman (2006, 2009) for canonical stress, and the properties of word stress in Persian

	Criterion	**Canonical stress**	**Persian**
a.	Obligatory	All words have a primary stress.	All lexical words are stressed.
b.	Culminative	No word should have more than one primary stress.	There is only one stress per lexical word.
c.	Predictable	Stress should be predictable by rule.	Word stress is always rightmost.
d.	Autonomous	Stress should be predictable without grammatical information.	Word stress is always predictable (rightmost).
f.	Demarcative	Stress should be calculated from the word edge.	Word stress is always rightmost.
g.	Edge-adjacent	Stress should be edge-adjacent (initial, final).	Word stress is always rightmost (final).
h.	Non-moraic	Stress should be weight-insensitive.	Stress is weight-insensitive.
i.	Privative	There should be no secondary stresses.	There are no secondary stresses at word level.
j.	Audible	There should be phonetic cues of the primary stress.	Stressed syllables associate with audible accents.

Based on the criteria developed in Hyman (2006, 2009) Persian can be considered as a prototypical ‘stress language’.

Chapter 2 Previous studies

Since this dissertation deals with both phonological and phonetic characteristics of prosodic phenomena, this chapter will review some of both phonological and phonetic studies on Persian prosody. Phonological studies and phonetic studies will be introduced in two separate sections.

2.1 Phonological studies

According to Windfuhr (1979), the first European study to deal with prosodic prominence in Persian is Chodzko (1852) whose description of the prosodic pattern in Persian nouns and adjectives is still valid. Ferguson (1957) is the first linguistic study devoted thoroughly to stress in Persian. This study describes the prominence pattern at word level in nouns, adjectives, adverbs, vocatives, infinitives and conjugated verbs, and provides a list of exceptionally non-final stressed particles and adverbs. A substantial amount of work has been done on Persian prosody in the past decades among which one may name Foadi (1933), Nye (1955), Hodge (1957), Lambton (1957), Jazayeri and Paper (1961), Atai (1964), Scott (1964), Yarmohammadi (1964), Bateni (1969), Towhidi (1974), Lazard (1992), Sepanta (1995), Same'i (1996), Amini (1997), Mahootian (1997), Hayati (1998), Eslami (2000, 2003), Vahidian-Kamyar (2000), Eslami and Bijankhan (2002), Mahjani (2003), Kahnemuyipour (2003, 2004), Parmoon (2006), Arbisi-Kelm (2007), Esposito and Barjam (2007), Sadat-Tehrani (2007), Scarborough (2007) and Hosseini (2012b, 2014a). In This section we will introduce and briefly review five major works on the prosodic structure of Persian which will be addressed directly in this dissertation. These works are Amini (1997), Kahnemuyipour (2003), Eslami (2000), Mahjani (2003) and Sadat-Tehrani (2007).

2.1.1 Amini (1997)

This study is a metrical approach to Persian stress and is one of the first attempts to explain the prosodic structure of Persian by Optimality Theory, carried out only a few years after the theory was proposed by Prince and Smolensky (1993). Amini's account accepts the strong version of Strict Layer Hypothesis, and does not allow non-exhaustive and recursive prosodic constituents. An immediate consequence of this is the fact that all syllables are parsed into iambic feet in her analysis. She does not clarify her position about secondary stress at word level, but her analysis obviously predicts secondary stresses on alternating syllables preceding the main stress. She explains the lack of stress on enclitics using the Extrametricality Theory (Hayes, 1982 and 1995), although her OT account does not explain why only one group of suffixed words are subject to extrametricality, and other types of suffixes can bear stress without being extrametrical. Amini (1997) proposes two different constraints that require the alignment of a PWord with its head. These constraints which are given in (1) and (2) are sensitive to lexical categories.

(1) Align (PWord, R; Head, R): Align the right edge of the Prosodic Word with the right edge of the head of the Prosodic Word.

(2) Align (PWord (=verb), L; Head, L): Align the left edge of the Prosodic Word (verb) with the left edge of the head of the Prosodic Word.

These two constraints explain the final-stress in nouns and adjectives, and the tendency of conjugated verbs to be initial-stressed.

2.1.2 Kahnemuyipour (2003)

Kahnemuyipour (2003) is a comprehensive study of prosodic structure in Persian carried out within the framework of Phrasal (Prosodic) Phonology. One of the significances of this study is the solution it provides to the longstanding problem of the asymmetrical behavior

of stress in nouns, adjectives, adverbs and verb roots on one hand, and conjugated verbs on the other. The stress is on the rightmost syllable in the former, and tends to be on the leftmost prefix in the latter. This had led most scholars to treat verbs differently and propose verb-specific stress assignment rules (cf. Amini (1997)). In Kahnemuyipour's analysis, for any prosodic category, the sentence is exhaustively parsed into a sequence of that category, and recursive prosodic structures are not allowed. All morpho-syntactic words (including suffixes and stressless clitics) are mapped onto PWords. The rightmost syllable in a PWord takes the PWord-level stress. This explains the final stress in nouns, adjectives, adverbs and verb roots. Syntactic maximal projections (XPs) are mapped onto PPhrases, and inside a PPhrase, the leftmost PWord takes the PPhrase-level stress. After PPhrase-level stresses are assigned, PWord-level stresses are deleted by a 'leveling rule', thus PPhrase-level stresses, unlike PWord-level stresses, are audible. This proposal explains the lack of stress on enclitics in Persian as shown in the examples in (3) and (4) based on Kahnemuyipour (2003), where PWord-level stresses and PPhrase-level stresses are shown by (X_ω) and (X_φ) respectively. The clitic and the host are PWords, but the clitics do not take the PPhrase-level stress, because they happen to follow their hosts, and the PPhrase-level stress rule puts the stress on the host.

(3) X_φ

(X_ω $X_\omega)_\varphi$

(*ketâb*)$_\omega$ (*i*)$_\omega$

book IND

'a book'

(4) X_φ

(X_ω $X_\omega)_\varphi$

(*ketâb*)$_\omega$ (*am*)$_\omega$

book POSS1SG

'my book'

This proposal also explains the left-most stress pattern in conjugated verbs. In a conjugated verb, each morpho-syntactic word maps onto a PWord, and the leftmost PWord takes the PPhrase-level stress. This is shown in (5) and (6) based on Kahnemuyipour (2003).

(5) X_{φ}

(X_{ω} $X_{\omega})_{\varphi}$

(mí)$_{\omega}$ (raft)$_{\omega}$

INDI go$_{\text{PAST3SG}}$

'She/he was going.'

(6) X_{φ}

(X_{ω} $X_{\omega})_{\varphi}$

(pás)$_{\omega}$ (dâd)$_{\omega}$

back give$_{\text{PAST3SG}}$

'She/he gave [it] back.'

In this analysis, syntactic clauses are mapped onto IPhrases and the IPhrase-level stress rule puts the IPhrase stress on the rightmost PPhrase inside it. Both PPhrase-level stresses and IPrase-level stresses are audible, but the latter is perceived more prominently. Kahnemuyipour takes the IPhrase-level stress as the primary stress and the PPhrase-level stress as the secondary stress in Persian utterances, although this language does not have secondary stresses at word level.

Kahnemuyipour (2003) based on Ghomeshi (1996) and Kahnemuyipour (2001) takes the nouns and adjectives in an Ezafe construction, to be non-projecting base-generated X^0 elements, and proposes that the entire Ezafe construction is mapped onto a single PWord. This seems to be consistent with the fact that the final syllable of the final word in an Ezafe construction is perceived more prominently than the other syllables. This is shown in an example in (7).

(7) X_{ω}

($)_{\omega}$

sag-e *siyâh-e* *gondé*

dog-EZ black-EZ big

'big black dog'

The present study accepts and adopts the fundamental idea in Kahnemuyipour (2003) that in Persian, PWords and IPhrases are right-headed, while PPhrases are left-headed. However, some of the proposals in Kahnemuyipour (2003) run into theoretical and empirical problems. This dissertation will discuss some of these problems in detail, and provide novel solutions for them. Following is a list of these problems addressed in chapter 3 and chapter

4:

i. Neglecting the prosodic difference between lexical words (nouns, adjectives, adverbs and verb roots) and weak function words (stressless clitics).

ii. Ignoring the existence of proclitics (weak function words that prosodically phrase with their following material) in Persian.

iii. Existence of leveling rule. Rightmost syllables of all PWords are assigned with stress, but these stresses are deleted after a higher-level stress is assigned.

iv. Treating the entire Ezafe constructions as a single PWord, which is in contrast with the fact that each lexical word in an Ezafe construction has an audible prominence.

2.1.3 Phonological studies within AM

Some of the works on the prosodic structure of Persian have been conducted within the framework of Autosegmental-Metrical Phonology. They are mainly based on studies such as Pierrehumbert (1980), Beckman and Pierrehumbert (1986), Pierrehumbert and Beckman (1988) and Ladd (1996), and use ToBI[1] system for transcribing the intonation and prosodic structure of Persian. The approach taken in these studies and some of the fundamental assumptions they make, differ from the ones in the theory adopted in this study. Nevertheless, some of them will be reviewed here briefly, since they address issues that Prosodic Phonology does not deal with, and also because some of the findings in these studies provide support for a few of the proposals made in the present study.

[1] Tones and break indices

Eslami (2000)

This work which is written in Persian is the first attempt to study the intonation of Persian within the theoretical framework of Autosegmental-Metrical Phonology and is very similar in methods of analysis to Pierrehumbert's account of English intonation system. It proposes three levels of prosodic category for Persian, namely, Accentual Phrase, Intermediate Phrase and Intonational Phrase. It also posits four pitch accents (H*, L*, L+H*, and L*+H), two phrase accents (L- and H-) marking the right edge of Intermediate Phrases, and two boundary tones (L% and H%) for Persian intonation system. Each pitch accent, phrase accent and boundary tone has a distinct meaning or pragmatic function. These meanings and functions are given in (8).

(8)	H*	New information
	L*	Given information
	L+H*	Contrast
	L*+H	Doubt
	L-	Complete message
	H-	Incomplete message
	L%	Statement
	H%	Interrogative

Eslami (2000) considers all possible combinations of these pitch accents, phrase accents and boundary tones and suggests a distinct pragmatic usage for each of the combinations.

Eslami (2000) proposes Head Avoidance Principle to locate the most prominent element in a syntactic phrase. According to this principle, the farthest word from the head of a syntactic phrase is prosodically the most prominent element of the phrase. This holds for both phrases (XPs) and sentences (CP/IPs). This principle is very similar to the stress rule proposed in Cinque (1993) according to which phrasal stress is assigned to the most deeply embedded element. The main problem of this analysis is the 'direct' approach it takes to the syntax-prosody relation, i.e., stress assignment rule takes a syntactic category as its domain.

Selkirk (2011) points out theoretical and empirical problems of such direct approaches and shows that domain-sensitive phenomena such as stress/accent assignment are only sensitive to phonological categories that are independent of syntactic constituency but related to it by a module of syntactic-prosodic constituency correspondence constraints. Moreover, it seems that Head Avoidance does not always make the right prediction in Persian, and some of the examples in Eslami (2000) are in contrast with others' works and native speakers' intuition. For example, in a simple sentence composed of a subject and a transitive verb, Eslami (2000:144) predicts the sentential stress to be on the subject because the inflection of the verb is the syntactic head of the sentence, thus the stress is assigned to the farthest element from this head. This is shown in (9) where an asterisk (*) marks the most prominent syllable.

(9) a.*li did
PR seePAST1SG
'Ali saw.'

According to Kahnemuyipour (2003), Sadat-Tehrani (2007) and also intuitively, the transitive verb is prosodically the most prominent element in this sentence.

Mahjani (2003)

This Master's thesis supervised by Robert Ladd is a description of Persian intonation system within the framework of Autosegmental-Metrical Phonology. Similar to Eslami (2000), this work postulates three levels, namely, Accentual Phrase, Intermediate Phrase and Intonational Phrase for prosodic structure in Persian. However, unlike Eslami (2000) it suggests only one pitch accent which is L+H* for most Accentual Phrases, and H* for initially-stressed words. Therefore, sentence types are mostly distinguished by their boundary tones. His analysis contains two Intermediate Phrase boundary tones L- and H-, and two Intonational Phrase boundary tones L% and H%. The account in Mahjani (2003) is

simpler than the one in Eslami (2000), but the relation between syntactic structure and prosodic structure is less clear in it.

Sadat-Tehrani (2007)

This dissertation is a detailed investigation of the phonology and phonetics of the intonation system of Persian, carried out within the framework of Autosegmental-Metrical Theory. This study assumes only two levels for the prosodic structure namely Accentual Phrase and Intonational Phrase, and only one pitch accent (L+H*) (with the allomorph H* for monosyllabic words). The final Accentual Phrase in an Intonational Phrase is assigned with a Nuclear Pitch Accent (NPA) and a low boundary tone (l), while other (pre-nuclear) Accentual Phrases are assigned with a non-nuclear pitch accent, and a high boundary tone (h). Statements have a low Intonational Phrase boundary tone, while interrogatives are marked by a high Intonational Phrase boundary tone on their right edge.

Lexical words together with their enclitics form Accentual Phrases, and pitch accents are associated with the metrically strong (stressed) syllables of lexical words. Accentual Phrases together form the Intonational Phrase which is the domain in which downstep occurs.

A substantial portion of this dissertation is devoted to determining the element which takes the NPA in a sentence, however no general rule or principle is provided for this purpose, rather, the place of NPA is identified by separate rules such as the ones given in (10).

(10) i. The NPA in affirmative copular constructions is on the complement, i.e. the element between the subject and the copula.

ii. In intransitive SV sentences, unergatives are always accented on the verb, whereas unaccusatives are accented either on the verb, or on the final element of the subject.

iii. In transitives, NPA goes on the verb, if the direct object is specific, and on the direct object if it is non-specific.

iv. Arguments attract the NPA, but adjuncts usually do not.

The crucial problem with Sadat-Tehrani (2007) is that the relation between syntactic and prosodic phrases is not clear in this work. For instance, each lexical word and its enclitics maps onto a Accentual Phrase, and the entire sentence is mapped onto an Intonational Phrase, however the phrase-level syntactic structure has no role in this account.

2.2 Phonetic studies

The phonetic characteristics of prosodic prominence in Persian had been the subject of controversy for many decades. Different parameters have been suggested to be the acoustic and perceptual correlates of prosodic prominence. The following quotations exemplify some of these suggestions. Fergusson (1957: 124) states that: “Identification of the exact nature of the prominence attributed to stress must await experimental investigation, but the auditory impression is that of relative loudness, so-called 'expiratory' stress”. Haghshenas (1977: 124) argues that: “Unlike English, in Persian stress, pitch plays the main role and the other correlates are subsidiary factors”. Sepanta (1977: 7－8) based on his instrumental studies draws a similar conclusions: “The decisive factor in Persian stress is pitch. Stressed syllables sometimes become longer, but this is an effect of stress, not its cause. Intensity increases only slightly on stressed syllables which is not decisive at all”. Lazad (1992: 48) expresses his different idea by stating that: “Stress in Persian is realized by a simultaneous increase in intensity and pitch”. Finally, in her widely referred book Mahootian (1997: 310)

maintains that stress in Persian involves both intensity and pitch. However, it is unclear whether pitch or intensity is the primary feature of stress.

However, in recent years, more precise instrumental researches both inside and outside Iran have shown that as Haghshenas (1977) and Sepanta (1977) have suggested, pitch (f0 rise) is the main correlate of acoustic prominence in Persian and other parameters do not change significantly on prominent syllables, and pitch alone can cue the prosodic prominence. By recent developments in easily accessible phonetic softwares such as Praat, laboratorial studies are spreading rapidly. Following sections introduce and review some of the recent works in this field.

2.2.1 Abolhasanizadeh et al. (2012)

This study raises two questions and tries to provide answers to those questions by conducting production and perception experiments. The first question is whether any phonetic cues other than f0 can signal the prosodic prominence in Persian. In order to answer this question the minimal pair in (11) and (12) are used as experiment material, in a production test.

(11) *un* *tâbéš-e*
that light- be3SG
'That is light.'

(12) *un* *tấb-eš-e*
that swing-POSS3SG-be3SG
'That is her/his swing.'

In (11), *tâbeš* is a noun, while in (12) the noun *tâb* together with the stressless enclitic *eš* form a Clitic Group (CG). In (11), the second syllable and in (12) the first syllable are prominent. The authors compare f0, duration, intensity and spectral measures in 'stressed' and 'stressless' syllables to determine the significance of each parameter. They report very small phonetic differences between the members of this minimal pairs outside the more evident f0 differences and attribute these to side effects of pitch accent placement.

The second question raised in this paper is whether post-focal words undergo complete deaccentation or not. The authors use a version of (11) and (12) in which the first word (*un*) has distinctive (narrow) focus on it, and measure the acoustic parameters in post-focal stressed syllables. They also conduct a perception test in which participants should distinguish the noun in (11) and the CG in (12) in a post-focal environment. Both production and perception tests reveal that there is only post-focal compression in Persian and complete deaccentation (total neutralization) does not occur. Their final conclusion is that Persian prominences are pitch accents, and words are not deaccented when the pitch range is reduced after the focus.

2.2.2 Sadeghi (2012)

This perception study is the continuation of the production investigations in Sadeghi (2011), which focuses on the phonetic correlates of prosodic prominence in Persian. This work is very similar in methods of experiment and analysis to Sluijter (1995), Sluijter and van Heuven (1996ab) and Sluijter et al. (1997). It analyzes and compares phonetic parameters such as f0, duration, overall intensity, and spectral balance in accented and accentless syllables, and determines the effects of each parameter on the perception of prosodic prominence. The results of this experiment and the production tests in Sadeghi (2011) show that f0 is the main acoustic and perceptual correlate of prosodic prominence in Persian. The other correlates may increase the reliability of the responses when they accompany f0 changes. However, in the absence of any f0 change, the listeners tend to rely on duration increase as a cue to prominence. Duration has a much stronger role in signaling prominence in comparison with intensity or spectral balance.

2.2.3 Taheri-Ardali and Xu (2012)

This study raises the question whether the main phonetic correlate of focus in Persian is on pre-focus, on-focus or post-focus elements. To answer this question, the authors investigate the phonetic realization of prosodic focus in Persian. The findings of this study reveal that focus dramatically changes the three regions. f0 and duration significantly increase in on-focus words without any significant change in intensity. Compared to their counterparts, pre-focus elements show weaker intensity but no change in duration and mean f0. Finally post-focus words show significant lowering of f0 and decrease of intensity. The duration of post-focus words remains intact. The authors conclude based on the results of their experiment that, f0 and duration are the main acoustic correlates of prosodic focus in Persian, and this language like English and Mandarin, falls into the category of PFC (Post-Focus Compression) languages.

Chapter 3 Phonological Word and below

3.1 Introduction

This chapter aims at determining the foot and Phonological Word structure in Persian. By adopting Optimality Theory we will illustrate how morphosyntax-prosody interface constraints together with prosodic markedness constraints determine the prosodic structure in Phonological Word level and below it. The core of our argument will be the fundamental prosodic distinction between lexical and function words set forth in Selkirk (1995), an issue which has often been disregarded in the literature on Persian. We will particularly address the prosodization of clitics and the direction in which they attach to their hosts. Finally, we will try to establish some regularities in the so called 'exceptionally stressed' words in the previous studies.

3.2 Foot structure

As described in chapter 1, Persian has an edge-oriented quantity insensitive stress system in which stress is always on the rightmost syllable in nouns, adjectives, verb roots and most adverbs, regardless of word length or syllable weight. Stress is also non-iterative in Persian with only one primary stressed syllable on the right edge and no secondary stresses at word level. Therefore, any analysis that postulates multiple feet for a Persian word encounters empirical problems, because such structure as shown in (1) will wrongly predict secondary stresses on alternating non-final syllables.

(1) (σ σ̀) (σ σ̀) (σ σ́)

Prince (1976) for the first time proposes unbounded feet (hence the name given to the type of stress system) for systems in which there are no limits on the distance between stresses or between the stress and word boundary. An unbounded foot may include an unlimited

number of weak syllables provided that only the head (rightmost or leftmost syllable) is strong (Hayes, 1995). An unbounded foot is shown in (2) for a five syllable final-stressed word.

(2) (σ σ σ σ σ́)

On the contrary, Prince (1985) suggests that metrical tree theory can be further improved by eliminating the unbounded foot, or as he puts it "by trimming the fat". In this view, unboundedness has an edge-seeking function, but this function is present for other devices in the theory, and the work performed by unbounded metrical constituents can be achieved with binary feet and stray adjunction. By accepting unfooted syllables instead of unbounded feet, one can explain non-iterative systems as well as phenomena such as extrametricality. Walker (1997) argues that while certain non-stress-related phenomena, such as reduplication, tone patterns, language games, expletive insertion, and word-minimality requirements provide independent support for bounded (binary) foot constituency, no such evidence has ever been found to support an unbounded foot as a prosodic constituent.

Hayes (1995: 299) reports that unbounded feet of any type face a difficulty which is the apparent absence of extrametricality, since according to Wheeler (1979) extrametricality of an unbounded foot is unattested. De lacy (1997: 21) argues that unbounded feet can easily be replaced by binary feet, resulting in the simplification of foot inventory. Thus, in his approach, even in 'quantity-insensitive' languages, feet are either mono- or di-syllabic. Hyde (2001: 53－58) also argues against unbounded feet and proposes FootCap Condition under which forms containing feet larger than two syllables cannot be considered as output candidates. In his approach, unbounded and ternary feet are not merely restricted; they are banned by the grammar. Hyde (2001) and Baković (2004) analyze the stress system of Uzbek by postulating a single iambic foot at the right edge of the words. Uzbek has a quantity insensitive non-iterative stress pattern which is very similar to that of Persian. This analysis is shown in (3).

(3) σ σ σ (σ σ́)

The assumption in this study is that the foot structure of Persian is identical to what Hyde (2001) and Baković (2004) put forward for similar languages such as Uzbek. We suggest that a structure containing multiple feet as in (1) should be avoided, because it wrongly predicts secondary stresses at word level, and constructs unnecessary structure. Unbounded feet shown in (2) should also be avoided, because admitting them will increase the foot inventory and will impose complication upon prosodic typology (See Prince, 1985 for a detailed discussion against unbounded feet). Unbounded feet were inevitable in approaches which assumed prosodic structure to be strictly layered. In such approaches, all strings were assumed to be exhaustively parsed at all prosodic levels, and no skipping of layers was allowed, and thus unfooted syllables were prohibited. This idea was mainstream when Strict Layer Hypothesis (Selkirk, 1981 & 1984; Nespor and Vogel, 1986) was dominant in the literature of Prosodic Phonology. However, in past two decades non-exhaustive phrasing has been shown to exist in almost all prosodic levels for different languages. For example, at foot level, as Selkirk (1995) reports, it has been widely shown that there exist cases where a syllable is immediately dominated by a PWord node (see also Inkelas, 1989; McCarthy and Prince, 1991, 1993b; Ito and Mester, 1992; Kager, 1993; and Prince and Smolensky, 1993). Therefore, a non-exhaustive foot structure as shown in (3) is theoretically acceptable and does not cause any problem to the principles of Prosodic Phonology.

In Optimality Theory's terms, Persian stress can be explained by assuming that the constraint ALLFEETRIGHT which demands all the feet to be aligned with the right edge of a PWord dominates PARSE-SYLLABLE which militates against unfooted syllables. The inviolable constraint FOOTFORM(IAMBIC) ensures that all feet are iambic. It is well known that there is a fundamental rhythmic requirement which requires feet to be binary (Prince, 1980; Kager, 1989; Prince and Smolensky, 1993). This requirement is expressed in terms of the constraint FOOTBINARITY. These constraints are defined in (4)–(7), and the tableau in (8) shows their role in deriving the foot structure of a trisyllabic lexical word.

(4) ALLFEETRIGHT (FT-BIN): Assign one violation mark for each foot which is not right-aligned with some PWord.

(5) PARSE-SYLLABLE (PARSE-σ): Assign one violation mark per syllable unparsed into a foot.

(6) FOOTFORM(IAMBIC) (FT-FORM):Assign one violation mark for every foot which is not iambic (σ σ́) or (σ́).

(7) FOOTBINARITY (FT-BIN): Assign one violation mark for each foot containing more or less than two syllables.

(8)

Input: /setâre/ 'star'	FT-FORM	FT-BIN	ALL-FT-R	PARSE-σ
☞ a. se (tâ.ré)				*
b. (se.tá̂) re			*!W	*
c. se.tâ (ré)		W*!		**W
d. se (tá̂.re)	*!W			*
e. (se.tâ.ré)		*!W		L
f. (sé) (tâ.ré)		*!W	*!W	L

The winner candidate (8a) violates the lower-ranked PARSE-σ, but satisfies all other constraints. As mentioned above, since Persian does not have an iterative stress pattern, any word structure with more than one foot will not be acceptable, thus the fact that for instance a four syllabic word has one iambic foot on the right rather than two iambic feet justifies the ranking ALL-FT-R » PARSE-σ.

One important issue to be noted here is the existence of Persian monosyllabic lexical words such as *bu*, 'smell', *lab* 'lip' and *dast* 'hand' which obviously violate FOOTBINARITY. We

suggest that as is assumed in Prosodic Phonology (Selkirk, 2011), in Persian, lexical words are always matched by a PWord. This will be discussed and explained in section 3.4 and it will be addressed that this matching between syntactic words and PWords occurs only for lexical words and function words are not mapped onto PWords. The existence of monosyllabic content words can be explained by assuming that the constraint MATCH(LEX,ω) defined in (9) is ranked higher than FOOTBINARITY.

(9) MATCH(LEX,ω): Assign one violation mark for every lexical word in syntactic constituent structure, that does not correspond to a PWord (ω) in phonological representation.

Following Ito and Mester (2009) we assume that the *headedness* condition of prosodic constituents which prohibits headless constituents is universally inviolable and should be assigned to GEN. According to *headedness*, PWords which do not dominate any foot cannot be an output candidate of GEN. the tableau in (10) demonstrates the result of ranking MATCH(LEX,ω) over FOOTBINARITY. As can be seen, an unfooted monosyllabic word $(lab)_{\omega}$ is not present as a candidate, because headedness bans the generation of footless PWords as candidates.

(10)

Input: /lab/ 'lip'	MATCHLEX	FT-BIN
☞ a. ($(láb)_{\pi}$)$_{\omega}$		*
b. $(láb)_{\pi}$	*!W	*
c. lab	*!W	L

As described in chapter 1, Persian enclitics are never stressed and the stress of a cliticized word falls on the rightmost syllable of the lexical host. This is further exemplified in (11):

(11)

i.	pedar	+	-am	→	*pe.dá.ram*	ii.	barâdar	+	-i	→	*ba.râ.dá.ri*
	father		POSS1SG		'my father'		brother		IND		'a brother'

In order to explain why enclitics in Persian are always stressless we adopt a constraint proposed in Ito and Mester (2009), namely HEAD-TO-LEX (12), according to which heads of prosodic categories must be contained in lexical (not functional) material. It will be discussed in detail in section 3.3 that Persian enclitics are weak function words that cliticize to their preceding lexical heads.

(12) HEAD-TO-LEX: Assign one violation mark for each prosodic head which is contained in a function word.

Following Ito and Mester (2009), the assumption here is that weak function words are syllable sized by their nature, and cannot be the head of a PWord i.e., cannot be a part of a foot. The constraint in (12) will be violated if the head of a PWord (the rightmost syllable of a foot) is contained in an enclitic. We know that enclitics are never stressed in Persian, thus the constraint HEAD-TO-LEX must be undominated in the language. The head of a PWord (its foot) is on the right edge, as far as it does not occur on an enclitic. Ranking HEAD-TO-LEX above ALL-FT-R will derive the well-formed outputs in (11). The tableau in (13) demonstrates how the output form (11ii) is chosen.

(13)

Input: /barâdar + -i/ brother + IND (lex + fnc) 'a brother'	HEAD-TO-LEX	ALL-FT-R	PARSE-σ
☞ a. ba (râ.dá)$_{\pi}$ ri		*	**
b. ba.râ (da.rí)$_{\pi}$	*!W	L	**

The tableau in (13) does not give us a thorough account of the prosodic structure of cliticized words in Persian. First of all, we should consider the role of the constraints in (4)–(7) as well, but more importantly, as we will see in section 3.3, enclitics are not the only weak function words found in Persian, and there are many weak function words that cliticize to their following lexical hosts and behave as a proclitic. In order to have a comprehensive account of the internal structure of the PWord in Persian, one should take both enclitics and proclitics into account, and provide a uniform explanation for the both types of clitics. Section 3.3 aims to accomplish this goal.

3.3 Proclitics and enclitics

Many studies have been conducted to deal with the syntactic status or prosodic structure of Persian enclitics, among which one can name Amini (1997), Kahnemuyipour (2003), Samvelian and Tseng (2010), Abolhasanizadeh et al. (2012). However, it seems that there are no studies postulating proclitics in Persian.

Following the standard prosodic distinction between lexical and function words set forth in Selkirk (1995) and subsequent work, we assume that morpho-syntactic words in Persian can be divided into two groups. The first group consists of words which can have stress, and since most of these belong to the category of lexical or content words, following the literature, we refer to them as lexical (*lex*) words. The second group comprises weak function words (*fnc*) which are always stressless, except when focused or used metalinguistically. It is worth noting that the difference between *lex* and *fnc* in Prosodic Phonology and in this study is a purely prosodic one. Following Ito and Mester (2009) this study assumes that weak function words are syllables sized, not foot sized. In Persian, demonstrative and personal pronouns such as *un* 'that, he/she' or *to* 'you', although being semantically function words, are stressed and thus belong to the first group, while prepositions such as *az* 'from' or *bâ* 'with' are always stressless and belong to the second group.

Weak function words are prosodic clitics; i.e., they combine with other material into a single prosodic unit rather than being autonomous units themselves (Anderson 2005: 42). In this sense, Persian has both enclitics and proclitics.

Some of the most frequently used enclitics in Persian are listed in (14):

(14)

i. the indefinite article and relative particle *-i*
ii. the direct object particles *-o* and *-râ*
iii. the conjunction *-o*
iv. the Ezafe particle *-e*
v. the particle meaning 'also, even' *-am/-ham*
vi. possessive (pronominal) suffixes:
-am 'my', *-et* 'yourSG', *-eš* 'her/his', *-emun* 'our', *-etun* 'yourPL', *-ešun* 'their'
vii. personal agreement suffixes when attached to past tense verbs:
-am '1SG', *-i* '2SG', ∅ '3SG' *-im* '1 PL', *-in* '2 PL', *-an* '3 PL'
viii. the copula verbs *-e* 'be3SG', *-an* 'be3PL'

Personal agreement suffixes when attached to present tense verbs are cohering (i.e., are part of the PWord rather being a clitic). For a detailed discussion refer to Kahnemuyipour (2003:372–5). Some examples of proclitics in Persian are given in (15):

(15)

i. class 1 prepositions: *az* 'from', *bâ* 'with', *be* 'to', *dar* 'at', *tâ* 'until', *joz* 'except'
ii. class 2a prepositions, when used without Ezafe linker: *tu* 'in', *ru* 'on'
iii. the conjunctions *va* 'and' and *yâ* 'or'
iv. the particle *ke* in some of its usages

According to the classification in Pantcheva (2006), Persian prepositions are mainly divided into two groups. Class 1 or true prepositions are always stressless and never take Ezafe morpheme. Class 2 prepositions are themselves divided into groups, namely, class 2a

and class 2b. Class 2b prepositions are always stressed and obligatorily take Ezafe morpheme, while class 2a prepositions optionally take Ezafe. For an introduction of Ezafe, see §4.3 of this dissertation. Class 1 prepositions are always weak function words and prosodize as proclitics and class 2a preposition are weak function words only when used without the Ezafe linker. Class 2b prepositions and also class 2a prepositions when used with Ezafe morpheme, are syntactically more similar to nouns (Samiian, 1994; Larson and Yamakido, 2005). All types of prepositions and their examples are listed in table 1:

Table 1 The classification of Persian prepositions based on Pantcheva (2006) (The types marked by gray are weak function words and phrase as proclitics. The others are lexical words and behave syntactically and prosodically in a similar way to nouns.)

Class 1	Class 2		
	Class 2a		Class 2b
	Without Ezafe	With Ezafe	
az xuné 'from the house'	*tu xuné* 'in the house'	*tú-ye xuné* 'in/inside the house'	*nazdík-e xuné* 'near the house'
be xuné 'to the house'	*ru xuné* 'on the house'	*rú-ye xuné* 'on the house'	*birún-e xuné* 'outside the house'

3.3.1 Previous attempts

As mentioned previously, the prosodic structure of proclitics in Persian has not been specifically explored in previous studies, though there are a number of works that study enclitics. This section briefly introduces four analyses of Persian enclitics proposed in the literature, and points out their theoretical and empirical problems.

Extrametricality

Amini (1997) proposes for Persian enclitics a non-recursive exhaustively parsed structure in which enclitics form a single PWord with their hosts. She then explains the lack of stress

of these enclitics using Extrametricality Theory. The examples in (16) demonstrate Amini's approach. The last syllable in (16i) and the last two ones in (16ii) are considered extrametrical.

(16) i. ketâb -i → *ke.tấ* <*bi*>
book IND
'a book'

ii. ketâb -ešun → *ke.tấ* <*be.šun*>
book POSS3PL
'their book'

Firstly, the constraint ranking in Amini's OT analysis does not explain why only cliticized words are subject to Extrametricality, and that Extrametricality does not apply to other types of words. More importantly, according to Hayes (1982, 1995) only constituents (segment, syllable, foot, PWord, etc.) may be marked as extrametrical. However, in Persian, we frequently encounter cases in which more than one enclitic is attached to a host, resulting in a clitic cluster. In most cases, such clusters, which are considered to be "extrametrical" in Amini's analysis, do not form a constituent in either prosodic or morphological sense. Example (17) illustrates such cases:

(17) i. ketấb -ešun -am → * *ke.tấ* <*be.šu.nam*>
book POSS3PL too
'their book too'

ii. ketấb -etun *-o* → * *ke.tấ* <*be.tu.no*>
book POSS2PL ACC
'your book- ACC'

If we assume with Amini (1997) that stresslessness of enclitics in Persian is due to extrametricality, then the entire strings within < > in example (17) must be considered as extrametrical constituents. However, these strings are not single prosodic or morphological constituents, and thus cannot be subject to extrametricality. Note that in cases like (17) an approach assuming a cyclic process in which each morpheme is marked as extrametrical in a different cycle is also unacceptable, because the extrametricality rule/constraint has access only to the final constituent of a prosodic category.

In sum, explaining the lack of stress on Persian enclitics by extrametricality is problematic, since it fails to account for a wide range of cases. Therefore, it is impossible to take enclitics and their hosts as a single non-branching PWord as suggested in Amini (1997).

Clitics as autonomous PWords

In Kahnemuyipour (2003), prosodic parsing is exhaustive in all levels and no recursive structure is allowed. All morpho-syntactic words including weak function words such as enclitics form their own PWords, and due to the phrasal stress rule, the leftmost PWord in the Phonological Phrase (PPhrase) takes stress. The reason enclitics never receive stress is that these functional morphemes happen to follow lexical words, and the phrasal stress rule puts the stress on the leftmost lexical word: $((\textit{ketâb})_{\omega}\ (\textit{i})_{\omega})_{\varphi}$ 'a book'.

As discussed in the literature (Selkirk 1995; Anderson 2005; Ito and Mester 2009) weak (stressless) function words cannot constitute PWords by themselves, and essentially tend to phrase with their adjacent lexical words. Therefore, theoretically it is inappropriate to take Persian enclitics as autonomous PWords in the absence of any phonological or phonetic evidence.

Taking Persian function words as independent PWords also suffers from a major empirical problem. Kahnemuyipour (2003) does not deal with proclitics, but if all morpho-syntactic words, including enclitics, can constitute PWords, then proclitics such as class 1 prepositions should also have the right to form PWords on their own. However, in the case of most proclitics, this analysis will wrongly predict the PPhrase stress on the phrase-initial proclitic as shown in (18).

(18) i. [az tehrân]$_{PP}$ → * ((*áz*)$_{\omega}$ (*tehrân*)$_{\omega}$)$_{\varphi}$
from Tehran

ii. [be un xune]$_{PP}$ → * ((*bé*)$_{\omega}$ (*un*)$_{\omega}$ (*xune*)$_{\omega}$)$_{\varphi}$
to that house

In brief, treating weak function words in Persian as independent PWords encounters theoretical and empirical problems and should therefore be avoided.

Two other Proposals

Clitic Group (CG) and Accentual Phrase (AP) categories have also been proposed for Persian.[1] Both CG and AP are composed of a lexical word and its enclitics. One major problem with these approaches is that they do not take proclitics into consideration, and thus fail to account for the asymmetry between proclitics and enclitics. More importantly, the necessity of such additional and non-universal categories as AP and CG has been questioned and rejected frequently in the literature (Selkirk 1995, 2011; Anderson 2005; Ito and Mester 2009; Féry 2010). As pointed out in Ito and Mester (2009: 168), "Occam's razor militates against any additional category such as Clitic Group, as long as the existing ones are sufficient to represent the prosodic phrasings". Another problem according to Wicka (2009) is that in fact, reference to categories such as Clitic Group would obscure generalizations that may be captured by limiting our discussion to include those constituents presented in §1.3 As Wiese (1996: 251) asserts, properties of [host+clitic] sequences are no different than those of the PWord.

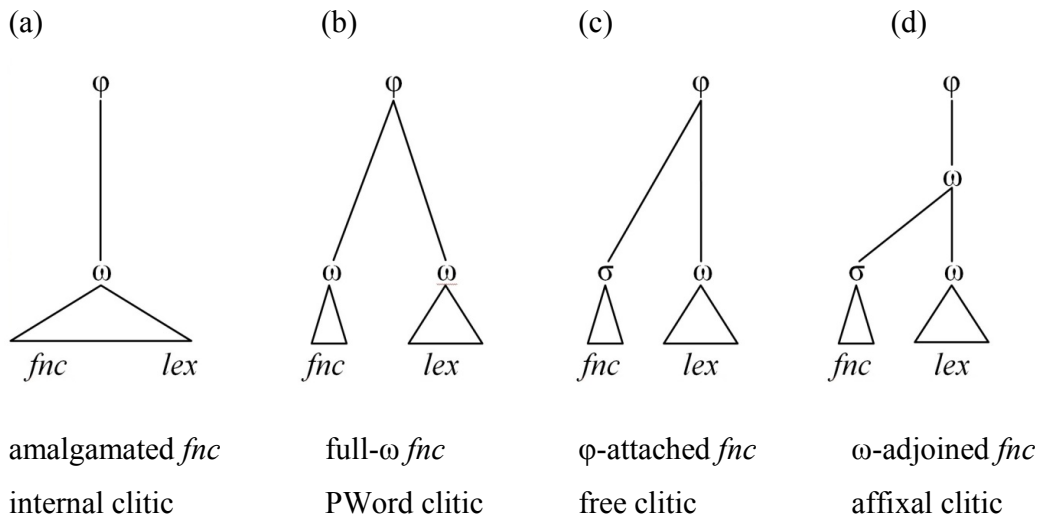
3.3.2 Proclitics

As Selkirk (1995) points out, for any given [host+clitic] combination, irrespective of their order, there are essentially four formal possibilities. These four possibilities are illustrated

[1] For a detailed discussion of AP see Sadat-Tehrani (2007), and for CG, refer to Abolhasanizadeh, Gussenhoven and Bijankhan (2012).

for proclitics in (19). Following the definitions in Selkirk (1995), affixal clitics are clitics in which the lexical host comprises its own PWord, and the clitic is adjoined to the host in a recursive structure. Free clitics on the other hand are incorporated directly into the PPhrase node, as a sister of the PWord containing the lexical word, in a non-exhaustive structure.

(19)



A crucial point in (19) is that the proclitic *fnc* is ω-initial (located at the initial position of a PWord) in (19a), (19b), and (19d), but not in (19c). The *fnc* in (19c) is the only one which is not parsed into ω, and is immediately dominated by the φ (PPhrase) node, although in all four structures the proclitic *fnc* is φ-initial.

Selkirk (1995) reports that the initial position in a PWord is often associated with effects involving the phonetic realization of segments. In English, for example, a word-initial voiceless stop is aspirated, even when the syllable to which it belongs is stressless.

Cooper (1991, 1994), as cited in Selkirk (1995), shows that there is a distinct word-initial aspiration effect which cannot be reduced to a simple syllable-initial effect. Prosodic Structure Theory takes such 'word-initial' effects to be PWord-initial effects.

Aspiration in Persian is studied in Samareh (1999: 27) according to which all voiceless stops are aspirated, but the degree of aspiration depends on the position of the stop. In the onset position of stressed syllables and word initially, stops are more aspirated. For example, /p/ in *pa.rvấz* 'flight' and *se.pấh* 'army' are more aspirated than in *se.pâ.ye* 'tripod'. If we accept the claim in Samareh (1985), then it becomes clear that Persian is similar to English in having a word-initial aspiration effect.

Therefore, if we can demonstrate that proclitic-initial stops are significantly less aspirated than word-initial stops, then we can claim that proclitics are not located at the initial position of a PWord, and this will support the structure in (19c).

An experiment was conducted to compare the degree of aspiration between proclitic-initial stops and word-initial ones.

PWord-initial aspiration, an experiment

In this experiment, the Persian class 1 preposition *tâ* and the class 2a preposition *tu* were used in two sentences before polysyllabic nouns starting with the syllables /ta-/ and /tu-/, in non-ι-initial positions.

Six Persian native speakers read each of the two sentences (20) and (21) twice, and the utterances were recorded at a sampling frequency of 44,100 Hz and were analyzed using the phonetics software Praat.

(20) *dišab* ***tu*** ***tunél*** *tasâdóf* *kard-am*
last night in tunnel accident do PAST-1SG
'I had an accident in the tunnel last night.'

(21) *behrúz* ***tâ*** ***tâbestún*** *kấr* *kard-Ø*
PR until summer work do PAST-3SG
'Behruz worked till summer.'

The Voice Onset Time (VOT) in proclitics and word-initial syllables was measured and the mean values and standard deviations were calculated. A paired Student's t-test was conducted to determine the significance of differences between VOT values in the two positions. The significance threshold of the t-test was 0.05, the degree of freedom being the number of participants minus one. Table 2 presents a summary of the statistical analyses.

Table 2 Summary of VOT analyses

	ta (proclitic)	*ta-* (word-initial)	*tu* (proclitic)	*tu-* (word-initial)
Mean VOT (ms)	22.35	51.35	26.55	62.99
Standard deviation	3.597	3.68	4.13	6.47
t-test results	$t(5) = 131.13, p < 0.05$		$t(5) = 106.2, p < 0.05$	

As is apparent from the t-test results, the VOT value (degree of aspiration) in proclitics is significantly less than that in word-initial position, showing that there is less aspiration on proclitics than on word-initial syllables. In Figure 1, two examples of the recorded data in this experiment are presented, in which the substantial difference between the aspiration of initial voiceless stop on proclitics and on the first syllable of lexical words is observable.

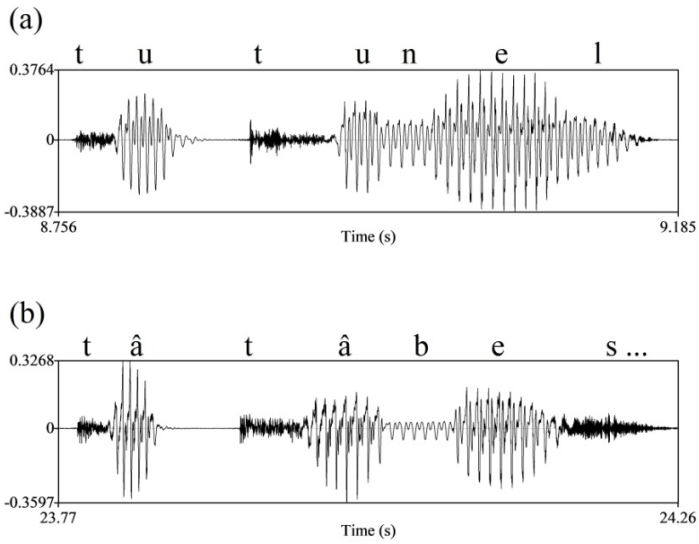


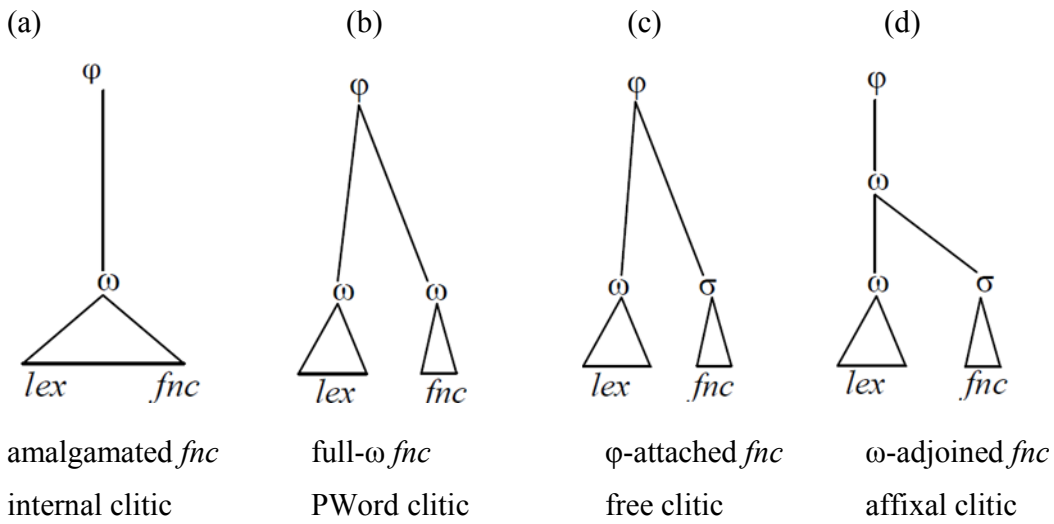
Figure 1 VOT differences observable in recorded data

Based on these findings, we can conclude that proclitics in Persian are not in PWord-initial position; thus they are not ω-adjoined, and rather are φ-attached free clitics as shown in (19c). Section 3.3.3 provides more evidence for this claim.

3.3.3 Enclitics

The structures in (22) illustrate the four possibilities of [host+clitic] combinations for enclitics.

(22)



The structure in (22a) is the one proposed in Amini (1997), and as we discussed in §2.1, it is not acceptable because if we take Persian enclitics as amalgamated function words, their lack of stress cannot be explained by extrametricality. The structure in (22b), in which weak function words are independent PWords, is the one proposed in Kahnemuyipour (2003), and as we saw in §2.2, this also is not acceptable due to both theoretical and empirical problems. Consequently, there are only two remaining possibilities for the structure of enclitics in Persian, namely (22c) and (22d). The main difference between (2c) and (2d) is that in the latter, the clitic and its host are parsed into the same PWord, but in

the former, the clitic is not parsed into PWord. The evidence for our argumentation comes from syllabification.

Syllabification

Syllabification is carried out on different levels. At the lexical level, syllabification rules/constraints require all segments in the domain to syllabify with each other, but at post-lexical levels, factors such as high rate of speed can cause syllabification to optionally take place at domain boundaries. Laeufer (1995: 118–119) states that syllabification applies cyclically to roots and is reapplied after each word formation process to incorporate newly added elements. In some languages, however, it is sensitive to morphological structure, while in others it is not. In other words, in some languages the domain of lexical syllabification is the morpheme, and in others the word.

The canonical syllable type in Persian is CV(C)(C) and the prosodic requirements of the syllable, such as obligatory onset, allowing no consonant clusters in onset position, and allowing at most two segments in coda position, determine the syllable structure in the word domain, ignoring the morpheme boundaries (Mahootian, 1997: 303–305). In (23i and ii) morpheme boundaries are shown by '-' on the left, while a '.' marks syllable boundaries on the right.

(23)

i.	dân-eš-mand-ân	→	*dâ.neš.man.dân* / **dân.(ʔ)eš.mand.(ʔ)ân*	'scientists'
ii.	ham-âhang-i	→	*ha.mâ.han.gi* / **ham.(ʔ)â.hang.(ʔ)i*	'harmony'

As can be seen, the syllabification is not sensitive to morphological structure and takes the whole word as its domain. Since in Prosodic Phonology PWords are the equivalents of words in syntactic constituent structure (Selkirk, 2011), we can conclude that in Persian lexical syllabification takes place in the PWord domain. However, resyllabification is also possible at higher (post-lexical) levels, ignoring PWord boundaries. This is shown in (24):

(24)

i. ((har)$_{\omega}$ (âdam)$_{\omega}$)$_{\varphi}$ → *har. (ʔ)â.dam / ha.râ.dam* 'each human'

ii. ((panj)$_{\omega}$ (angošt)$_{\omega}$)$_{\varphi}$ → *panj. (ʔ)an.gošt / pan.jan.gošt* 'five fingers'

The syllabic structures to the left of the slash (/) show the lexical level syllabification, while the structures to the right show the post-lexical resyllabification, which ignores PWord boundaries. The optional existence of such resyllabifications proves that the obligatory syllabification is only a characteristic of the PWord domain and does not necessarily apply to higher domains such as PPhrase.

A close investigation of syllabification in cliticized words reveals that content words and their enclitics are necessarily syllabified together and it is not possible to syllabify them independently, as shown in (25).

(25)

i. [márd]$_{lex}$ [i]$_{fnc}$ → *már.di / *márd.(ʔ)i* 'a man'

ii. [pedár]$_{lex}$ [am]$_{fnc}$ → *pe.dá.ram / *pe.dár.(ʔ)am* 'my father'

iii. [ketấb]$_{lex}$ [o]$_{fnc}$ → *ke.tấ.bo / *ke.tấb.(ʔ)o* 'book-ACC'

iv. [gárm]$_{lex}$ [e]$_{fnc}$ → *gár.me / *gárm.(ʔ)e* 'It's hot.'

This shows that [host+enclitic] combinations are more similar to the structures in (23) than to those in (24), meaning that there are no maximal PWord boundaries between a host and its enclitic. If we return to the two possible structures for enclitics mentioned in §3.2, we now have enough evidence to claim that the only acceptable structure for [host+enclitic] combinations is the one shown in (8d), which is a ω-adjoined affixal clitic.

There are two PWord nodes (shown by ω) in (22d). The lower PWord, which does not dominate any other PWord, is called a Minimal PWord, and the higher PWord, which is not dominated by any other PWord, is called a Maximal PWord. In Persian, the domain of obligatory syllabification is Maximal PWord, and stress is culminative and obligatory in

this domain. Minimal PWord is right-aligned with a *lex*, and stress is rightmost in this domain.

3.3.4 Further evidence

In this section, three additional pieces of evidence are provided, that support the claim that proclitics and enclitics behave asymmetrically in Persian, and that there is a maximal PWord boundary between a proclitic and its host, while enclitics and their hosts are parsed into a single maximal PWord.

Proclitics and Syllabification

In the previous section we argued that the obligatory syllabification of [host+enclitic] combinations in Persian suggests that there is no PWord boundary between a host and its enclitic. Proclitics, on the other hand, behave like the structures in (24), suggesting that they are not parsed into PWords together with their hosts and that there is a PWord boundary between them and their lexical hosts. This is exemplified in (26):

(26)

i.	[az]$_{fnc}$	[injấ]$_{lex}$	→	*(ʔ)a.zin.jấ* / *(ʔ)az.(ʔ)in.jấ*	'from here'
ii.	[dar]$_{fnc}$	[âsemấn]$_{lex}$	→	*da.râ.se.mấn* / *dar.(ʔ)â.se.mấn*	'in the sky'
iii.	[joz]$_{fnc}$	[irấn]$_{lex}$	→	*jo.zi.rấn* / *joz.(ʔ)i.rấn*	'except Iran'

Uninterruptibility

It is uncontroversial that it is prosodically unacceptable to insert a pause in the middle of a PWord:*[pe…dár]$_{\omega}$. This principle is known as *uninterruptibility* and is used as one of the major criteria of wordhood (Bauer 2003: 63–64). An examination of clitics in Persian shows that the insertion of a pause between a *lex* and its enclitics is not possible, however, speakers can make a pause between a *lex* and its proclitic. As shown in (27) pause insertion is not possible between a noun and the Ezafe particle (27i), between a verb and a personal agreement suffix (27ii), or between a noun and the conjunction -*o* (27iii). However, a pause

can be made between a proclitic and its host, for instance, between the complementizer *ke* and the material following it (27iv), between a preposition and a noun (27v), and between the conjunction *va* and the following noun (27vi).

(27) i. **mard...-e dânấ* ‘wise man’ ii. **ráft...-im* ‘we went’ iii. **zán...-o márd* ‘woman and man’

iv. *behrúz ke...dirúz* ‘Behruz that yesterday…’ v. *az…šomấ* ‘from you’ vi. *zan va…márd* ‘woman and man’

Glide epenthesis

In Persian words, no adjacent vowels are allowed (Windfuhr and Perry 2009: 429). Adjacent vowels are also avoided when the first of two vowels is the final vowel of a *lex* and the second one is the initial vowel of an enclitic. In this case, an epenthetic segment (often a glide) is inserted between the two vowels to resolve the hiatus. In contrast, there is no such epenthetic segment insertion between a proclitic-final vowel and a *lex*-initial vowel. In (28i, ii and iii) word-final vowels are followed by a clitic vowel, and to avoid vowel adjacency, a glide is epenthesized between the two vowels.

(28) i. [xodấ]$_{lex}$ [i]$_{fnc}$ → *xodấ-**y**i*
God IND
‘a God’

ii. [pesté]$_{lex}$ [e]$_{fnc}$ → *pesté-**y**e*
pistachio EZ
‘pistachio-EZ’

iii. [xoš-bu]$_{lex}$ [-i]$_{fnc}$ → *xošbú-**y**i*
sweet-smelling be2SG
‘You are sweet-smelling.’

In (29i, ii and iii) we have exactly the same vowel sequences as in (28), but unlike (28), glide epenthesis between the two vowels is unacceptable.

(29) i. [bâ]$_{fnc}$ [ín]$_{lex}$ → **bâ-**y**-ín*
with this
'with this'

ii. [be]$_{fnc}$ [eskâtlánd]$_{lex}$ → **be-**y**-eskâtlánd*
to Scotland
'to Scotland'

iii. [tu]$_{fnc}$ [estaxr]$_{lex}$ → **tu-**y**-estáxr*
in pool
'in the pool'

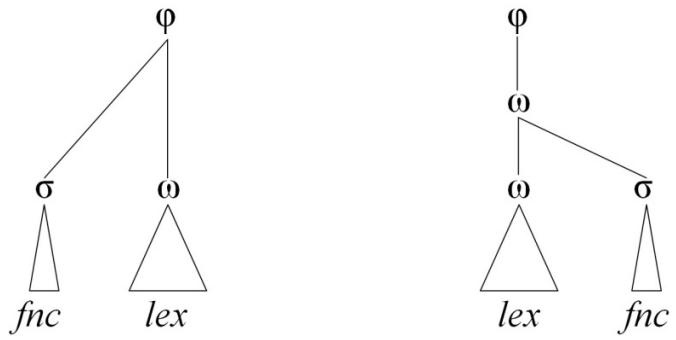
In cases such as *bâ*+-*am* → *bâhâm* 'with me', it may seem that an epenthetic segment is inserted after the preposition *bâ*, suggesting that *bâ* may not exclusively be a proclitic. However, as Naderi and van Oostendorp (2011: 164–165) observe, the insertion of /h/ in *bâhâm* or *behem* 'to me' is not productive, and happens in only a few frozen phrases, and should not be considered an epenthesis process in Persian. In these cases, the second part of the combination (the nominal suffix -*am*) is a prosodically weak function word itself and cannot be a legitimate host for a proclitic. These fixed fossilized combinations of two clitics form a single PWord in these examples, while *bâ* and *be*, when used as true prepositions before lexical words, are always proclitics.

Sometimes epenthetic consonants other than glides can appear between a vowel-final word and a vowel-initial enclitic. For instance, before the accusative marker -*o*, the consonant /r/ appears due to the effect of the full form of the clitic -*râ*, and before the conjunction -*o*, the consonant /v/ is inserted probably due to the effect of the other conjunction *va*. However, such epenthetic insertions are never observed between vowel-final proclitics and vowel-initial words.

3.4 Phonological Word: an OT account

As we saw the previous section, weak function words in Persian prosodize asymmetrically depending on weather they are proclitics or enclitics. Enclitics regardless of their morpho-syntactic characteristics or semantic content, always attach to their hosts as an affixal clitic, forming a recursive PWord. This is apparent, since enclitics phonologically behave as a part of the PWord containing their host. They obligatorily syllabify with their hosts, they are uninterruptible from their lexical hosts, and vowel hiatus is not tolerated at the boundary between a host and an enclitic. On the other hand, proclitics regardless of their morpho-syntactic characteristics or semantic content, behave independently from the PWord containing their hosts. They do not necessarily syllabify with their hosts, they are interruptible from their hosts, and vowel hiatus can be tolerated in their boundary with their hosts. These two structures are repeated in (30):

(30) (a) (b)



Prosodic structure of proclitics

Prosodic structure of enclitics

Considering the structure in (30b), we can conclude that all the function words following a lexical element have to be parsed into PWord in some level. This means that between the right edge of a PWord and the right edge of the PPhrase containing it, nothing should be left unparsed into PWord.

Elfner (2009) argues that in Irish, prosodically weak pronoun objects never occur at the left edge of PPhrases. In Elfner's analysis, weak function words are free clitics and constraints

on prosodic structure are in direct competition with constraints on linearization. Thus, in order to avoid PPhrase-initial free clitics, weak pronoun objects appear phrase-finally or medially but not phrase-initially in outputs. The Irish examples in (31) are taken from Elfner (2009) and show this interaction between prosodic constraints and syntactic linearization constraints. The phrase within the parentheses constructs a PPhrase, and as can be seen, the weak function word *é* 'it' is avoided at the left edge of a PPhrase.

(31)

i. ? Léigh Liam (*é* ar an traein aréir)
read Liam it on the train last-night

ii. Léigh Liam (ar an traein *é* aréir)
read Liam on the train it last-night

iii. Léigh Liam (ar an traein aréir *é*)
read Liam on the train last-night it
'Liam read it on the train last night.'

Elfner (2009) explains the tendency in Irish to avoid starting PPhrases with prosodically weak elements in terms of the constraint PARSE-EDGE-L(φ,ω). This constraint defined in (32) demands all PPhrases to start with a PWord.

(32) PARSE-EDGE-L(φ,ω): Assign one violation mark for every PPhrase whose left edge does not correspond to the left edge of a PWord.

We saw above that Persian does not allow free clitics between the right edge of a PWord and the right edge of the PPhrase containing it. This can be captured by the mirror image constraint of the constraint proposed in Elfner (2009), namely PARSE-EDGE-R(φ,ω):

(33) PARSE-EDGE-R(φ,ω): Assign one violation mark for every PPhrase whose right edge does not correspond to the right edge of a PWord.

The constraint in (33) is violated by any PPhrase that directly dominates a syllable on its right edge. This will occur when a free clitic is located at the right edge of a PPhrase. Since enclitics are never free clitics, this constraint is inviolable in Persian.

The reason why a constraint such as ALIGN(φ,R;ω,R) which requires the right-alignment of PPhrases with some PWord is not adopted in this analysis is that this constraint implies PPhrases to be right-headed. In fact, ALIGN(φ,R;ω,R) is used in the same way as ALIGN(φ,R;Head,R) in the literature, however, as we will see in chapter 4, PPhrases in Persian are left-headed. Therefore, the use of the ALIGNMENT constraint has been abandoned and it has been replaced by the PARSE-EDGE constraint which does not make any implications about the direction of the head of a PPhrase.

We introduced the constraint MATCH(LEX,ω) in the §3.2 briefly. MATCH(LEX,ω) is an instance of S(yntax)-P(honology) faithfulness constraints proposed in Selkirk (2011). According to this constraint, a word in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it ω, in phonological representation. We suggest that this constraint is never violated in Persian, because as we saw in this chapter, each lexical word is necessarily matched with a PWord. In Prosodic Phonology, each S-P faithfulness constraint has a P-S counterpart which requires that prosodic constituency be a faithful reflection of syntactic constituency. The P-S counterpart for MATCH(LEX,ω) would be the constraint in (34):

(34) MATCH(ω,LEX): Assign one violation mark for every PWord in phonological representation, that does not correspond to a lexical word in syntactic constituent structure.

Free clitics (30a) do not violate the constraint MATCH(ω, LEX) because they are located outside the PWord domain, while this constraint is violated in case of affixal clitics (30b), since the maximal PWord does not match with any lexical item. This can be seen in (35) in which the PWord is matched by the *lex* in the procliticized example (35i), but the maximal PWord is not matched by any *lex* in the encliticized example (35ii).

(35)

i. $(fnc\ (lex)_{\omega}\ )_{\varphi}$ ii. $(\ (\ (lex)_{\omega}\ fnc)_{\omega}\ )_{\varphi}$

Another constraint that should be considered to determine the prosodic structure of Persian is PARSEINTO-φ defined in (36) following Ito and Mester (2009):

(36) PARSEINTO-φ: Assign one violation mark for any element of terminal string which is not parsed into a PPhrase.

As asserted in Ito and Mester (2009), in prosodic hierarchy, the elements of the terminal string are the phonological elements. PARSEINTO-φ only requires that segmental strings belong to PPhrase, and is crucially not "PARSE-INTO-WELLFORMED-φ".

We propose that the prosodic structure of proclitics and enclitics in Persian can both be explained uniformly, by identical ranking of the four constraints HEAD-TO-LEX, PARSE-EDGE-R, MATCH(ω,LEX), and PARSEINTO-φ. The interaction of these constraints is shown in the tableau in (37) for enclitics. The fact that weak function words are never matched with PWords shows that HEAD-TO-LEX is undominated and is never violated by winner candidates. The constraint PARSE-EDGE-R and PARSEINTO-φ are also undominated in Persian. As mentioned in §3.3.2, the obligatory syllabification of enclitics and their hosts shows that they belong to the same PWord, thus the structure in candidate (37b) in which there is a PPhrase boundary between the enclitic and its host is not acceptable. Candidate (37c) is less harmonic than the winner, because the prosodic head is contained in a weak function word in this candidate. In candidate (37d), the function word is parsed as a PWord. This candidate also violates HEAD-TO-LEX, because we know that all PWords contain at least one foot, and headedness condition of prosodic constituents assigned to GEN prevents it from generating headless constituents as candidates. Candidates (37e) and (37f) in the combination tableau (37), show that PARSE-EDGE-R outranks MATCH(ω,LEX) and candidate (37b) being a looser proves that PARSEINTO-φ is ranked over MATCH(ω,LEX).

(37)

Input: /barâdar + -i/ brother + IND (*lex* + *fnc*) 'a brother'	PARSEINTO-φ	HEAD-TO-LEX	PARSE-EDGE-R	MATCH(ω,LEX)
☞ a. $(((\text{ba râ.dá})_{\omega}\ \text{ri})_{\omega})_{\varphi}$				*
b. $((\text{ba râ.dá})_{\omega})_{\varphi}\ \text{ri}$	*!W			L
c. $((\text{ba râ.da rí})_{\omega})_{\varphi}$		*!W		*
d. $((\text{ba râ.dá})_{\omega}\ (\text{rí})_{\omega})_{\varphi}$		*!W		*
e. $((\text{ba râ.dá})_{\omega}\ \text{ri})_{\varphi}$			*!W	L
f. $(\text{ba râ.da.ri})_{\varphi}$			*!W	L

The tableau in (38) demonstrates the interaction of the same constraints in case of proclitics. The winner candidate (38a) does not violate any of the constraints, while the other candidates violate at least one constraint. In candidates (38b) and (38c) there is one PWord which does not match with any lexical word. Candidate (38d) violates HEAD-TO-LEX as well as MATCH(ω,LEX), because the preposition parsed as a PWord does not match with a *lex* and contains a foot due to the headedness condition. In candidate (38e), the PPhrase is not right-aligned with a PWord, and in candidate (38f), the preposition is not parsed at PPhrase level.

(38)

Input: /be + pedar/ to + father (*fnc* + *lex*) 'to father'	PARSEINTO-φ	HEAD-TO-LEX	PARSE-EDGE-R	MATCH(ω,LEX)
☞ a. (be (pe.dár)$_{\omega}$)$_{\varphi}$				
b. ((be pe.dár)$_{\omega}$)$_{\varphi}$				*!W
c. ((be (pe.dár)$_{\omega}$)$_{\omega}$)$_{\varphi}$				*!W
d. ((bé)$_{\omega}$ (pe.dár)$_{\omega}$)$_{\varphi}$		*!W		*!W
e. (be.pe.dar)$_{\varphi}$			*!W	
f. be ((pe.dár)$_{\omega}$)$_{\varphi}$	!*W			

The tableau in (39) shows the interactions of the same constraints in case of a lexical word with both enclitics and proclitics. In the winner candidate (39a), the maximal PWord violates MATCH(ω,LEX). In candidate (39b), both function words are free clitics. This candidate violates the higher ranked PARSE-EDGE-R. In candidate (39c) on the other hand, both function words are affixal clitics. In this candidate, the two non-minimal PWords violate MATCH(ω,LEX), which makes this candidate less harmonic than the winner (39a) which violates MATCH(ω,LEX) only once. In candidate (39d), the two *fnc*s which are mapped onto PWords violate HEAD-TO-LEX, and candidate (39e) also violates this constraint because the head of the PWord is contained in a *fnc*. Finally, candidate (39f) violates PARSEINTO-φ, because one *fnc* is not parsed into PPhrase.

(39)

Input: /az + pedar + -eš/ from + father + POSS3SG (*fnc* + *lex* + *lex*) 'from her/his father'	PARSEINTO-φ	HEAD-TO-LEX	PARSE-EDGE-R	MATCH(ω,LEX)
☞ a. (az ((pe.dá)$_{\omega}$ reš)$_{\omega}$)$_{\varphi}$				*
b. (az (pe.dá)$_{\omega}$ reš)$_{\varphi}$			*!W	L
c. ((az ((pe.dá)$_{\omega}$ reš)$_{\omega}$)$_{\omega}$)$_{\varphi}$				**
d. ((áz)$_{\omega}$ (pe.dá)$_{\omega}$ (réš)$_{\omega}$)$_{\varphi}$		**!W		**
e. (az (pe.da.réš)$_{\omega}$)$_{\varphi}$		*!W		*
f. az (((pe.dá)$_{\omega}$ reš)$_{\omega}$)$_{\varphi}$	*!W			*

3.5 The direction of cliticization

The direction in which a clitic attaches to its host has been taken for granted in this study so far. Weak function words cliticize to their adjacent lexical item with which they have the strongest morpho-syntactic relation. For example, the indefinite article suffix /-i/ is uncontroversially an enclitic, while a preposition such as *az* 'from' takes its following material as its host and cliticizes as a proclitic. This is easily explainable by syntax-prosody interface constraints which demand morpho-syntactic units to be matched by prosodic constituents (see chapter 4 for the details). Prosodic phrase boundaries are determined by the interaction of syntax-prosody interface constraints and prosodic markedness constraints, and prosodic phrase boundaries determine the direction of cliticization. For example in (40) the PP *bâ rezâ* 'with Reza' is mapped onto a PPhrase, and the function word *bâ* is cliticized to the lexical element in the same PPhrase. The NP *film-o* 'movie-ACC' is also matched with a PPhrase and the accusative case marker /-o/ does not have any chance to cliticize to its following verb. This will be discussed in more details in chapter 4 where we concentrate on the prosodic structures higher than the PWord level.

(40) *diruz* *bâ* *rezâ* *film-o* *did-im*
yesterday with Reza movie-ACC seePAST-1PL
'We watched the movie with Reza yesterday.'

However, there are cases in which a weak function word has exactly equal syntactic relations with its preceding and following materials, and does not belong to either of them. Since they are not a part of their preceding or following syntactic phrase (XP), we call them 'XP-external function words' or 'XP-external-clitics'. An interesting example of XP-external clitics in Persian is the case of the two conjunctions *-o* and *va-* both meaning 'and'. The conjunction *-o* is always an enclitic and necessarily syllabifies with its preceding host, while *va-* is always phrased with its following material.
The conjunction *-o* is originally a native Persian word which can be traced back to Old Persian as summarized in (41) based on Abolghassemi (2008), while *va-* is originally an Arabic conjunction (*wa*) borrowed into Persian.

(41)

Old Persian	Middle Persian	Classical Persian	Modern Persian
uta	*ud*	*-u*	*-o*

The example in (42) adopted from Kahnemuyipour (2003), shows the syntactic structure of two noun phrases which are connected to each other by a conjunction. Kahnemuyipour (2003) only mentions *-o* as a conjunction in Persian, and uses the fact that *-o* is phrased with the preceding material, as an evidence for left-alignment of syntactic phrases with PPhrases in Persian. However, taking the other conjunction (*va-*) into account will provide evidence against Kahnemuyipour's claim.

(42) (NP) & (NP)

The prosodic structures of the two conjunctions are given in (43). Needless of mentioning, according to the suggestions in this chapter, the conjunction *-o* being an enclitic is an affixal clitic, while *va* being a proclitic is a free clitic.

(43)

i. $(lex\text{-}o)_{\varphi}$ $(lex)_{\varphi}$ ii. $(lex)_{\varphi}$ $(va\ lex)_{\varphi}$

Many approaches to the study of the prosodic structure of clitics have suggested that the direction of cliticization is lexically specified for weak function words. For example Klavans (1985) proposes a parameter for specifying the direction of cliticization for each clitic. In this approach, the directionality parameter is specified individually and is a property of particular clitics, rather than being reducible to general properties of the language. Nespor and Vogel (1986) take the role of syntactic structure into account in determining the direction of cliticization. In their view, clitics are lexically specified to attach in only one direction, otherwise, they phrase with the host with which they share the most syntactic structure. Booij (1995, 1996) deals with the clitics in Dutch and observes that the same clitic may attach either leftward or rightward in this language. He concludes that these elements could not be lexically specified for a direction of attachment. Rather, the direction in which these clitics are attached is determined by the general properties of Dutch prosody. However, Booij points out two Dutch pronominal clitics, for which a lexical specification of direction of attachment is necessary. In contrast, Anderson (2005) proposes a view which treats direction of attachment of all clitics as a matter that follows from the overall prosodic properties of the language, and which is not available for lexical specification with respect to individual items.

This is the view adopted in this study for the direction of cliticization in Persian. We suggest that when the syntax-prosody interface constraints cannot determine the direction of cliticization, prosodic constraint takes the turn to determine it. For instance, in the case of the two conjunctions, both structures in (43) are predictable by the constraint ranking in (37). The suggestion is that the candidate that satisfies the syllable well-formedness conditions of the language wins over the other. Onsetless syllables are known to be marked and avoided universally (McCarthy and Prince, 1994). As we saw in chapter1, Persian prohibits syllables without onset. This tendency is captured by the constraint ONSET which is defined in (44).

(44) ONSET: Assign one violation mark for every onsetless syllable in the output.

If syntactic evidence can be provided supporting that the conjunction *-o* is syntactically related to its preceding XP, whereas the conjunction *va* is syntactically associated to its following XP, then the syntax-phonology constraints (discussed in chapter 4) can explain the direction of cliticization in these weak function words without reference to syllable well-formedness constraints. However, in the absence of such evidence, following Kahnemuyipour (2003) we take the conjunctions to be equally unrelated to their preceding and following XPs, i.e., to be 'XP-external clitics'.

When the syntax-prosody interface constraints cannot determine the direction of cliticization, if the weak function word located between two lexical words begins with a vowel, it prefers to phrase as an affixal clitic with its preceding *lex*, to avoid forming an onsetless syllable. While, if the *fnc* begins with a consonant, it will not form an onsetless syllable if it phrases with its following *lex*, and the ranking in (37) will automatically choose a candidate in which the *fnc* is a free proclitic. We will need the ranking in (45), to be able to eliminate the vowel-initial weak function words as proclitics:

(45) PARSEINTO-φ, HEAD-TO-LEX, PARSE-EDGE-R, ONSET » MATCH(ω,LEX)

The tableau in (46) shows how this ranking prefers *-o* to be an enclitic, and the tableau in (47) demonstrates how *va-* is preferred to be a proclitic by the same ranking. In tableau (46), if we do not consider the constraint ONSET, candidate (46b) will be optimal with only one violation of MATCH(ω,LEX). However, ONSET, being ranked higher than MATCH(ω,LEX) will eliminate this candidate in favor of candidate (46a) which satisfies ONSET. Candidate (46c) is disharmonic, because the conjunction is not parsed into PPhrase. This is inconsistent with the fact that the conjunction *-o* obligatorily syllabifies with its preceding host. In candidate (46d), the weak function word *-o* is mapped onto a PPhrase. This is obviously impossible, because according to the headedness condition, being a PPhrase entails being parsed into a PWord and being parsed into a PWord entails having foot and

this will violate HEAD-TO-LEX. In (46e) the PPhrase is not right-aligned with a PWord, and (46f) violates both ONSET and MATCH(ω,LEX).

(46)

Input: /zan + -o + mard/ woman + and + man (*lex* + *fnc* + *lex*) 'woman and man'	PARSE[NTO-φ	HEAD-TO-LEX	PARSE-EDGE-R	ONSET	MATCH(ω,LEX)
☞ a. $(((za)_{\omega}\ no)_{\omega})_{\varphi}\ ((mard)_{\omega})_{\varphi}$					*
b. $((zan)_{\omega})_{\varphi}\ (o\ (mard)_{\omega})_{\varphi}$				*!W	L
c. $((zan)_{\omega})_{\varphi}\ o\ ((mard)_{\omega})_{\varphi}$	*!W			*W	L
d. $((zan)_{\omega})_{\varphi}\ (o)_{\varphi}\ ((mard)_{\omega})_{\varphi}$		*!W		*W	*
e. $((za)_{\omega}\ no)_{\varphi}\ ((mard)_{\omega})_{\varphi}$			*!W		L
f. $((zan)_{\omega})_{\varphi}\ ((o\ (mard)_{\omega})_{\omega})_{\varphi}$				*!W	*

In tableau (47), ONSET is not violated by any of the candidates, and the ranking of MATCH(ω,LEX) which is lower than the other constraints as discussed in §3.4 will choose candidate (47a) as the output. In candidate (47b), the maximal PWord on the left does not match with a lexical word and thus violates MATCH(ω,LEX). Candidates (47c) and (47d) are very similar to (46c) and (46d) in the previous tableau.

(47)

Input: /zan + va + mard/ woman + and + man (*lex* + *fnc* + *lex*) 'woman and man'	PARSEINTO-φ	HEAD-TO-LEX	PARSE-EDGE-R	ONSET	MATCH(ω,LEX)
☞ a. $((zan)_{\omega})_{\varphi}$ $(va\ (mard)_{\omega})_{\varphi}$					
b. $(((zan)_{\omega}\ va)_{\omega})_{\varphi}$ $((mard)_{\omega})_{\varphi}$					*!W
c. $((zan)_{\omega})_{\varphi}$ va $((mard)_{\omega})_{\varphi}$	*!W				
d. $((zan)_{\omega})_{\varphi}$ $(va)_{\varphi}$ $(mard)_{\omega})_{\varphi}$		*!W			
e. $((zan)_{\omega}\ va)_{\varphi}$ $((mard)_{\omega})_{\varphi}$			*!W		
f. $((zan)_{\omega})_{\varphi}$ $((va\ (mard)_{\omega})_{\omega})_{\varphi}$					*!W

The same ranking of constraints can explain the prosodization of other conjunctions such as *yâ* 'or' as a proclitic, and some other weak function words such as the Ezafe morpheme as an enclitic. We will explore Ezafe constructions in detail in chapter 4.

3.6 Multiple function words

In Persian, more than one weak function word can occur before or after a lexical word. In cases with multiple proclitics, e.g., *az tu xune* 'from inside the house', we observe that none of the proclitics are stressed, suggesting that they do not form PWords on their own. Even in literary forms such as *vazấn* and *kazín*, which are contracted forms of *va az ân* 'and from that' and *ke az in* 'that from this' respectively, the whole contracted form is cliticized to the following *lex* without itself being stressed.[2] Also, in multiple proclitics, there is more than one possible syllabification, similar to what we saw in (24) and (26). For example, a noun preceded by two class 1 prepositions such as *joz az irấn* 'except from Iran', when produced without any focus, is a sequence of two proclitics and a host, and can be syllabified in four ways:

[2] Please also note that contracted forms like *kazín* or *kín* also support the proclitic nature of *ke* in such cases.

(48)

a. *jo.za.zi.rấn* b. *joz.(ʔ)a.zi.rấn* c. *jo.zaz.(ʔ)i.rấn* d. *joz.(ʔ)az.(ʔ)i.rấn*

The possibility of syllabifying each proclitic independently from the host and other proclitics suggests that neither of the proclitics is parsed into the PWord, a domain in which, only one specific syllabification is allowed. Thus, the structure for multiple proclitics would be multiple free clitics ($fnc_2\ fnc_1\ (lex)_\omega\ )_\varphi$. This is exactly what our ranking predicts. As can be seen in the tableau in (49), candidate (49a) is the only one which does not violate any of the constraints. Candidates (49b) and (49c) violate MATCH(ω,LEX) once and two times respectively. In candidate (49d) the head of a PWord is contained in a weak function word. In candidate (49e) the two function words are not a part of the PWord, but form a foot together. This structure also fails to satisfy HEAD-TO-LEX due to the definition in (12). Finally (49f) is eliminated, because it contains material unparsed into PPhrase.

(49)

Input: /az + tu + xune/ from + in + house (*fnc* +*fnc*+ *lex*) 'from inside the house'	PARSEINTO-φ	HEAD-TO-LEX	PARSE-EDGE-R	MATCH(ω,LEX)
☞ a. $(az.tu\ (xu.ne)_\omega\ )_\varphi$				
b. $(az\ (tu\ (xu.ne)_\omega\ )_\omega\ )_\varphi$				*!W
c. $((az\ (tu\ (xu.ne)_\omega\ )_\omega\ )_\omega\ )_\varphi$				**!W
d. $((az.tu)_\omega\ (xu.ne)_\omega\ )_\varphi$		*!W		
e. $((az.tu)_\pi\ (xu.ne)_\omega\ )_\varphi$		*!W		
f. $az.tu\ ((xu.ne)_\omega\ )_\varphi$	*!W			

So far, we have argued that enclitics in Persian are affixal clitics that adjoin to PWords in a recursive manner. In the case of multiple enclitics, e.g., *pedáreto* 'father-POSS2SG-ACC', there can be two possible structures:

(50) i. $(((lex)_{\omega} fnc_1 fnc_2)_{\omega})_{\varphi}$

ii. $(((lex)_{\omega} fnc_1)_{\omega} fnc_2)_{\omega})_{\varphi}$

In the relatively plain structure in (50i), there are only two levels of PWord, namely a minimal and a maximal one, while in the more nested structure in (50ii), there is an intermediate level PWord between the minimal and the maximal one. In fact, there seems to be neither a phonological process nor phonetic evidence in Persian supporting either of the structures in (50). However, the ranking we proposed in this study prefers (50i) over (50ii). Moreover, (50i) has simpler structure, and all other things being equal, should be preferred to (50ii), because (50ii) violates the general structure-prohibiting constraint *STRUC introduced in Prince and Smolensky (1993). In tableau (51), candidate (51b) is less harmonic than the winner candidate (51a), because there is one more PWord in (50ii) which is not matched by a *lex*. Candidates (51c) and (51d) violate HEAD-TO-LEX in a similar way as candidates (d) and (e) do in tableau (49). Candidate (51e) contains material unparsed into PPhrase and in candidate (51f), PPhrase is not aligned with a PWord.

(51)

Input: /pedar + -et + -o/ father +POSS2SG+ ACC (*lex* + *fnc* + *fnc*) 'yor fatherACC'	PARSEINTO-φ	HEAD-TO-LEX	PARSE-EDGE-R	MATCH(ω,LEX)
☞ a. $((((pe.da)_{\omega}\ re.to)_{\omega})_{\varphi}$				*
b. $((((pe.da)_{\omega}\ re)_{\omega}\ to)_{\omega})_{\varphi}$				**!W
c. $(((pe.da)_{\omega}\ (re.to)_{\omega})_{\varphi}$		*!W		*
d. $((((pe.da)_{\omega}\ (re.to)_{\pi})_{\omega})_{\varphi}$		*!W		*
e. $((pe.da)_{\omega})_{\varphi}$ re.to	*!W			L
f. $((pe.da)_{\omega}$ re.to $)_{\varphi}$			*!W	L

3.7 'Exceptionally stressed' words

Studies on word stress in Persian often list a group of lexical entries which are considered to be "exceptionally stressed" or to have "lexically non-final" stress. Ferguson (1957), Lazard (1992), Amini (1997), Mahootian (1997) among others provide such lists which mostly include initially stressed disyllabic discourse markers or particles. This section will try to shed further light on these exceptions and establish some regularities in them.

We should first make an important distinction between two different classes of these so-called 'exceptionally stressed' words. The first class comprises Arabic loanwords. Persian has borrowed a substantial portion of its lexicon from Arabic, and Arabic loans in Persian are predominantly nominal in nature (Versteegh, 2001). Arabic nouns and also adjectives and verb roots have been borrowed as lexical items into Persian, and naturally match with a PWord in prosodic representation in a similar way as other lexical items, and the word stress is on their rightmost syllable, no matter what their original prosodic structure was in the donor language. However, loan function words such as discourse markers, do not follow the Persian final stress pattern, since function words are essentially stressless in Persian. Interestingly, Arabic discourse markers when enter Persian, rather than becoming prosodically weak, seem to preserve their original stress pattern in Arabic. Some examples of these loanwords are given in (54).

(52)

i.	*lǻken/líken*	'but'	v.	*ámmâ*	'but'
ii.	*ǻmmin*	'Amen'	vi.	*háttâ*	'even'
iii.	*bále* [3]	'yes'	vii.	*ávvalan*	'firstly'
iv.	*éllâ*	'except'	viii.	*élâ*	'until'

According to Watson (2011: 3003) in Classical Arabic, disyllabic words are always initial stressed except when the second syllable is a super heavy (CVVC or CVCC). The

[3] From the Arabic word *bala*:

loanwords in (52) are all initial-stressed in Classical Arabic, the language which has been the source of borrowing for Persian.

The second class of exceptionally stressed words consists of lexicalized [host+enclitic] combinations. In these cases, the word had originally been a combination of a lexical word and its following weak function words, but diachronically has fossilized as a single lexical entry, however, for some reasons, still preserves its non-final stress pattern. A list of some lexicalized [host+enclitic] words is given in (53).

(53)

i.	*xéyli*	'very'	*xéyl* 'swarm' + indefinite *-i*
ii.	*básí*	'very much'	*bás* 'many' + indefinite *-i*
iii.	*ári*	'yes'	*âvár* 'certainty' (from Middle Persian *ēvar*[4]) + indefinite *-i*
iv.	*čéra*	'why'	*čé* 'what' + *-ra* (from Middle Persian *râd*/*rây* 'for'[5])
v.	*zírâ*	'because'	*az* 'from'+*ēd*[6] 'this' +*rāy* 'for' (*az-ēd-rāy*>*azērâ*> *zirâ*)
vi.	*čónke*	'because'	čón 'because' + the particle *ke*
vii.	*bálke*	'rather'	*bál* 'but' (from Arabic) + the particle *ke*
viii.	*báske*	'so much'	*bás* 'many'+ the particle *ke*
ix.	*káški*	'I wish'	*káš* 'I wish' + the particle *ke*

First, it should be made clear that these words have undergone some degree of lexicalization, and are adopted into the lexicon. As Wicka (2009) maintains, a lexicalized phrase takes on new semantics which cannot be derived from its original constituent parts. If we examine the lexicalized [host+enclitics] listed in (53), we can see that their meanings cannot be easily predicted by knowing the meanings of their constituents. Persian speakers intuitively treat them as single words, and they are listed as independent entries in

[4] For a detailed discussion refer to Sadeghi (1988)
[5] For the historical evolution of *-ra* see Dabirmoghaddam (2006)
[6] See Abolghassemi (2008)

dictionaries. Moreover, some of them have been borrowed into other languages as single words. For instance *čónke* and *kấški* have been borrowed into Turkish as *çünkü* and *keşke*. This phenomenon seems to be a common process in other languages as well. Wicka (2009) reports similar processes in Germanic languages. She calls this process *form fossilization* and defines it as the following:

(54) Form fossilization: a process through which a [host+clitic] sequence, originally the product of an active phonological process of cliticization, becomes lexically listed, at which point the [host+clitic] sequence is produced wholesale by the speaker.

Wicka (2009) argues that the frequency with which a particular phrase is attested relates to the question of fossilization. In other words, a phrase must be commonly attested in the spoken language, in order for it to become part of the lexicon as one fixed unit.

Wicka (2009) also states that after fossilization has occurred, it is likely that the affected form would initially still enjoy a strong semantic and syntactic connection to the non-fossilized active phrase. In fact Persian examples in (53) suggest that lexicalized forms can also maintain prosodic connections with their original non-fossilized forms. There is possibility that these prosodic connections gradually vanish in the future if the lexicalization process proceeds further, as is the case with semantic and syntactic connections. In (53iii) *ấri* 'yes' is initial stressed, but the more recent and informal form of this word *âré* is final stressed. Sadeghi (1988) considers such shifts in the stress locus, as an ongoing change in Persian.

It should be noted that there are still a few exceptionally stressed words that do not belong to either of the two groups introduced here, and the reason they are not final stressed is not clear and needs further studies.

3.8 Summary

In this chapter, we examined the possible foot structures for Persian words and concluded that a single iambic foot can best explain the prosodic prominence pattern at word level in this language. The foot structure in encliticized words can be derived by the constraint ALL-FT-R ranked lower than HEAD-TO-LEX and higher than PARSE-σ.

We also introduced existing proposals for the prosodic structure of function words in Persian, and showed that taking function words as amalgamated clitics and independent PWords are theoretically and empirically problematic. The proposal made in this paper is that enclitics are grouped with their hosts in a recursive structure, while proclitics are not parsed into PWords, and rather are immediately dominated by a PPhrase in a non-exhaustive structure. In recursive PWords, syllabification is obligatory in the Maximal PWord domain, and stress is also obligatory and culminative in this domain. Minimal PWords are always right-aligned with a *lex*, and a single iambic foot is also right-aligned with the minimal PWord.

We analyzed cases in which multiple *fnc*s attach to a host, and also introduced cases in which a *fnc* can prosodize with both its preceding and following lexical words. We suggested that the prosodic markedness constraint ONSET will be responsible to determine the direction for clitics to attach to their hosts in XP-external function words.

We also dealt with lexically non-final stressed words and argued that some of them are loanwords which have preserved their original stress pattern in the donor language, and some are historically lexicalized [host+enclitic] combinations.

Chapter 4 Phonological Phrase and above

4.1 Introduction

The previous chapter dealt with foot structure and Phonological Words in Persian. This chapter will explore syntactic phrases and their relation with Phonological Phrases. First we will introduce Kahnemuyipour's prosodic accounts for DPs and VPs which we adopt in this study, and then will focus on Ezafe constructions, examine the existing accounts of their prosodic status, and will provide a novel account. We will also investigate lexicalized Ezafe constructions and the relation between information structure and the prosody of Ezafe constructions briefly. Finally, we will aim at explaining the prosodic structure of Persian syntactic phrases including DPs, VPs and Ezafe constructions by an identical ranking of syntax-phonology interface constraints and prosodic markedness constraints.

4.2 Prosodic structure of DPs and VPs: a summary

Before discussing Ezafe constructions, we will summarize the prosodic structure of DPs and VPs in Persian based on Kahnemuyipour (2003). The structures in (1) and (2) show determiner phrases in Persian in which only the leftmost determiner is always prominent, and thus, the highest DP is mapped onto a PPhrase. Determiners such as *in* 'this' and *un* 'that' as discussed in chapter 3 are not prosodically weak, and belong to the lexical class of words, thus their maximal projections are eligible to map onto PPhrases.

(1) $[\text{D NP}]_{\text{DP}} \rightarrow (\text{D NP})_{\varphi}$ (*ín âdam*)$_{\varphi}$ 'this person'

(2) $[\text{D } [\text{D NP}]_{\text{DP}}]_{\text{DP}} \rightarrow (\text{D D NP})_{\varphi}$ (*ún do nafar*)$_{\varphi}$ 'those two people'

The structures in (3) and (4) are VPs containing an object NP and a verb, but the difference between (3) and (4) is that in (3) the object NP is indefinite/unspecific and thus combines with the verb to form a single VP, while in (4), the object NP is definite/specific and the

verb projects into a maximal projection, and combines with the object NP in a higher level. The view that definite/specific objects are syntactically in a higher position than nonspecific ones has been put forward in the literature for several languages including Persian.[1] Kahnemuyipour (2003) following Browning and Karimi (1994) and Ghomeshi (1996) suggests that the definite/specific object is in a VP-external position, and argues that the indefinite object and the verb are mapped onto the same PPhrase, while the definite/specific object and its following verb are mapped onto two distinct PPhrases.

(3) $[\text{NP V}]_{\text{VP}} \rightarrow (\text{NP V})_{\varphi}$ (*qazá́ xord*)$_{\varphi}$ 'ate food'

(4) $[\text{NP VP}]_{\text{VP}} \rightarrow (\text{NP})_{\varphi}\,(\text{VP})_{\varphi}$ (*qazá́-râ*)$_{\varphi}$ (*xórd*)$_{\varphi}$ 'ate the food'

The structure in (5) is the case of VPs in which the verb has stressed prefixes marking mood. These prefixes are prosodically strong and as Kahnemuyipour (2003) suggests, they form an autonomous PWord and phrase with the following verb into the same PPhrase. These lexical items include frequently used indicative marker *mi-* and the subjunctive marker *be-*.

(5) $[\textit{Mood}\ \text{V}]_{\text{VP}} \rightarrow (\textit{Mood}\ \text{V})_{\varphi}$ (*mí-xord*)$_{\varphi}$ 'was eating'

To sum up, DPs and VPs in Persian have only one prosodic prominence on their leftmost PWord. According to Kahnemuyipour (2003), in DPs and VPs the highest maximal projection is mapped into a PPhrase and the PPhrase-level stress rule puts the stress on the leftmost PWord of a PPhrase. Then the IPhrase-level stress rule assigns the IPhrase stress to the rightmost PPhrase. After PPhrase and IPhrase stresses are assigned, the leveling rule will delete the stress of all non-head PWords. As a consequence, there will be only two levels of audible prominence, that of PPhrase stress, and that of IPhrase stress.

[1] See Diesing (1992) for Turkish, Mahajan (1990) for Hindi, Koopman and Sportiche (1991) for Dutch, and Moorcroft (1995) for Icelandic and German.

4.3 Ezafe Constructions

Persian is considered to be a "mixed-headed" SOV language in which both head-initial and head-final phrases can be observed (Zepter, 2003). Verb phrases are always head-final, while in non-verbal categories, a head can have both preceding and following modifiers. When the head is followed by certain complements and modifiers, an unstressed morpheme /-e/ appears between the head and its following material. The term used for this morpheme traditionally has been *Ezafe*, a loanword from Arabic /idˤaːfa/ literally meaning "adding". It generally appears between any two items that have some sort of connection (Ghomeshi, 1996). Ezafe is semantically vacuous and is cliticized to its preceding lexical word as an affixal clitic (Hosseini, 2012b). Ezafe can appear in the following contexts:

1. between a noun and a modifier:

âdam-e bad
person-EZ bad
'a bad person'

2. between a noun and a possessor:

pedar-e dâvud
father-EZ PR
'Davud's father'

3. between a noun and its complement:

xaridan-e nân
buy-EZ bread
'buying bread'

4. between an adjective and its complement:

nârenji-ye rowšan
orange-EZ light
'light orange'

5. between a class 2b preposition and its complement:

nazdik-e *bâzâr*

near-EZ market

'near the market'

Ezafe is a feature of certain Western Iranian languages such as Persian and Kurdish (Samvelian, 2007). It is also present in Urdu, a language highly influenced by Persian.

4.3.1 Ezafe constructions as single PWords

The syntax of Ezafe constructions has been a controversial matter in Persian linguistics and many proposals have been made to explain their structure. Ghomeshi (1996) following Samiian (1983) proposes that Ezafe constructions are formed by base-generated X^0 adjunction, and thus nouns and adjectives are not projected in these constructions. Consequently, all the elements in the Ezafe domain are X^0's, and the Ezafe construction cannot be considered as an XP. Kahnemuyipour (2000) in a minimalist approach adopts this idea and suggests that in an Ezafe construction, the adjectives (or modifiers) are located in the heads of functional projections above NP. These adjectives (or modifiers) bear the feature [Mod] (for modifier), and the functional projections are thus called Mod(ifier) P(hrase)s. The noun, which also has the feature [Mod] (morphologically realized by the Ezafe vowel), moves up and head-adjoins to the adjective, and checking takes place. This is shown in figure 1 adopted from Kahnemuyipour (2003). As can be seen, all the elements in the Ezafe domain are X^0's, and the final structure of this phrase (circled in the tree diagram) is an X^0-level element.

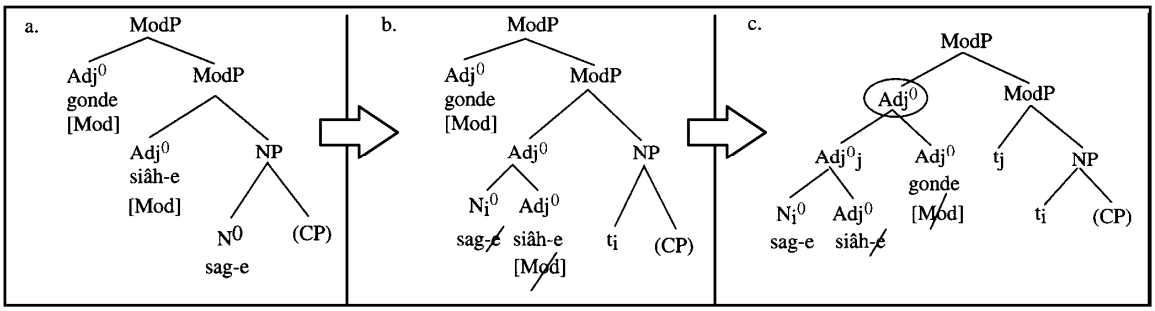


Figure 1 Ezafe construction as non-projecting X^0's, adopted from Kahnemuyipour (2003)

In this analysis, possessor nouns which always occur at the right edge of Ezafe constructions behave differently from other nouns and adjectives. Possessors are maximal projections and always coincide with a PPhrase. We will discuss possessors separately in §4.3.4 below.

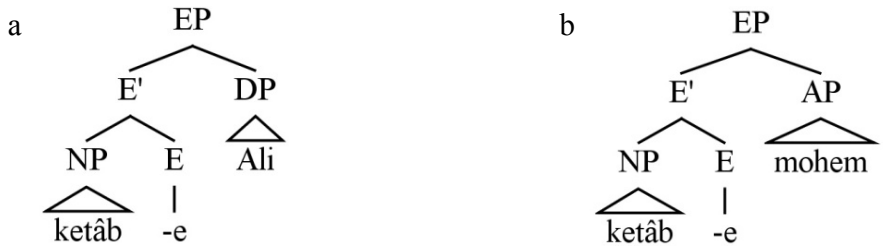
Many researchers such as Eslami (2000) and Kahnemuyipour (2003) claim that in an Ezafe construction, the last lexical word is prosodically more prominent than the others in the construction. In (6), an example from Kahnemuyipour (2003), the last syllable of the rightmost lexical word (*gonde*) is perceived more prominently. Kahnemuyipour (2003) adopts the above mentioned syntactic analysis that takes the whole Ezafe construction to be an X^0 element. Since Ezafe constructions have rightmost prosodic prominence, and since they are syntactically X^0-level elements, Kahnemuyipour (2003) proposes that the whole Ezafe construction should be treated as a single PWord.

(6) sag-e siyâh-e gonde → (*sag-e siyâh-e gondé*)$_\omega$
dog-EZ black-EZ big
'big black dog'

Ito and Mester (2012: 297) based on Kahnemuyipour's analysis, suggest that the Ezafe construction in Persian can be regarded as a recursive PWord rather than a plain one.

4.3.2 Ezafe constructions as projections of the functional head Ezafe

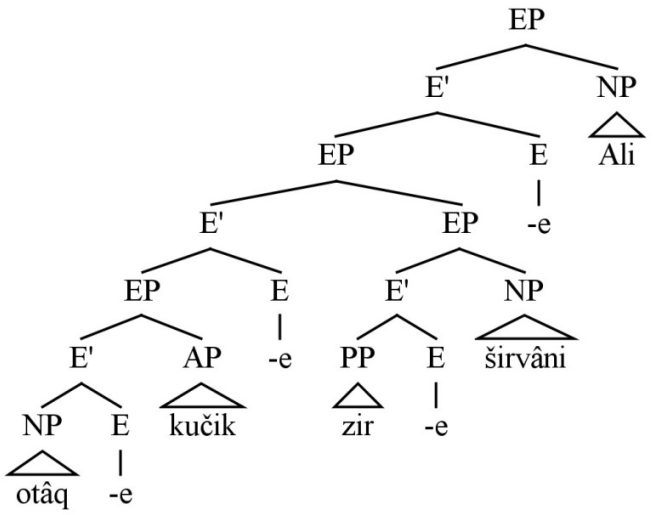
Mahootian (1993) for the first time suggests that the whole Ezafe construction is a phrase which she calls Ezafe Phrase (EP), with the Ezafe morpheme as its head. Ezafe is a functional head, and EP is a functional category. Her analysis of simple Ezafe constructions is shown in figure 2.



ketâb-e	*ali*	*ketâb-e*	*mohem*
book-EZ	PR	book-EZ	important
'Ali's book'		'important book'	

Figure 2 Mahootian's analysis of Ezafe constructions adopted from Mahootian (1993)

Mahootian (1993) provides a similar analysis for complex Ezafe constructions in which one Ezafe construction is nested in another. An example of such constructions is given in figure 3.

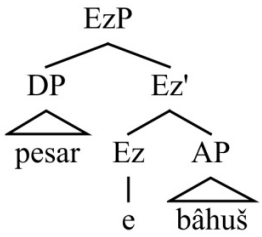


otâq-e	*kučik-e*	*zir-e*	*širvâni-ye*	*ali*
room-EZ	small-EZ	under-EZ	roof-EZ	PR

'Ali's small room under the roof'

Figure 3 Mahootian's analysis of a complicated nested Ezafe construction adopted from Mahootian (1993)

Moinzadeh (2005) in his minimalist approach applies Mahootian's analysis and takes Ezafe constructions as a functional category (EzP), but based on syntactic evidence, suggests that the head of the EzP is the morpheme /-e/ with its complement to the right.



pesar-e	*bâhuš*
boy-EZ	intelligent

'the intelligent boy'

Figure 4 Moinzadeh's analysis of Ezafe constructions adopted from Moinzadeh (2005)

Larson and Yamakido (2005) take Ezafe to be a case-marker very similar to the particle -*no* in Japanese. Butt et al. (2008) also propose an analysis for Ezafe constructions in Urdu which is very similar to that in Moinzadeh (2005).

4.3.3 Prominence pattern of Ezafe constructions

The account in Ghomeshi (1996) and Kahnemuyipour (2000) that takes the Ezafe construction to be a string of non-projected X^0's may be syntactically grounded, but is not compatible with phonological facts of the language. As we saw in chapters 1 and 3, Persian prominence is culminative at word level, i.e., there can be only one prominent syllable in a PWord, and there are no secondary stresses at word level in Persian. However, in case of Ezafe constructions, there are audible prominences on the last syllable of every lexical word present in the construction. For instance, in (6), there are audible prominences on the final syllables of *sag-e* and *siyâh-e* as well as the final word *gonde*. As mentioned in §4.3.1, Eslami (2000) and Kahnemuyipour (2003) claim that the final word in an Ezafe construction is perceived more prominently than the others. In fact, their observation is accurate, but the actual reason for this extra prominence is not what they claim. We will argue in the following sections that each lexical word in an Ezafe construction is in fact mapped onto a PPhrase. Provided the fact that only PPhrase and IPhrase level prominences are audible in Persian, if we utter an Ezafe construction such as (6) in isolation, the Ezafe construction will form an IPhrase and the final PPhrase of this IPhrase (the word *gonde*) will be perceived more prominently. Thus, the extra prominence of the final word in an Ezafe construction is due to IPhrase-level prominence which is associated to the rightmost PPhrase. In other words, the final word in the Ezafe construction is perceived more prominently when it is uttered in isolation, only because it bears the final audible accent of the utterance. In fact, this claim is very easy to justify: if we put a phrase like (6) in a non-final position of a carrier sentence, the final word of the construction (*gonde*) will not be perceived more prominently. This is shown in (7) in which the IPhrase-level prominence

will be associated to the intransitive verb *mí-raqs-e* and all the three words in the Ezafe construction will have the same degree of prominence.

(7) *sàg-e siyâh-e gondè mí-raqs-e*
dog-EZ black-EZ big INDI-dance-3SG
'The big black dog is dancing.'

Treating Ezafe constructions as recursive PWords is also not supported by the data, because in that case all PWords except the minimal one will have more than one audible prominence which is not acceptable for a PWord in Persian. The fact that each lexical word in an Ezafe construction has its own audible prominence is readily observable in the pitch contour of utterances of these constructions. This can be seen in figure 5 which shows an utterance of the sentence in (8) uttered by a female speaker. The Ezafe construction in (8) was embedded in the carrier sentence shown in (9).

(8) *barâdar-e ârum-e moʔaddab-e leylâ*
brother-EZ calm-EZ polite-EZ PR
'Leyla's calm and polite brother'

(9) *diruz* _____ *umad*
yesterday come PAST3SG
'_____came yesterday.'

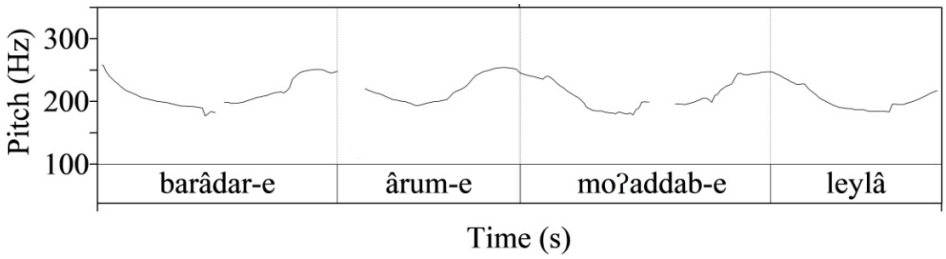


Figure 5 An utterance of the sentence in (8) uttered by a female speaker

As discussed in details in chapter 3, the heads of PWords are the final syllables of their lexical words, thus one may expect to observe the f0 peaks on the *lex*-final syllables. However, as can be seen in figure 4, the actual peaks are realized on the final syllables of maximal PWords which are prosodically weak enclitics. This is due to the '*peak delay*' phenomenon which has been proven to occur in various languages. As shown and discussed for English by Silverman and Pierrehumbert (1990), non-final peaks often occur after the syllable with which they are intuitively associated. Therefore, in the words with penultimate stress, the f0 peaks will occur on the rightmost non-stressed syllables. Peak delays on non-IPhrase-final PPhrases will be discussed in more detail in chapter 5. In sum, actual prosodic pattern of Ezafe construction suggest that they cannot be regarded as neither recursive nor plain PWords.

4.3.4 Are possessors prosodically different?

Another issue with Kahnemuyipour's account of Ezafe constructions is the prosodic distinction he makes between the Ezafe constructions which contain a possessor and those which do not. He argues that whereas all the other elements in Ezafe construction are X^0's, the possessor is an XP. The syntactic argumentation for this claim is that in Ezafe constructions, only possessors can be expanded into more complex phrases such as complex DPs or CPs. The left-alignment rule will derive the prosodic structure of an Ezafe phrase containing a possessor as shown in example (10) taken from Kahnemuyipour (2003: 368).

(10)

( )$_{\varphi}$ ( )$_{\varphi}$

[$_{DP}$N^0-EZ Adj^0-EZ NP (POSS)]

ketâb-e *jâleb-e* *ali*

book-EZ interesting-EZ PR

'Ali's interesting book'

This analysis will predict different prominence patterns for the Ezafe constructions that contain a possessor and those that do not contain a possessor. For example, in the two Ezafe constructions in (11), the word *zibâ* has been used in two different senses. In (11i) *zibâ* is an adjective meaning 'beautiful' and the Ezafe construction does not contain a possessor. In Kahnemuyipour's analysis, this Ezafe construction is mapped onto a single PWord which is parsed into a PPhrase at a higher level as shown in (12i). This structure will predict a single audible prominence on the rightmost syllable of the whole construction. In (11ii) however, *zibâ* is used as a proper feminine name and appears as a possessor in the Ezafe construction. Kahnemuyipour's analysis will predict that the possessor maps onto a PPhrase, while the two other words *xune-ye* and *bozorg-e* are first mapped onto a single PWord, since they are non-projected x^0 material, and then, this PWord will be parsed into a PPhrase. Therefore, this approach predicts that there will be two prominent points in (11ii), one on the rightmost syllable of the string *xune-ye bozorg-e*, and one on the second syllable of the possessor *zibâ* as shown in (12ii)

(11)

i. *xune-ye bozorg-e zibâ*
house-EZ big-EZ beautiful
'the big beautiful house'

ii. *xune-ye bozorg-e zibâ*
house-EZ big-EZ PR
'Ziba's big house'

(12)

i. ((*xune-ye bozorg-e zibá*)$_{\omega}$)$_{\varphi}$

ii. ((*xune-ye bozórg-e*)$_{\omega}$)$_{\varphi}$ ((*zibá*)$_{\omega}$)$_{\varphi}$

However, this prediction is inconsistent with the attested prosodic prominence patterns of the two cases. Both (11i) and (11ii) have three prominent points (pitch peaks) on the last syllables of their lexical words which renders the utterances of (11i) and (11ii) absolute homophones. Figure 6a shows the pitch contour of an utterance of the Ezafe construction in (11i) uttered by a native speaker, and figure 6b demonstrates the pitch contour of an utterance of the Ezafe construction in (11ii) uttered by another speaker. Both figures 6a and 6b have been uttered in the carrier sentence in (13):

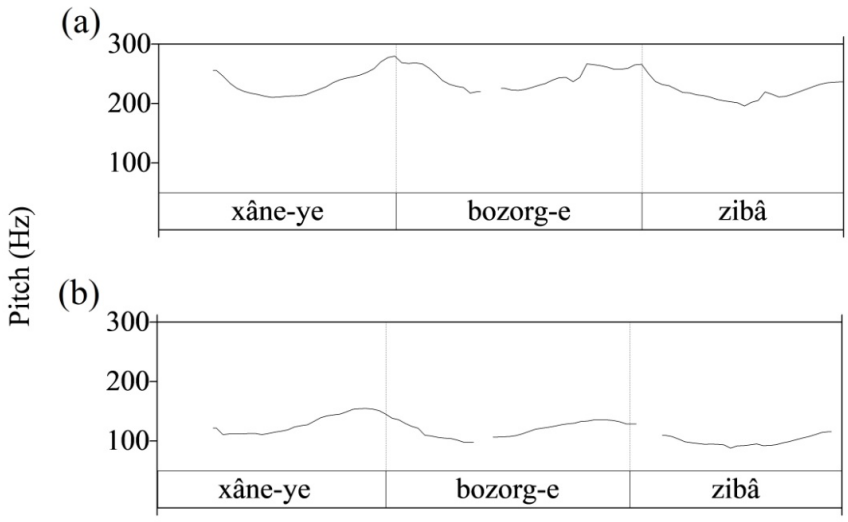


Figure 6 (a): pitch contour of an utterance of the sentence in (11i) uttered by a female speaker, and (b): pitch contour of an utterance of the sentence in (11ii) uttered by a male speaker

(13)

_____ *injâ-st*

'_____ is here.'

Figure 6 shows that there are no prosodic differences between the Ezafe constructions containing a possessor and those which do not contain.

In an Ezafe construction, each lexical word has an audible prominence on its final syllable. However, there are cases in which the Ezafe construction has a single prominence on its rightmost syllable. These cases are Ezafe constructions which have undergone some degree of lexicalization due to their high frequency of co-occurrence.

4.3.5 Lexicalized Ezafe Constructions[2]

Some Ezafe constructions with high frequency of co-occurrence have diachronically combined with each other and formed compounds. In this process, a syntactic construction becomes fossilized and forms a lexical item which maps onto a single PWord in the same way as other compounds. Lexicalization, as argued in chapter 3, is a gradual and gradient process, thus we can find Ezafe constructions which are lexicalized more or less than the others. The example in (14) is highly lexicalized, and is always uttered with a single prominence on the final syllable.

(14)

toxm-e mórq
egg-EZ hen
'egg'

On the other hand, for instance in (15), the phrase is often uttered with only a single prominence on the second word *fârs* as shown in (15i), however an utterance with audible prominence on each of the words is also possible as shown in (15ii). Nevertheless, the former version which treats the whole Ezafe construction as a single PWord is more common, probably due to the frequent usage of the construction as a place-name.

[2] Some of the ideas and examples of this section are taken from Solhju et al. (2003)

(15)

i.	*xalij-e*	*fárs*	ii.	*xalíj-e*	*fárs*
	gulf-EZ	Persia		gulf-EZ	Persia
	'Persian Gulf'			'Persian Gulf'	

In (16) however, the Ezafe construction is uttered with each word having an audible prominence, which shows that this construction is far from being lexicalized. In fact, a pronunciation in which there is only one prosodic prominence on the last syllable of the second word might also be possible for (16), but is acceptable only in the contexts in which 'Gulf of Mexico' is regarded as an old (given) information. Deaccentation as an effect of information structure is a productive process and differs from lexicalization and will be addressed in §4.7 in this chapter.

(16)

xalíj-e	*mekzík*
gulf-EZ	Mexico
'Gulf of Mexico'	

One further point about lexicalization is that it seems that if the lexicalization progresses further, the Ezafe linker may disappear granted that its deletion does not lead to a phonologically ill-formed structure. The examples in (17) illustrate such cases.

(17)

i.	pedár-e	zán	→	*pedarzán*
	father-EZ	wife		'Father in law'
ii.	tút-e	farangí		*tutfarangí*
	berry-EZ	European	→	'strawberry'

4.3.6 Prosodic structure of Ezafe constructions

We adopt the analysis of Persian prosodic structure in Kahnemuyipour (2003) in which there are only two levels of audible prominence which he calls "PPhrase stress" and "IPhrase stress". PPhrase stress is assigned to the head PWord, and IPhrase stress is assigned to the rightmost PPhrase head. Thus, the words that bear audible prominence are necessarily PPhrase heads. If we look at Ezafe constructions closely, since every word in an Ezafe construction has an audible prominence, each of them must be a PPhrase head. In other words, each lexical word in an Ezafe construction must be located at the leftmost edge of some PPhrase. Therefore, for an Ezafe construction containing three lexical words there would be three possible prosodic structures that can explain its prosodic pattern.

The structure shown in (18i) for an Ezafe construction consisting of three lexical elements, is a recursive structure in which, each word present in the Ezafe construction is the leftmost PWord (the head) of a PPhrase. The other possibility is that each word forms a separate PPhrase as shown in (18ii), and the third structure is a combination of (18i) and (18ii) with one autonomous PPhrase on the left, and a recursive one on the right.

(18)

i. (*ság-e* (*siyáh-e* (*gondé*)$_\varphi$)$_\varphi$)$_\varphi$

ii. (*ság-e*)$_\varphi$ (*siyáh-e*)$_\varphi$ (*gondé*)$_\varphi$

iii. (*ság-e*)$_\varphi$ (*siyáh-e* (*gondé*)$_\varphi$)$_\varphi$

Recursive PPhrases have been shown to exist in several languages among which one can name Truckenbrodt (1999) for Kimatuumbi, Gussenhoven (2005) for English and Dutch, Selkirk and Kawahara (2005), Ishihara (2007), Selkirk (2009) and Ito and Mester (2012) for Japanese and Féry (2009) for German. In most of these languages, there are independent phonological or phonetic evidence supporting the recursive structure. For example, in Japanese recursive PPhrases, cumulative rise of the pitch marks the left edge of

maximal PPhrases, and a metrical boost marks the left edge of Intermediate Phrases (Ito and Mester, 2012). However, in Persian, there seems to be no such phonetic evidence available supporting a recursive structure for Ezafe constructions.

More importantly, the assumption in Prosodic Phonology is that only syntactic maximal projections of lexical elements can be mapped onto PPhrases, and functional projections are not allowed to coincide with PPhrases (Selkirk; 1995, 2011). This fact is stated as 'Lexical Category Condition' in Truckenbrodt (1999) and is as follows:

(19) Lexical Category Condition (LCC):

> Constraints relating syntactic and prosodic categories apply to lexical syntactic elements and their projections, but not to functional elements and their projections, or to empty syntactic elements and their projections.

The reason why in some languages syntactic maximal phrases do not precisely correspond to PPhrases, is that some kinds of prosodic markedness constraints outrank interface constraints that map syntactic phrases onto PPhrases. In the languages that prosodic and syntactic structures match with each other, interface constraints are ranked over prosodic markedness constraints. Therefore, in order to have a recursive structure as in (18i) firstly, a corresponding recursive syntactic structure similar to the one shown in (20i) is necessary, in which a lexical maximal projection (hereafter XP) is embedded in another XP. Secondly, the prosodic markedness constraints need to be ranked lower than syntax-prosody interface constraints, so that recursive syntactic structures be mapped onto recursive PPhrases.

(20)

i. [$_{XP}$ *sag-e* [$_{XP}$ *siyâh-e* [$_{XP}$ *gonde*]]]

ii. [$_{XP}$ *sag-e*] [$_{XP}$ *siyâh-e*] [$_{XP}$ *gonde*]

The syntactic structure in (20i) is not acceptable for a Persian Ezafe construction, because there is no syntactic evidence suggesting that Ezafe constructions in Persian form recursive

XPs similar to the structure shown in (20i). On the contrary, substantial evidence has been provided in the literature supporting the fact that Ezafe constructions cannot be regarded as either plain or recursive XPs. If we summarize the syntactic structures proposed far for Ezafe constructions, Ghomeshi (1996) and Kahnemuyipour (2000) among others suggest that lexical elements in an Ezafe construction are non-projected X^0 words, and other researches such as Mahootian (1993), Moinzadeh (2005), and Butt et al. (2008) propose that Ezafe constructions are projections of Ezafe as a functional head. Thus, a recursive PPhrase cannot emerge as a result of syntax-prosody interface constraints. This is also true about the structure in (18iii) in which the recursive PPhrase on the right cannot be formed by syntax-prosody mapping.

On the other hand, to have a structure similar to the one in (18ii), we will only need each lexical element in an Ezafe construction to be a lexical maximal projection. This is acceptable if we postulate a syntactic structure in which the Ezafe morpheme is a functional syntactic head and the lexical element preceding or following Ezafe projects to an XP to combine with the functional head. This is in accordance with MAXIMALITY constraint in X-bar theories that requires all non-head daughters to be an XP (Kornai and Pullum, 1990). In fact, this analysis is identical to the approach taken in Mahootian (1993), Moinzadeh (2005), and Butt et al. (2008). As we observed in the examples in figures 2 and 3, the Ezafe morpheme as a functional head combines with its preceding or following lexical element which has projected to an XP. We suggest that the Ezafe morpheme phrases with its preceding element, rather than the following one, for purely phonological reasons. As discussed in chapter 3, if the Ezafe morpheme phrases as a proclitic, it will violate the constraint ONSET, and it is this constraint that militates against a procliticized Ezafe, and prefers this morpheme to be an enclitic.

Therefore, the general prosodic structure of an Ezafe construction shown in (21) would be as follows:

(21) *lex*-EZ *lex*-EZ… *lex* → $(lex\text{-EZ})_\varphi$ $(lex\text{-EZ})_\varphi$ … $(lex)_\varphi$

This analysis is consistent with any syntactic account of Ezafe constructions in which Ezafe constructions are considered to be projections of the functional head Ezafe.

According to our proposal, the prosodic structure of the example in figure 3 taken from Mahootian (1993) would be the structure shown in (22):

(22) (*otâq-e*)$_{\varphi}$ (*kučik-e*)$_{\varphi}$ (*zir-e*)$_{\varphi}$ (*širvâni-e*)$_{\varphi}$ (*ali*)$_{\varphi}$

Recall from chapter 3 that the Ezafe morpheme being an enclitic, is an affixal clitic which attaches to its preceding host in a recursive manner as explained in chapter 3, but the internal structure of the PPhrases has not been shown here for convenience sake.

If we accept that words bearing audible prominences are the heads of some PPhrase, then we can conclude that in complex Ezafe constructions, similar to the one shown in figure 3, each lexical word in the construction is in fact a PPhrase, because each lexical word bears an audible prominence. If we observe the pitch contour of an utterance of the sentence in (22), which is illustrated in figure 7, we can see that each *lex* has an observable peak which is realized on its rightmost syllable.

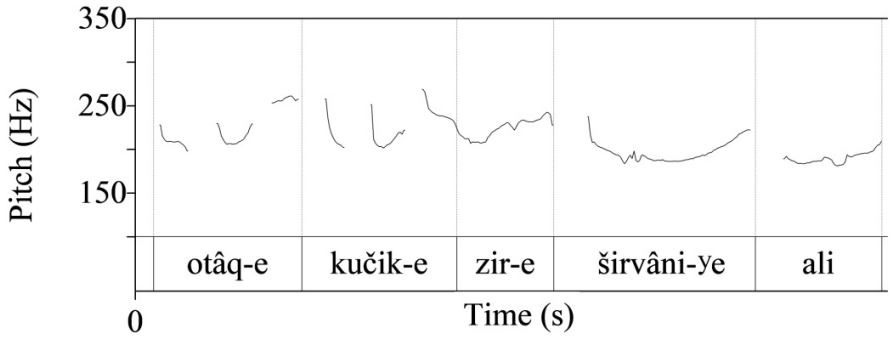


Figure 7 Pitch contour of an utterance of the sentence in (22)

Observable f0 peaks on each lexical element in the above example supports the idea that each lexical element in an Ezafe construction is mapped onto a PPhrase. In the next section we will try to provide an OT account for this mapping.

4.4 Syntax-Phonology interface and OT constraints

The goal of this section is to explain the prosodic structure of DPs and VPs on the one hand, and that of Ezafe constructions on the other hand uniformly, using an identical ranking of syntax-prosody interface constraints and prosodic markedness constraints. The structure of DPs and VPs and their corresponding prosodic structures are repeated in (23).

(23)

i.	$[\text{D NP}]_{\text{DP}}$	→ $(\text{D NP})_{\varphi}$	$(\textit{ín âdam})_{\varphi}$	'this person'
ii.	$[\text{NP V}]_{\text{VP}}$	→ $(\text{NP V})_{\varphi}$	$(\textit{qazá́ xord})_{\varphi}$	'ate food'
iii.	$[\text{NP VP}]_{\text{VP}}$	→ $(\text{NP})_{\varphi}\,(\text{VP})_{\varphi}$	$(\textit{qazá́-ro})_{\varphi}\,(\textit{xórd})_{\varphi}$	'ate the food'
iv.	$[\textit{Mood}\ \text{V}]_{\text{VP}}$	→ $(\textit{Mood}\ \text{V})_{\varphi}$	$(\textit{mí xord})_{\varphi}$	'was eating'

4.4.1 The problem of XP-external clitics

As we discussed in chapter 3, XP-external clitics are weak function words that are not a part of an adjacent XP, but prosodically tend to phrase with one of them. The constructions in (24i-iii) are examples of XP-external clitics introduced in chapter 3 and in this chapter. The structure in (24i) shows an XP followed by a weak functional word which is not syntactically a part of the XP but phrases with it to the same PPhrase, and (24ii) is the mirror image structure of (24i). Finally the structure in (24iii) is the Ezafe construction discussed in this chapter. The syntactic structure of Ezafe constructions adopted here is the one proposed by Mahootian (1993), Moinzadeh (2005) and others.

(24)

i.	XP *fnc*	→ $(\text{XP}\ \textit{fnc})_{\varphi}$	$(\textit{zán-o})_{\varphi}$	'woman and'
ii.	*fnc* XP	→ $(\textit{fnc}\ \text{XP})_{\varphi}$	$(\textit{yâ márd})_{\varphi}$	'or man'
iii.	XP EZ XP	→ $(\text{XP-EZ})_{\varphi}\,(\text{XP})_{\varphi}$	$(\textit{pedár-e})_{\varphi}\,(\textit{ârmín})_{\varphi}$	'Armin's father'

Since in the case of XP-external clitics, as can be seen in (24i-iii), the left and the right edges of XPs do not coincide with the left and the right edges of PPhrases, these structures

will cause problems to any theory that suggests exact alignment or matching between XPs and PPhrases. In the next sections, we will introduce two theories proposed in the framework of Prosodic Phonology, and point out their problems in dealing with XP-external clitics.

Alignment

In her end-based theory of the syntax-prosody relation, Selkirk suggests that interface constraints demand alignment of the left or right edges of XPs with those of Phonological Phrases (Selkirk, 1986, 1995). This theory has been widely used in the past two decades. The constraints responsible for this alignment are given in (25) and (26):

(25) ALIGN(XP, R; φ, R) (ALIGN-XP, R): Assign one violation mark for each XP whose right edge does not coincide with the right edge of a PPhrase.

(26) ALIGN(XP, L; φ, L) (ALIGN-XP, L): Assign one violation mark for each XP whose left edge does not coincide with the left edge of a PPhrase.

Kahnemuyipour (2003) adopts Selkirk's edge alignment theory and argues that in Persian XPs are left-aligned with PPhrases. One piece of evidence he uses in his discussion is the existence of the XP-external clitic conjunction -*o* which is not a part of its preceding NP, but prosodizes as an enclitic with it as shown in (24i). He argues that postulating right-alignment of XPs with PPhrases in Persian will lead to the ill-formed structure in (27):

(27) NP o NP → * $(\text{NP})_{\varphi}$ $(\text{o NP})_{\varphi}$

He does not deal with proclitic conjunctions of Persian such as *va* 'and' or *yâ* 'or', but if we take XP-external proclitics into consideration, left-alignment will also fail to derive the right prosodic structure:

(28)

i. NP *va* NP → * $(\text{NP } va)_{\varphi}$ $(\text{NP})_{\varphi}$

ii. NP *yâ* NP → * (NP *yâ*)$_{\varphi}$ (NP)$_{\varphi}$

Therefore, as can be seen in (27) and (28), neither left-alignment nor right-alignment can explain the prosodic structure of the phrases in (24). As we discussed in chapter 3, if only XPs are mapped onto PPhrases, the prosodic markedness constraint ONSET will determine whether XP-external clitics are enclitics or proclitics. The tableau in (29) shows that left-alignment derives the prosodic structure of the XP-external enclitic -*o* correctly, but fails to predict the structure of XP-external proclitics.

(29)

Input: [XP] & [XP]	ALIGN-XP, L	ALIGN-XP, R
☞ a. (NP *o*)$_{\varphi}$ (NP)$_{\varphi}$		*
💣 b. (NP *va*)$_{\varphi}$ (NP)$_{\varphi}$		*

The tableau in (30) on the other hand, shows that right-alignment will predict the structure of XP-external proclitics, but fails to derive the structure of the XP-external enclitics.

(30)

Input: [XP] & [XP]	ALIGN-XP, R	ALIGN-XP, L
☞ a. (NP)$_{\varphi}$ (*va* NP)$_{\varphi}$		*
💣 b. (NP)$_{\varphi}$ (*o* NP)$_{\varphi}$		*

Please note that placing the constraint ONSET over ALIGNMENT will require all words present in a PPhrase to obligatorily syllabify with each other.

Match theory

Selkirk (2011) in her Match Theory of the syntax-prosodic constituency, proposes that interface constraints call for a match between syntactic and prosodic constituents. She formalizes the tendency of XPs to match with PPhrases in a syntax-prosody interface constraint, namely, MATCH-XP:

(31) MATCH-XP: assign one violation mark for every syntactic XP that does not correspond to a PPhrase.

MATCH-XP can be interpreted as simultaneous right and left alignments of XPs with PPhrases, and it makes no preferences for a single edge alignment, therefore it seems more appropriate for a language like Persian in which both XP-external proclitics and enclitics are found. The constraint PARSE-INTO-φ introduced in (36) chapter 3, was shown to be an undominated constraint in Persian. In fact, ranking this constraint over MATCH-XP, seems to be able to derive structures that contain two XPs and one XP-external clitic. The tableau in (33) demonstrates this ranking and its prediction for cases such as (32i) and (33ii):

(32)

i.	pedar	e	ârmin →	$(pe.da.re)_{\varphi}(âr.min)_{\varphi}$	ii.	zan	o	mard→	$(za.no)_{\varphi}\ (mard)_{\varphi}$
	father	EZ	PR	'Armin's father'		woman	and	man	'woman and man'
	[XP]	*fnc*	[XP]			[XP]	*fnc*	[XP]	

(33)

Input: [XP] *fnc* [XP]	PARSE-INTO-φ	MATCH-XP
☞ a. $(\text{XP}\ fnc)_{\varphi}\ (\text{XP})_{\varphi}$		*
b. $(\text{XP})_{\varphi}\ fnc\ (\text{XP})_{\varphi}$	*!W	L
c. $(\text{XP}\ fnc\ \text{XP})_{\varphi}$		**!W

Candidate (33a) violates MATCH-XP once, because the XP on the left is not matched with a PPhrase. Candidate (33b) does not satisfy the higher-ranked PARSE-INTO-φ, and candidate (33c) is eliminated, because neither of the XPs matches with a PPhrase.

The same ranking will derive the prosodic structure of XP-external proclitics such as the one shown in (34). This is shown in the tableau in (35).

(34)

zan	va	mard→	$(zan)_{\varphi}$ $(va\ mard)_{\varphi}$
woman	and	man	'woman and man'
[XP]	*fnc*	[XP]	

(35)

Input: [XP] *fnc* [XP]	PARSE-INTO-φ	MATCH-XP
☞ a. $(XP)_{\varphi}$ $(fnc\ XP)_{\varphi}$		*
b. $(XP)_{\varphi}$ *fnc* $(XP)_{\varphi}$	*!W	L
c. $(XP\ fnc\ XP)_{\varphi}$		**!W

However, if there are two XPs followed or preceded by XP-external function words, the ranking in (33) and (35) would not be able to derive the correct prosodic structure anymore. This is shown in the tableau in (36) for multiple XP-external enclitics, but the same problem exists with proclitics as well. The example in tableau (35) can be an Ezafe construction or three lexical words conjoined by the conjunction *-o*.

(36)

Input: [XP] *fnc* [XP] *fnc* [XP]	PARSE-INTO-φ	MATCH-XP
☞ a. $(XP\ fnc)_{\varphi}$ $(XP\ fnc)_{\varphi}$ $(XP)_{\varphi}$		**
b. $(XP)_{\varphi}$ *fnc* $(XP)_{\varphi}$ *fnc* $(XP)_{\varphi}$	*!W	L
c. $(XP\ fnc\ XP\ fnc\ XP)_{\varphi}$		***!W
💣 d. $(XP\ fnc\ XP\ fnc)_{\varphi}$ $(XP)_{\varphi}$		**

Candidate (36a) violates MATCH-XP two times, because it contains two XPs that are not matched with any PPhrases. Candidate (36b) fails to satisfy PARSE-INTO-φ, and in candidate (36c), neither of the XPs is matched with a PPhrase. The crucial case is candidate (36d) which violates MATCH-XP the same times as the winner. In fact, if one accepts the OT constraints to be gradient, it may be possible to argue that candidate (36d) is less harmonic than candidate (36a), because the mismatch between the XPs and the PPhrases is

more severe in candidate (36d) comparing to that in candidate (36a). However, as we argued in chapter 1, following McCarthy (2003), this study does not postulate gradient constraints due to their ad hoc nature, and also due to the wrong predictions they might make.

Therefore, it is problematic to use MATCH-XP for case in which XPs are prosodically phrased with their adjacent XP-external weak function words, because in these structures shown in (37), XPs are not actually matched by any PPhrase:

(37)

i. *fnc*[XP] → (*fnc* XP)$_{\varphi}$

ii. [XP]*fnc* → (XP *fnc*)$_{\varphi}$

iii. *fnc*[XP]*fnc* → (*fnc* XP *fnc*)$_{\varphi}$

It may seem in tableau (36) that the constraint ALIGN-XP,L is able to derive the expected output (candidate a), but it should be noted again that if the XP-external function words are consonant-initial, they will phrase as proclitics as shown in (37i) and ALIGN-XP,L will not be able to derive the expected structure. As demonstrated above, single-edge alignment constraints fail to explain the prosodic structure of a language like Persian in which XP-external weak function words can phrase either as proclitics or enclitics.

By examining the candidates in tableau (36) closely, we realize that the most appropriate constraint that can render candidate (36a) win over (36d) is a one which militates against a PPhrase containing two XPs. Selkirk's MATCH-XP requires that each XP be exactly matched with one PPhrase, but this hypothetical constraint should ban PPhrases containing more than one XP. In fact, a constraint with such function has been proposed in the literature.

4.4.2 MAP-XP

Büring (2001) examines the syntactic and prosodic structure of focused double objects in German. In his Optimality Theory analysis, syntactic maximal phrases tend to coincide with Accent Domains (ADs). In this approach, ADs are intermediate prosodic categories that intervene between PWords and IPhrases. He adopts this term following Uhmann (1991) and states that it is similar to 'Focus Domain' in Gussenhoven (1984), 'Intermediate Phrase' in Pierrehumbert and Hirschberg (1990), and 'PPhrase' in Truckenbrodt (1999) and others' work. One of the Accent Domain formation constraints Büring (2001) proposes is the constraint called XP which is defined in (38):

(38) XP: AD contains an XP. If XP and YP are within the same AD, one contains the other (where X and Y are lexical categories).

According to the constraint in (38), two syntactic maximal projections XP and YP cannot be contained in a single PPhrase, unless when one of them is embedded in the other in the syntactic representation. Therefore, the constraint XP will allow the prosodic phrasing in (39i), but will ban the prosodic structures in (39ii)-(39iv).

(39)

i. $[X\ [YP]\ ]_{XP} \rightarrow (X\ \ YP)_{\varphi}$

ii. $[XP\ YP]_{XP} \rightarrow *(XP\ YP)_{\varphi}$

iii. $[\textit{fnc}\ XP\ YP]_{FncP} \rightarrow *(XP\ YP)_{\varphi}$

iv. $XP\ YP \rightarrow *(XP\ YP)_{\varphi}$

Zimmermann (2006) and Lovestrand (2009) adopt the constraint XP proposed in Büring (2001) to explain the prosodic structure of focused constituents in West Chadic languages and particularly in Hausa. In the present study, we will also adopt the constraint XP, but for more clarity, we will use the name 'MAP-XP' for it. MAP-XP is defined in (40):

(40) MAP-XP: Assign one violation mark for every PPhrase which contains two or more sister XPs.

MAP-XP requires each XP to be mapped onto some PPhrase, hence the name of the constraint. If a PPhrase contains more than one XP, then each of the XPs has not been mapped onto a PPhrase properly and this will violate the constraint MAP-XP.

To compare Selkirk's MATCH-XP with MAP-XP defined in (40), let us return to tableau (36). If we change the constraint MATCH-XP in tableau (36) with MAP-XP, candidate (36a) will win over all the other candidates. This is shown in the tableau in (41). Candidate (41a) does not violate MAP-XP, because none of the PPhrases contain more than one XP. However, candidates (41c) and (41d) violate this constraint by allowing a single PPhrase contain more than one XP.

(41)

Input: [XP] *fnc* [XP] *fnc* [XP]	PARSE-INTO-φ	MAP-XP
☞ a. (XP *fnc*)$_{\varphi}$ (XP *fnc*)$_{\varphi}$ (XP)$_{\varphi}$		
b. (XP)$_{\varphi}$ *fnc* (XP)$_{\varphi}$ *fnc* (XP)$_{\varphi}$	*!W	L
c. (XP *fnc* XP *fnc* XP)$_{\varphi}$		*!W
d. (XP *fnc* XP *fnc*)$_{\varphi}$ (XP)$_{\varphi}$		*!W

By examining the structures in (23) and (24), we realize that the constraint MAP-XP is inviolable in Persian, because no PPhrase contains more than one XP in these prosodic structures. If we adopt the ranking in chapter 3, repeated here in (42), and assume that the inviolable constraint MAP-XP is undominated in this ranking, we can derive the prosodic structure of constructions containing XP-external weak function words.

(42) PARSEINTO-φ, HEAD-TO-LEX, PARSE-EDGE-R, MAP-XP, ONSET » MATCH(ω,LEX)

This is shown in the tableau in (43) for an XP-external enclitic, and in the tableau in (44) for an XP-external proclitic. The tableaux in (43) and (44) are the more complete versions of the tableaux in (46) and (47) in chapter 3. Candidate (43a) violates MATCH(ω,LEX), because it contains one maximal PWord that does not match with any lexical word. Candidate (43b) satisfies this constraint, but contains an onsetless syllable. In candidate

(43c) the PPhrase contains two XPs and also an onsetless syllable. Candidate (43d) violates both ONSET and MATCH(ω,LEX). In candidate (43e), a weak function word is mapped onto a PPhrase. This obviously violates HEAD-TO-LEX, because being a PPhrase entails being a PWord and PWords dominate at least one foot according to the headedness condition. In candidate (43f) a syllable interferes between a PWord and the PPhrase right edge, and in candidate (43g) there is an element which is not parsed into a PPhrase.

(43)

Input: /zan + -o + mard/ woman + and + man (*Lex*P + *fnc* + *Lex*P) 'woman and man'	PARSE-INTO -φ	HEAD-TO-LEX	PARSE-EDGE-R	MAP-XP	ONSET	MATCH(ω,LEX)
☞ a. $(((za)_{\omega}\ no)_{\omega})_{\varphi}\ ((mard)_{\omega})_{\varphi}$						*
b. $((zan)_{\omega})_{\varphi}\ (o\ (mard)_{\omega})_{\varphi}$					*!W	L
c. $((zan)_{\omega}\ o\ (mard)_{\omega})_{\varphi}$				*!W	*W	L
d. $((zan)_{\omega})_{\varphi}\ ((o\ (mard)_{\omega})_{\omega})_{\varphi}$					*W	*
e. $((zan)_{\omega})_{\varphi}\ (o)_{\varphi}\ ((mard)_{\omega})_{\varphi}$		*!W			*W	L
f. $((za)_{\omega}\ no)_{\varphi}\ ((mard)_{\omega})_{\varphi}$			*!W			L
g. $((zan)_{\omega})_{\varphi}\ o\ ((mard)_{\omega})_{\varphi}$	*W				*W	L

The tableau in (44) demonstrates the same ranking of constraints and its consequences for an XP-external proclitic. The winner candidate (44a) satisfies all the constraints. Candidate (44b) violates MATCH(ω,LEX), because it contains one recursive PWord. Candidate (44c) is not harmonic, because the PPhrase in it contains two XPs. Please note that this candidate satisfies all the other constraints in the tableau. The candidates (44d)-(44g) are similar to the candidates (43d)-(43g) in the previous tableau.

(44)

Input: /zan + va + mard/ woman + and + man (*Lex*P + *fnc* + *Lex*P) 'woman and man'	PARSE-INTO -φ	HEAD-TO-LEX	PARSE-EDGE-R	MAP-XP	ONSET	MATCH(ω,LEX)
☞ a. $((zan)_{\omega})_{\varphi}$ $(va\ (mard)_{\omega})_{\varphi}$						
b. $(((zan)_{\omega}\ va)_{\omega})_{\varphi}$ $((mard)_{\omega})_{\varphi}$						*!W
c. $((zan)_{\omega}\ va\ (mard)_{\omega})_{\varphi}$				*!W		
d. $((zan)_{\omega})_{\varphi}$ $((va\ (mard)_{\omega})_{\omega})_{\varphi}$						*!W
e. $((zan)_{\omega})_{\varphi}$ $(va)_{\varphi}$ $(mard)_{\omega})_{\varphi}$		*!W				
f. $((zan)_{\omega}\ va)_{\varphi}$ $((mard)_{\omega})_{\varphi}$			*!W			
g. $((zan)_{\omega})_{\varphi}$ va $((mard)_{\omega})_{\varphi}$	*!W					

4.5 Phonological Phrases: a unified OT account

The constraint MAP-XP bans PPhrases which contain more than one XP, but our analysis still lacks a constraint that requires mapping of the entire XP onto a PPhrase. We adopt the constraint WRAP-XP proposed by Truckenbrodt (1999) which demands each XP be contained in a PPhrase.

(45) WRAP-XP: Assign one violation mark for every XP that is not contained in a PPhrase.

Since WRAP-XP requires XPs only to be contained in some PPhrase and does not demand an exact matching, it seems to be an appropriate syntax-prosody interface constraint for analyzing a language like Persian that has both XP-external proclitics and enclitics. The prosodic structure of maximal projections dominating more than one maximal projection demonstrated in (39ii) and repeated here in (46) shows that the requirement for PPhrases to contain only one XP is more important than the requirement for XPs to be wrapped in a PPhrase, because in the prosodic structure in (46), the higher XP is not wrapped by any

PPhrase to avoid a PPhrase containing two XPs. This proves that the constraint MAP-XP is ranked over WRAP-XP.

(46) $[XP\ YP]_{XP} \rightarrow (XP)_{\varphi}\ (YP)_{\varphi}$ MAP-XP » WRAP-XP

Finally, following Kahnemuyipour (2003), and and the discussions in this chapter, this study assumes that PPhrases are not recursive in Persian. This fact is explained by adopting the constraint NO-REC-φ from Ito and Mester (2009).

(47) NO-REC-φ: Assign one violation mark for each PPhrase which is embedded in another PPhrase.

Now let us return to the phrase structures in Persian, and see if our constraint ranking can derive the prosodic structures suggested in Kahnemuyipour (2003) as well as the structures proposed in this study. As is asserted in the literature, DPs in Persian have a single prosodic prominence on the leftmost determiner. Since audible prominences are PPhrase heads, we can conclude that the whole DP is mapped onto a single PPhrase. It is worth mentioning again that determiners are lexical elements in Persian and can bear prosodic prominence. The tableau in (48) demonstrates the prosodic phrasing of a DP. Since only a single prosodic prominence is audible in this phrase, which is located on the first determiner *in*, all structures that predict audible prominences on other words are not acceptable. Please note that none of the candidates violate the constraint MAP-XP, because the syntactic maximal projections are embedded in each other in a recursive manner. The NP is contained inside the lower DP, and the lower DP is dominated by the higher DP. Therefore, by the definition in (40) phrasing the whole structure in a single PPhrase does not violate MAP-XP. The winner candidate (48a) satisfies all the constraints. The prosodic structure in candidate (48b) reflects the syntactic structure of the input most closely, but makes a fatal violation of the constraint NO-REC-φ. The structure in candidate (48b) cannot be accepted, because it predicts audible prominence on *do* and *mard* as well. In candidate (48c), each word is parsed into a PPhrase. Please note that this structure does not violate the previously mentioned constraint HEAD-TO-LEX, because all the three words are lexical words, but this

structure violates WRAP-XP two times, since there are two XPs, namely, the two DPs which are not wrapped by any PPhrases. Candidate (48d) is similar to candidate (48c), but it violates WRAP-XP once. In candidate (48e) none of the XPs are wrapped by PPhrases, so it violates WRAP-XP three times, but more importantly, it also violates the higher-ranked PARSE-INTO-φ.

(48)

Input: /in do mard/ this two man (*Lex* *Lex* *Lex*) [D [D [NP]]$_{DP}$]$_{DP}$ 'these two men'	PARSE-INTO-φ	NO-REC-φ	MAP-XP	WRAP-XP
☞ a. (ín do mard)$_{\varphi}$				
b. (ín (dó (márd)$_{\varphi}$)$_{\varphi}$)$_{\varphi}$		**!W		
c. (ín)$_{\varphi}$ (dó)$_{\varphi}$ (márd)$_{\varphi}$				**W!
d. (ín)$_{\varphi}$ (dó mard)$_{\varphi}$				*W!
e. (ín)$_{\varphi}$ do mard	*!W			***W

Now let us turn to various VP structures in Persian. The tableau in (49) demonstrates a prefixed verb. The prefix is the indicative mood marker *mi-* which is a lexical word and can become the PPhrase head. Candidate (49a) satisfies all the constraints, while candidates (49b) and (49c) violate WRAP-XP, since the VP is not wrapped by any PPhrase in these candidates. Candidate (49b) also violates PARSE-INTO-φ.

(49)

Input: /mi raft/ INDI goPAST3SG (*Lex* *Lex*) [Mood V]$_{VP}$ 'was going'	PARSE-INTO-φ	NO-REC-φ	MAP-XP	WRAP-XP
☞ a. (mí raft)$_{\varphi}$				
b. (mí)$_{\varphi}$ (raft)$_{\varphi}$				*!W
c. (mí)$_{\varphi}$ raft	*!W			*W

The tableau in (50) shows a VP containing a verb, and an indefinite/unspecific object. Again, candidate (50a) does not violate any of the constraints. In candidate (50b), the VP is not wrapped by any PPhrase and violates WRAP-XP. Candidate (50c) violates WRAP-XP and also PARSE-INTO-φ. The prosodic structure in candidate (50d) is very similar to the syntactic structure of the input, but it violates NO-REC-φ.

(50)

Input: /qazâ xord/ food eatPAST3SG (*Lex* *Lex*) [[NP] V]$_{VP}$ 'She/he ate food.'	PARSE-INTO-φ	NO-REC-φ	MAP-XP	WRAP-XP
☞ a. (qazá̂ xord)$_{\varphi}$				
b. (qazá̂)$_{\varphi}$ (xórd)$_{\varphi}$				*!W
c. (qazá̂)$_{\varphi}$ xord	*!W			*W
d. ((qazá̂)$_{\varphi}$ xord)$_{\varphi}$		*!W		

As we saw in §4.2, definite/specific objects combine with verbs projected onto VPs, thus a VP consisting of a definite/specific object and a verb will contain two XPs as shown in the input of the tableau in (51). A crucial point in these structures is that the weak function word *-ro* is a part of the NP. The winner candidate (51a) violates WRAP-XP, because the higher VP is not wrapped by any PPhrase in this candidate. Candidate (51b) satisfies

WRAP-XP but makes a fatal violation of MAP-XP because the PPhrase in it contains two XPs.

(51)

Input: /qazâ-ro xord/ food-ACC eatPAST3SG (*Lex fnc* *Lex*) [[NP] [VP]]$_{VP}$ 'She/he ate food.'	PARSE-INTO-φ	NO-REC-φ	MAP-XP	WRAP-XP
☞ a. (qazá̂-ro)$_{\varphi}$ (xórd)$_{\varphi}$				*
b. (qazá̂-ro xord)$_{\varphi}$			*!W	L
c. (qazá̂)$_{\varphi}$ ro (xórd)$_{\varphi}$	*!W			*
d. ((qazá̂-ro)$_{\varphi}$ (xórd)$_{\varphi}$)$_{\varphi}$		*!W	*W	L

The tableau in (52) shows the prosodic phrasing of an Ezafe construction. Candidate (52a) does not violate any of the constraints. In candidates (52b) and (52c), there is one PPhrase that contains more than one XP, thus they fail to satisfy MAP-XP. In candidate (52d) the Ezafe morphemes are not parsed into PPhrases, and candidate (52e) contains recursive PPhrases and also violates MAP-XP.

(52)

Input: /doxtar -e šâh -e pari-ân/ daughter-EZ king-EZ fairy-PL (*Lex fnc* *Lex fnc* *Lex*) [NP] EZ [NP] EZ [NP] 'fairy king's daughter'	PARSE-INTO-φ	NO-REC-φ	MAP-XP	WRAP-XP
☞ a. (doxtáre)$_{\varphi}$ (šá̂he)$_{\varphi}$ (pariyá̂n)$_{\varphi}$				
b. (doxtáre šá̂he pariyá̂n)$_{\varphi}$			*!W	
c. (doxtáre šá̂he)$_{\varphi}$ (pariyá̂n)$_{\varphi}$			*!W	
d. (doxtár)$_{\varphi}$ e (šá̂h)$_{\varphi}$ e (pariyá̂n)$_{\varphi}$	*!W			
e. ((doxtáre)$_{\varphi}$ (šá̂he)$_{\varphi}$ (pariyá̂n)$_{\varphi}$)$_{\varphi}$		*!W	*W	

As discussed in §4.3.4, the syntactic account in Kahnemuyipour (2000, 2003) maintains that in Ezafe constructions only possessors can be expanded into more complex phrases such as DPs. However, in the syntactic account proposed in Mahootian (1993) which has been adopted in this study, the functional phrase EP can contain XPs such as DPs or APs in it. If an Ezafe construction contains an XP, the XP maps onto an independent PPhrase, and the other lexical words together with their following Ezafe morphemes map on PPhrases. This is readily predictable from the constraint ranking proposed in this section, and is shown in the examples in (53). In (53i), the possessor is a DP which maps onto a PPhrase, and in (53ii), there is an AP inside the Ezafe Phrase which coincides with a PPhrase in phonological representation.

(53)

i. [xune-ye [un mard]$_{DP}$]$_{EP}$ → (*xuné-ye*)$_{\varphi}$ (*ún mard*)$_{\varphi}$
house-EZ that man
'that man's house'

ii. [xune-ye [xeyli bozorg]$_{AP}$]$_{EP}$ → (*xuné-ye*)$_{\varphi}$ (*xéyli bozorg*)$_{\varphi}$
house-EZ very big
'a very big house'

4.6 Prosodic heads and accent associations

As introduced in chapter 2, the analysis in Kahnemuyipour (2003) assumes that stress assignment takes place in three different levels in Persian. PWord-level stress is assigned to the rightmost syllable of a PWord, PPhrase-level stress is assigned to the leftmost PWord inside a PPhrase, and finally, IPhrase-level stress is assigned to the rightmost PPhrase inside an IPhrase. After the PPhrase-level stress is assigned, the PWords lose their PWord-level stress. As a consequence, there are only two levels of audible prominence in Persian, that of PPhrase level, and that of IPhrase-level.

This study, following the literature on Autosegmental-Metrical Phonology and Prosodic Phonology, makes a distinction between the abstract metrical strength and the physical audible prominence. As Fagyal et al. (2006) put forward, metrical structure provides the docking sites for pitch-accents and boundary tones. Metrically prominent constituents are prosodic heads, while the actual audible prominence is due to the pitch accents that are associated with these prosodic heads. In other words, the ranking of morphosyntax-prosody interface constraints and prosodic markedness constraints determine the prosodic structure and prosodic heads, and pitch-accents are associated with these heads. The PWord in Persian as discussed in detail in chapter 3, contains only one foot, which is right aligned with the minimal PWord. Thus, the head of a PWord is its foot, and since feet are binary and iambic in Persian, the head of the foot is its strong syllable. Following Kahnemuyipour (2003) we suppose that the head of a PPhrase is left-aligned with it, and the head of an IPhrase is right-aligned with it. These can be explained by assuming that the two constraints in (54) and (55) are undominated.

(54) ALIGN(φ, L; Hd, L): Assign one violation mark for every PPhrase whose head PWord is not left-aligned with it.

(55) ALIGN(ι, R; Hd, R): Assign one violation mark for every IPhrase whose head PPhrase is not right-aligned with it.

Having the entire prosodic structure determined now, we need a single constraint to associate pitch-accents to the prosodic heads. Selkirk (1995) proposes the constraint ASSOC-PA according to which a pitch accent must be associated to (or aligned with) a stressed syllable. We propose a similar constraint that requires all PPhrases be associated with an accent. In this study we adopt the term 'accent' instead of 'pitch-accent', because as we will see in chapter 5, we consider the possibility that factors other than pitch (f0) might also have a role in prominence of accented syllables.

(56) ASSOC-AC-φ: Assign one violation mark for every PPhrase which is not associated with an accent.

Following Hayes and Lahiri (1991), we assume that pitch accents are only associated with metrically strong (stressed) constituents. In Prosodic Phonology, metrically strong constituents are prosodic heads. Therefore, if a PPhrase is associated with an accent, the accent will be associated with its head i.e., leftmost PWord. Inside the head PWord, the accent will be associated with the foot, and inside the iambic foot, it will prefer the strong syllable on the right. This is shown in figure 8 which demonstrates the prosodic structure of an utterance of a simple sentence shown in (57). The morpho-syntax/prosody interface constraints and prosodic markedness constraints introduced in the previous and current chapters determine the prosodic structure shown in figure 8. The syllables to which pitch accents are associated are shown with an acute accent (´) on the vowel. The IPhrase in figure 8 consists of four PPhrases, and the constraint ASSOC-AC-φ will require each of them to be associated with an accent. Then inside the PWord, the accent will be associated to the strong syllable of the foot.

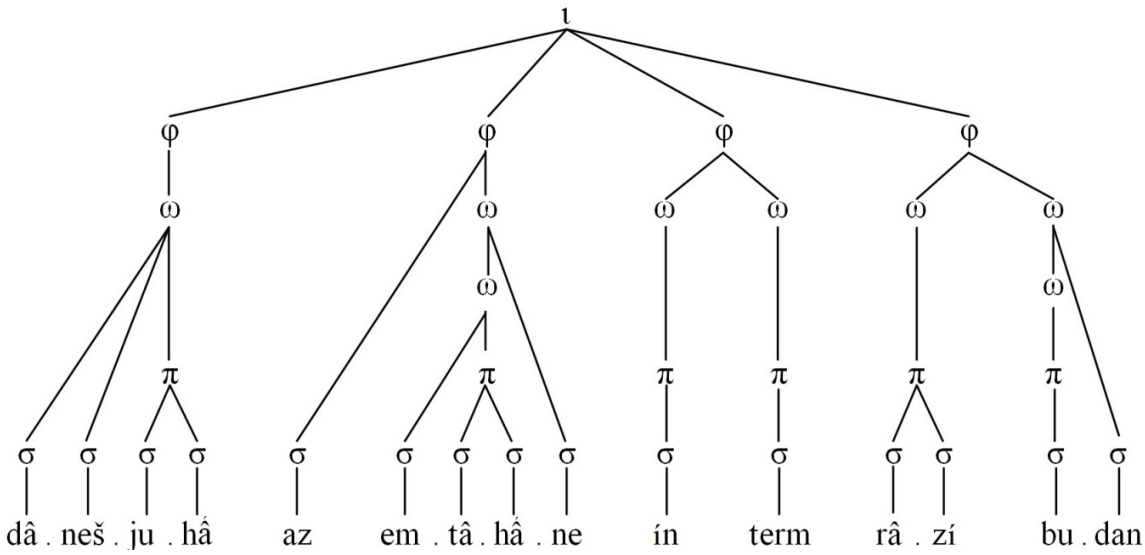


Figure 8 The prosodic structure of the sentence in (57)

(57) *dânešju-hấ az emtâhấn -e ín term râzí bud-an*
student-PL from exam EZ this term satisfied be-PAST3PL
'The students were satisfied with this term's exam.'

(58) [*dânešju-hâ*]$_{NP}$ [*az* [*emtâhân*]$_{NP}$ -e [*in term*]$_{DP}$]$_{PP}$ [[*râzi*]$_{NP}$ [*bud-an*]$_{V}$]$_{VP}$

Kahnemuyipour (2003) maintains that the rightmost PPhrase inside an IPhrase will be assigned an IPhrase-level stress, which will render this PPhrase to be perceived more prominently than the other PPhrases. Therefore, in the structure in figure 8, the rightmost PPhrase (precisely, the final syllable of the leftmost PWord inside the rightmost PPhrase) will be prosodically the most prominent part of the entire IPhrase.

The question here is why the final pitch accent (the one on the final syllable of *râzí*) is perceived more prominently than the other pitch accents. According to Ladd (2008), the last accent in an Utterance is always perceived more prominently than the others. Autosegmental-Metrical Phonology (Pierrehumbert, 1980) treats the final [pitch] accent of an IPhrase as the 'nuclear' pitch accent, as opposed to other 'pre-nuclear' ones. If we adopt the notions of nuclear and pre-nuclear accents into our analysis, we can maintain that all PPhrases are associated with a pre-nuclear accent, while the IPhrase head (the rightmost PPhrase) is associated with a nuclear accent.

There seems to be no comprehensive explanation in the literature about the fact that nuclear accents are perceived more prominently than the others. Bolinger (1986: 58) states the intuition that the final accent is the most prominent one is nothing but an illusion. Ladd (2008: 267) also maintains that there is no clear phonetic justification for treating the last accent in a sentence as the most prominent one. Sadat-Tehrani (2007) argues that the reason why the final (nuclear) pitch accent is perceived more prominently in Persian is that there is no substantial pitch rise after this accent, and also it seems that the syllable associated with the final accent has slightly higher f0 than the other accented syllables.

The next chapter aims at providing an answer for the questions raised in this section. Are the accents associated to the rightmost PPhrases perceived more prominently only as a

consequence of being located at the end of an IPhrase, or there are phonetic explanations for their special prominence?

4.7 Ezafe constructions and information structure

In §4.3.5 we introduced lexicalization as a process in which an Ezafe construction becomes a single lexical element with only one prosodic prominence, and also pointed out that beside lexicalization, sometimes lexical words present in an Ezafe construction may lose their prosodic prominence as a result of carrying old or given information. Sadat-Tehrani (2007: 46-49) based on works such as Brown (1983), Steedman (1991), Baumann and Grice (2006) and Cruttenden (2006) argues that prosodic structure in Persian is highly dependent on the information structure. He gives the example in (59) and points out the possible variations in producing the Ezafe construction in it.

(59) *mardom-e injâ xeyli mehrabun-an*
people-EZ here very kind-bePL3
'The people here are very kind.'

In Sadat-Tehrani (2007), each lexical word in an Ezafe construction together with its following Ezafe morpheme forms an Accentual Phrase. The string *mardom-e injâ* in (59) can be produced with a prosodic prominence on each of the words *mardom* and *injâ*. This can be the case when (59) is used as part of the new information, e.g., in answer to the question "How do you like your new town?" The second alternative is when the whole Ezafe construction (*mardom-e injâ*) has a single prominence on its final syllable. This can happen when the Ezafe construction carries given information, e.g., as part of the answer to the question "How do you find the people there?"

Sadat Tehrani treats the Ezafe constructions with one prominence as a single AP, and the ones with two prominences as two APs. The relation between prosodic structure and information structure is beyond the scope of the present study, but in the approach adopted

here, the non-prominent state of the words carrying given information may be due to lack of accent association rather than different prosodic structures. Bader (2001) in his OT analysis, explains the lack of prominence on given material using the OT constraint *GIVEN which bars given constituents from being prosodically prominent. If we define the constraint *GIVEN as in (60) and then assume that it is ranked over ASSOC-AC-φ, then PPhrases whose heads carry given information will not be associated with any accent, and will not be perceived prominently.

(60) *GIVEN: Assign one violation mark for each informationally given constituent which is associated with an accent.

A final issue which is worth addressing here is that in utterances containing focused constituents, pre-focal and post-focal material are considered informationally given, and tend not to be prominent. The relatively plain pitch curve on these parts which denotes lack of prominence can be attributed to the information structure of the sentence.

The discussion provided here for the prosodic structure of Ezafe construction is only valid for informationally neutral or all-new constituents. Determining the prosody of constructions containing focused elements or discourse-given material need further studies of phonology-semantics and phonology-pragmatics interface.

To sum up, in Ezafe constructions in Persian, except for informationally given cases and lexicalized constructions, in all other utterances there will be audible prosodic prominence on the final syllable of each lexical word present in the Ezafe construction as discussed in this chapter.

4.8 Summary

In this chapter, we examined the prosodic structure of syntactic phrases in Persian. First we introduced DPs and VPs and their prosodic structures based on the existing literature. Then,

we focused on Ezafe construction and introduced the suggestion in Kahnemuyipour (2003) on the prosodic structure of Ezafe constructions and pointed out the problems with this analysis. Based on the syntactic account proposed in Mahootian (1993) and others for Ezafe constructions, we proposed that in an Ezafe construction, each lexical word together with its following Ezafe morpheme forms a PPhrase. In the next sections, we explained the prosodic structure of Persian syntactic phrases including DPs, VPs and Ezafe constructions by the identical ranking of syntax-phonology interface constraints and prosodic markedness constraints. We pointed out that XP-external clitics introduced in chapter 3 cause problems for any theory that suggests exact alignment or matching between XPs and PPhrases. Then we introduced and adopted the constraint MAP-XP by which we can explain the prosodic structure of all kinds of syntactic phrases in Persian. We also suggested that after the prosodic structure is determined, an accent associates with the head of each PPhrase. We also pointed out two possible cases in which there are no audible prominences on some of the lexical words present in an Ezafe construction, namely lexicalized Ezafe constructions, and Ezafe constructions containing words carrying discourse-given information.

The prosodic structure proposed here for Persian utterances differs significantly from that in Kahnemuyipour's analysis in two main aspects. The first one comes from the fundamental prosodic distinctions between lexical and functional words. In Kahnemuyipour (2003) there are no distinctions between lexical and function words, and every morpho-syntactic word forms an autonomous PWord. The present study demonstrates the theoretical and empirical problems with the analysis and proposes that prosodically weak elements cannot be independent constituents and necessarily cliticize to their adjacent elements. As a result, this study allows recursive PWords and also syllables left unparsed into PWords. The second difference is that Kahnemuyipour (2003) treats Ezafe constructions as single lexical items and consequently as autonomous PWords. The present study proposes that each lexical item in an Ezafe construction together with its following Ezafe morpheme forms a PPhrase. The next chapter will deal with acoustic and perceptual differences between the accents that associate with PPhrases, and the accents that associate with IPhrases.

Chapter 5 Phonetic differences between pre-nuclear and nuclear accents

5.1 Introduction

In chapter 4 we sketched our account of prosodic phrasing in Persian, according to which syntax-phonology interface constraints and prosodic markedness constraints form the prosodic structure, and a constraint requiring PPhrases to be accented associates each PPhrase with an accent. As mentioned in previous chapters, the existing literature suggests that in an IPhrase, the accent of the rightmost PPhrase is perceived more prominently than the other ones. This fact has been expressed in various ways in different approaches.

The earlier approaches and descriptive works such as Vahidian-Kamyar (2000) and Khazayi-Far (2004) often make a distinction between the word stress and the so-called 'sentence stress'. In these approaches, one syllable in every word bears the word stress, and one word in each sentence bears the sentence stress which makes this word the most prominent word in the sentence.

In Autosegmental-Metrical approaches such as Mahjani (2003) and Sadat-Tehrani (2007), it is assumed that there are two levels of prosodic prominence in Persian utterances, that of the pre-nuclear pitch accent and that of the nuclear pitch accent. For instance in Sadat-Tehrani (2007), each lexical word together with its enclitics forms an Accentual Phrase, and each Accentual Phrase bears a pitch accent on the final syllable of its lexical word. The rightmost Accentual Phrase is associated with so-called 'Nuclear Pitch Accent' which is perceived more prominently than the other pitch accents.

In Kahnemuyipour (2003) which is carried out within the framework of Prosodic Phonology, PPhrase-level stress is assigned to its head, and IPhrase-level stress is assigned

to the rightmost PPhrase inside the IPhrase. According to this analysis, the stress on the rightmost PPhrase is perceived more prominently than the stresses on other PPhrases.

Following examples demonstrate and compare different perspectives towards the duality of prominence levels in Persian utterances. In an unmarked utterance of the sentence in (1), the words *diruz*, *havâ* and *sard* have audible prominences, but the word *sard* is perceived more prominently than the other words.

(1) *diruz* *havâ* *sard* *bud*
yesterday weather cold be$_{\text{PAST3SG}}$
'It was cold yesterday.'

According to a descriptive approach (Khazayi-Far, 2004), in all predicates except for predicates consisting of one simple verb, the preverbal element of the predicate takes the sentence stress. Thus, in the sentence in (1), the preverbal element of the predicate (*sard*) will be perceived more prominently than the other words, as shown in (2).

(2) *diruz havâ* ***sard*** *bud.*

Based on the Autosegmental-Metrical approach in Sadat-Tehrani (2007), copular verb declarative sentences have their Nuclear Pitch Accent (NPA) on the final element of the complement, and the copula is always deaccented. In the example in (1) the complement *sard* takes the NPA, and the post-NPA portion of the utterance becomes deaccented. Since everything including pitch accents and AP boundaries disappear after an NPA, the copula verb *bud* becomes deaccented and forms a single AP together with the NPA-bearing *sard* as shown in (3). The pre-NPA APs have a high boundary tone, while the NPA bearing AP has a low boundary tone on its right edge. The two APs preceding the NPA have their own audible pitch accents, but the rightmost Nuclear Pitch Accent is perceived more prominently than the other two.

(3) L+H* h L+H* h H* l

diruz *havâ* *sard bud*

$(\quad)_{AP}$ $(\quad)_{AP}$ $(\quad)_{AP}$

Finally, in Kahnemuyipour's approach, each word in (1) forms a separate PWord, the adjunct *diruz*, the subject *havâ*, and the verb phrase *sard bud* map onto separate PPhrases, and in the final PPhrase (*sard bud*), the leftmost PWord (*sard*) takes the PPhrase-level stress. The PPhrase *sard bud* being the rightmost PPhrase, takes the IPhrase-level stress which is the most prominent stress of the entire utterance. The three words *diruz*, *havâ* and *sard* are PPhrase heads and receive a PPhrase-level audible stress, while the word *sard* is also assigned with IPhrase-level stress which is perceived more prominently than the others. After the assignment of IPhrase-level stresses, the leveling rule deletes the PWord-level stress on *bud*. The prosodic structure of the sentence in (1) based on Kahnemuyipour's analysis is shown in (4).

(4) X_ι

$(X_\varphi$ X_φ X_φ $)_\iota$ IPhrase

$(X_\omega)_\varphi$ $(X_\omega)_\varphi$ $(X_\omega)_\varphi$ PPhrase

$(\quad)_\omega$ $(\quad)_\omega$ $(\quad)_\omega$ $(\quad)_\omega$ PWord

diruz *havâ* *sard* *bud*

The account presented in this dissertation treats the sentence in (1) in a similar way to Kahnemuyipour's approach. The difference is that in the present study, syntax-phonology interface constraints and prosodic markedness constraints form the prosodic structure of the utterance, and the constraint ASSOC-AC-φ requires all the PPhrases to be associated with an accent.

All the above approaches have one thing in common: they all reflect the fact that there are two levels of audible prominence in Persian utterances, namely that of phrase-level (PPhrase or AP) prominence, and that of sentence-level (IPhrase) prominence, and that the latter is the final prominence in the utterance, and is perceived more prominently than the

other prominent constituents. This study treats the former as pre-nuclear accent and the latter as nuclear accent and this chapter aims at determining the phonetic differences between these two accents. The main question in this chapter is whether nuclear accent is perceived more prominently only because it is the final accent of the utterance or because it is phonetically different from the pre-nuclear accent. As mentioned in chapter 4, Bolinger (1986: 58) argues the intuition that the final accent is more prominent than the other accents is nothing but an illusion, and Ladd (2008: 267) maintains that there is no clear phonetic justification for treating the last accent in a sentence as the most prominent one.

In previous researches that aimed at determining the acoustic correlates of stress and accent in Persian, no distinction has been made between these two types of accent. For instance, Sadeghi (2011, 2012) and Abolhasanizadeh et al. (2012) use their target words in accented (almost always nuclear accented) positions, unaccented (post-focal) positions and focused position, and try to measure and compare the acoustic correlates of prosodic prominence in these three positions. No experiments known to the author has yet been conducted in order to compare the phonetic characteristics of pre-nuclear accented words with those of nuclear accented words. As discussed in the previous chapter, nuclear and pre-nuclear accents are associated with PWords that are PPhrase heads. In this chapter, when referring to the head PWords associated with an accent, the terms 'nuclear accented words' or 'pre-nuclear accented words' have been used for the sake of convenience.

5.1.1 Delayed peaks

As mentioned in chapter 4, Silverman and Pierrehumbert (1990) have shown that there is a tendency in the languages that have both nuclear and pre-nuclear accents, for the pre-nuclear accents (manifested as non-final f0 peaks) to occur after the syllable with which they are intuitively associated. Basically, this type of f0 peak delay does not appear in case of nuclear accents (final f0 peaks). In Persian, the f0 peak on a nuclear accented word occurs almost in the middle of metrically strong syllables of the words. However, in pre-

nuclear accented words, if the word is followed by enclitics, the f0 peak occurs on the enclitic(s) as a result of peak delay, while the accent is intuitively associated with the final syllable of the host as shown in figure 5 in chapter 4. If the pre-nuclear accented word is not cliticized, the f0 peak tends to occur at the right edge of the final syllable of the accented word, and sometimes on the left edge of the following word due to the peak delay phenomenon. This is shown in figure 1 which illustrates the pitch contour of an utterance of the sentence in (5).

(5) *dišab* *kozovo* *nesbatan* *sard* *bud-e*
last night Kosovo rather cold be$_{\text{PAST-PFV}}$
'Kosovo was rather cold last night.'

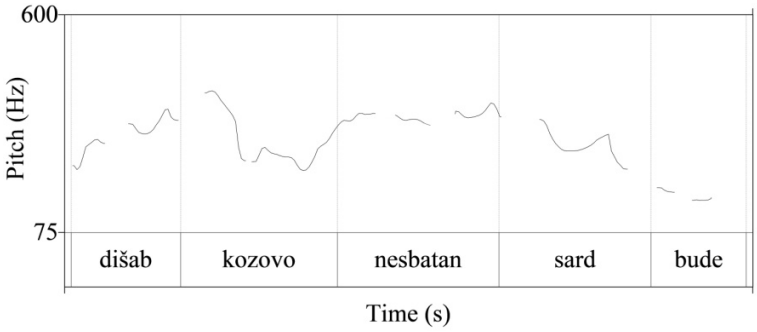


Figure 1 Pitch contour of an utterance of the sentence in (5) with delayed f0 peaks on pre-nuclear accented words

In the pitch contour in figure 1, the left edge of the second word *kozovo* has obviously a very high f0 due to peak delay. The same phenomenon can also be observed on the left edge of the third word *nesbatan*.

Sadat-Tehrani (2007) in his Autosegmental-Metrical approach treats this phenomenon as an effect of AP boundary tones. In his analysis, the right edge of the final AP which bears a nuclear accent is associated with a low boundary tone, while the right edge of other non-final APs are associated with high boundary tones. His approach successfully explains the occurrence of f0 peaks on the right edges of pre-nuclear accented words. It also explains the

occurrence of f0 peaks on the enclitics in cliticized words, because the high boundary tone spreads over the enclitic(s). However, this analysis fails to explain the occurrence of the f0 peak on the left edge of the following word, a phenomenon frequently found in Persian. Attributing the difference between the locations of the f0 peak in syllables associated with the two types of accents to delayed peaks (an explanation adopted in the current study), can explain all the above problems, and moreover, has been shown to be a common issue in other languages as well.

The delay of f0 peaks on pre-nuclear accented words and the lack of this delay on nuclear accented words, leads to a difference in the shape of the pitch contour on the words associated with these two types of accents. On pre-nuclear accented words, the pitch contour often has a rising slope at the right edge of the word, while on nuclear accented words, the pitch contour has a falling slope at the right edge of the word, because the peak occurs in the middle of the metrically strong syllable. This is shown in the idealized schema in figure 2. This hypothetical utterance consists of four words in which the third one is associated with the nuclear accent.

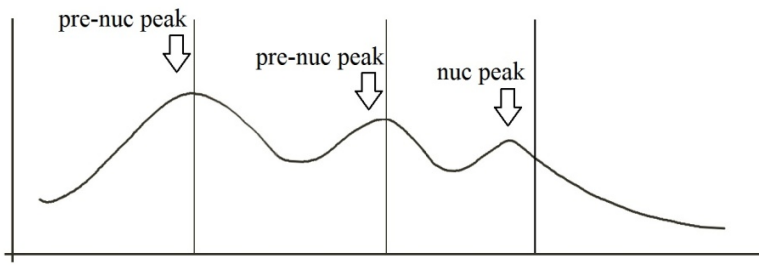


Figure 2 An idealized schema of the pitch contour of a hypothetical Persian utterance

This study assumes that one difference between pre-nuclear and nuclear accents is the delay of f0 peaks on pre-nuclear accented words and the lack of this delay on nuclear accented words. This difference leads to a contrast between the shapes of the pitch contour on the two types of accented words. Two questions arise out of this assumption. The first question is, whether this difference between the shapes of the pitch contours can actually cue the

difference between the two types of accents. The next question is whether the two types of accents have any other phonetic differences which render nuclear accented words to be perceived more prominently than pre-nuclear accented words.

This chapter presents two experiments carried out in order to provide an answer to these questions. The first experiment is a production experiment that investigates the differences between the acoustic correlates of the pre-nuclear accent and the acoustic correlates of the nuclear accent by using non-words produced in pre-nuclear accent bearing positions and in nuclear accent bearing positions. The second experiment is a perception experiment designed to determine which acoustic correlate(s) can render the nuclear accent to be perceived more prominently than pre-nuclear accents. This experiment uses manipulated speech as stimuli.

5.2 Acoustic correlates of pre-nuclear and nuclear accents

This experiment aims at measuring and comparing the acoustic parameters in nuclear accented syllables and pre-nuclear accented ones. First the methods of the experiment will be presented, and results and discussions will follow the methods section.

5.2.1 Methods

This section will introduce the material used in this experiment (target words and carrier sentences), participants, procedures (recordings and measurements) and statistical analyses of the experiment.

Experiment materials

In this experiment two non-words were used in carrier sentences in pre-nuclear accent bearing and nuclear accent bearing positions. One of the two non-words consisted of four identical open syllables (*mâmâmâmâ*), and the other consisted of three identical closed syllables (*dandandan*). Each participant first read the carrier sentences with the words *âlâbâmâ* 'Alabama' and *bargaštan* 'coming back' as target words, and then was asked to read the carrier sentences again and replace *âlâbâmâ* with *mâmâmâmâ*, and *bargaštan* with *dandandan*. The non-words were composed of reiterated speech, segmentally and prosodically similar to their real word counterparts, and the participants were instructed to treat non-words as naturally as possible, and read them as if they are uttering the real words in daily conversation.

Non-words made of reiterated speech have been used in previous studies of acoustic correlates of stress and accent, among which one can name Morton and Jassem (1965), Katwijk (1974), Liberman and Steeter (1978), Berinstein (1979), Sluijter (1995), Sluijter and van Heuven (1996ab), Sluijter et al. (1997) and Qian (2008). The above-mentioned non-words were adopted in this study firstly because the syllable boundaries and segment boundaries in them are accurately identifiable. Secondly, since all syllables are identical in each non-word, effects caused by intrinsic characteristics of vowels and consonants, and also micro-prosodic effects will also be identical in all syllables. Thus, all acoustic correlates of prosodic prominence including f0, intensity and duration can be readily compared among the syllables of the same word, without need to control these effects. Fricatives and voiceless consonants were avoided in target words to prevent their effects on the f0. As demonstrated in chapter 3, the left edges of PWords in Persian are strong positions, and in fact as we will see in the results of this experiment, initial syllables of PWords have longer duration and higher intensity than their adjacent syllables. F0 is also higher on leftmost syllables in most cases, due to the delayed f0 peaks of their preceding words. Therefore, if disyllabic target words are used in this experiment and similar experiments, the increase of duration, intensity and f0 on the first syllable, due to strong

start effects, may interfere with the increase of these parameters on the second syllable as a result of accent association. Previous studies on Persian such as Sadeghi (2011, 1012) and Abolhasanizadeh et al. (2010, 2012) adopted disyllabic target words, and compared the acoustic correlates of prosodic prominence on the second syllable with those on the first syllable. In the present study, however, in order to prevent this problem, three-syllabic and four-syllabic target words have been adopted.

Each of the target words *dandandan* and *mâmâmâmâ* was embedded in two carrier sentences, in one in a pre-nuclear accent bearing position, and the other in the nuclear accent bearing position in order to measure and compare the acoustic correlates of pre-nuclear accent with those of the nuclear one. Target words were posited in the middle of the carrier sentences far from IPhrase boundaries, to avoid boundary effects on target words.

It is a well-known fact in phonetic studies that fundamental frequency decreases gradually from the beginning to the end of an utterance. The fundamental frequency also tends to decrease substantially after each peak, so that in an IPhrase, each f0 peak is lower than its preceding peak. The former phenomenon is known as 'declination', and the later as 'downstep', and the overall rubric used for these phenomena is 'downtrends' (Connell, 2001). Since this experiment will compare the f0 values on final syllables of target words in two different carrier sentences, if the downtrend effects differ in the two carrier sentences, the comparison will not be valid. In order to control the downtrend effects, the target words in all carrier sentences were placed at exactly the same distance from the beginning of the carrier sentence. The syllable-based distance between a target word and the left edge of the carrier sentence and the number of accents (f0 peaks) preceding the target word were identical in all carrier sentences. In all carrier sentences, in the portion preceding the target word, there were two words, one cliticized three-syllabic word, followed by a monosyllabic word. Thus the number of syllables, and the number and location of audible prominences preceding a target word was identical in all types of carrier sentences. This is shown in (6) below, where syllable boundaries are shown by (.), word boundaries by (#) and the accented syllables by (′).

(6) ○.ó.○# ó# target word post-target portion

The two carrier sentences in which the target word was placed in a pre-nuclear accent bearing position are given in (7). In (7i) the VP *bâruni-ye* and in (7ii) the VP *nâmomken-e*, being the rightmost PPhrase, receive the nuclear accent, and this guaranties that the underlined target words, receive a pre-nuclear accent. Please note that if we take the descriptive approach or the Autosegmental-Metrical approach in Sadat-Tehrani (2007), the underlined target words in (7i and 7ii) will still be predicted to have an audible accent less prominent than their following words.

(7) Carrier sentences embedding the target word in pre-nuclear positions:

i. *hafte-ye* *baʔd* *<u>âlâbâmâ</u>* *bâruni-ye*
week-EZ next Alabama rainy-be3SG
'Next week it's going to be rainy in Alabama.'

ii. *hafte-ye* *baʔd* *<u>bargaštan</u>* *nâmomken-e*
week-EZ next returning impossible-be3SG
'Next week it's impossible to return.'

The two carrier sentences in which the target word is placed in a nuclear accent bearing position are given in (8). In both (8i and 8ii) the target word together with the copula *bud* form the rightmost PPhrase of the IPhrase, and the underlined target words which are the heads of these PPhrases, are associated with the nuclear accent which are perceived more prominently than the preceding accented words. Again, if we adopt the descriptive or Autosegmental-Metrical approach, the underlined words are predicted to have a more prominent accent than their preceding words.

(8) Carrier sentences embedding the target words in nuclear positions:

ii. *xune-ye* *ted* *<u>âlâbâmâ</u>* *bud*

house-EZ PR Alabama bePAST3SG

'Ted's house was in Alabama.'

i. *hadaf-e* *man* *<u>bargaštan</u>* *bud*

purpose-EZ I returning bePAST3SG

'My purpose was to return.'

In order to create a context for the carrier sentences to be read without any contrastive focus, introductory sentences were used before carrier sentences. This context was designed in a way that participants read the carrier sentences as informationally all-new sentences. During recordings, if focused utterances were produced, the participants were asked to read the introductory and carrier sentences again. After recordings, utterances containing contrastive focused elements and other erroneous utterances were eliminated from the analyses.

Before recording the actual utterances, training data were presented to the subjects, and they exercised treating non-words sufficiently. During recordings, when the subjects felt their utterance was unnatural in any sense, they repeated the sentences as many times they wanted.

Participants and recordings

Participants were ten native Persian speakers, six females and four males, with an age range from 23 to 32 and an average age of 28.1 at the time of recording. All were monolinguals born and raised in Iran, but had studied English and Japanese in their adulthood. They were all unaware of the purpose of the study.

Data were recorded in a soundproof recording booth in the Sound Lab. at Komaba Campus, The University of Tokyo. The recordings were done using the software Praat at a sampling rate of 44100 Hz and acceptable noise rate, and the recorded data were directly saved as .wav files to the hard disk of a laptop computer. The microphone used in this experiment was a SONY ECM-MS975, and the mouth-microphone distance was around 5 cm, and the subjects were asked to maintain this distance during recording.

Measurements

The main measurement intervals in this experiment were the sonorant parts of the rhyme in all syllables of each target word. Recorded utterances were annotated to text grid files using Praat. There were five annotation tiers in each text grid file, namely, point tier in which maximum f0 of the final syllable of the target word was marked, segment tier, interval tier, target word tier and sentence tier. Segmentations and annotations were done manually following the guidelines in Ladefoged (1996, 2006) and Sluijter (1995). The main criterion used for segmentation was the spectrum of the recorded speech. Sound wave, vowel formants and auditory criteria were also used when it was necessary. Figure 3 demonstrates an example of segmentation and annotation to text grids.

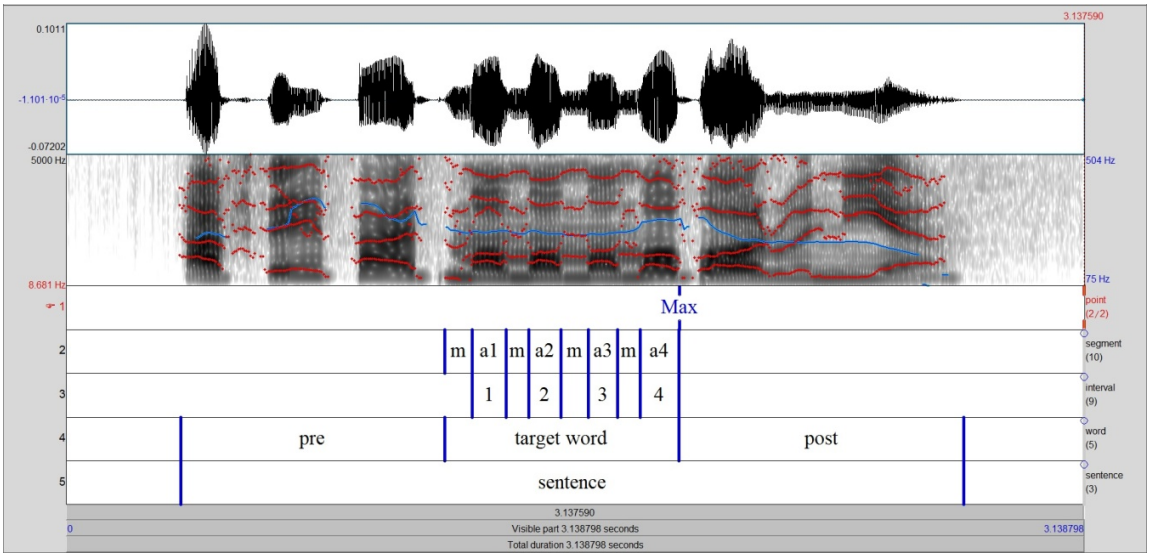


Figure 3 An example of segmentation and annotation of an utterance containing the target word *mâmâmâmâ* in a pre-nuclear accent bearing position (The measurement intervals are shown in the interval tier by the numbers 1-4.)

Pitch

In order to measure pitch, a Praat script was used to measure and extract the f0 maxima in Hertz in all measurement intervals. As mentioned in chapter 3 and chapter 4, in pre-nuclear accents, sometimes the f0 peak appears after the syllable with which it is intuitively associated. All the utterances were examined for delayed peaks, and where the f0 peak occurred after the final interval, it was marked in the point tier, and the value of the delayed peak was used instead of the f0 maximum value in the final interval. The difference between the maximum f0 value of the final interval with that of the penultimate interval was calculated for all utterances. This difference was compared and statistically analyzed between a nuclear accented target word and a pre-nuclear accented word. The statistical analysis was conducted separately for male and female speakers.

Duration

As for duration, the length of all intervals and all target words were measured and extracted using a Praat script. The duration analyzed in this experiment was the relative duration of each interval to the duration of the entire target word. This is shown in the formula in (9).

(9) $\text{Relative interval duration} = \frac{\text{interval duration(ms)} * 100}{\text{target word duration(ms)}}$

The difference between the duration of final and penultimate intervals was also calculated and compared between a nuclear accented target word and a pre-nuclear accented word.

Overall intensity

A Praat script was used to measure and extract the mean RMS intensity in all intervals in dB. The difference between the mean RMS intensity value of the final interval with that of the penultimate interval was calculated for all target words. Then, this difference was compared between nuclear accented target words and pre-nuclear accented words.

Vowel quality

A Praat script was used to measure and extract the first, the second and the third formants of all the vowels in each target word at the temporal midpoint of the vowels, where all formants of the vowel tend to be more stable and more reliable. The statistical analyses of formant values were conducted separately for male and female speakers.

Spectral tilt

Traditionally, it was believed that stress-accent languages like English use intensity to cue stress, however in 1950s, laboratory phonetics studies such as Fry (1955) showed that overall intensity does not provide a reliable cue to stress in these languages. Four decades later, Sluijter and van Heuven (1996b) showed that stressed syllables are produced with more effort, but more glottal effort does not necessarily lead to greater amplitude of the waveform. In fact, increased vocal effort generates a more strongly asymmetrical glottal pulse: low frequency components are hardly affected and the energy increase is concentrated on the higher harmonics only. In increased vocal effort, the closing phase is

shortened, and as a result, there is a shift of intensity over the spectrum such that the intensity increase is concentrated on the higher harmonics only. This phenomenon is called 'spectral balance' or 'spectral tilt' and is considered to be a reliable acoustic correlate of stress in languages like English. On the other hand, a more gradual pattern of glottal closure leads to low-frequency emphasis which cues unstressed syllables. Waveform produced with more effort has a more abrupt opening and a briefer closing portion. The ratio of open time to total period duration is called 'open quotient' (OQ). Sluijter (1995) argues that the difference between the first two harmonics (H1-H2) can be used as a reliable measure for OQ. The difference between H1 and A1, A2 and A3 which are the amplitudes of F1, F2 and F3 respectively, can also be used in analyzing the spectral tilt. Please note that f0 does not affect OQ. A summary of the physiological and acoustic aspects of spectral tilt, and their phonological role in perception of stress is given in table 1.

Table 1 An overview of psychological origins of spectral tilts and their acoustic equivalents and phonological impacts

Physiology	Acoustic	Relation to stress
Open Quotient (OQ)	H1-H2	Lower in stressed syllables
Degree of glottal opening	H1-A1	Lower in stressed syllables
Duration of closing portion	H1-A2 and H2-A3	Lower in stressed syllables

Studies such as Sluijter (1995) and Sluijter and van Heuven (1996b) have shown that in languages like English, the above three parameters -unlike overall intensity- change significantly when the syllable is stressed. In the present study, spectral measurements were also carried out to examine if spectral balance can cue the difference between pre-nuclear and nuclear accents in Persian. Spectral measurements were done using a Praat script. Two different versions of the same script were used for male and female speakers. Reference frequencies for F1, F2 F3, F4 and F5 were 500 Hz, 1485 Hz, 2475 Hz, 3465 Hz and 4455 Hz respectively in male version, and 550 Hz, 1650 Hz, 2750 Hz, 3850 Hz and 4950 Hz respectively in the female version. The maximum frequency was 5000 Hz in the male script

and 5500 Hz in the female one. After measurements were done, negative values were offset by adding 30dB to all the resulting measures.

Statistical analyses

There were four types of reading materials in this experiment, namely, open syllable target word (*mâmâmâmâ*) in pre-nuclear accent bearing position, open syllable target word in nuclear accent bearing position, closed syllable target word (*dandandan*) in pre-nuclear accent bearing position, and closed syllable target word in nuclear accent bearing position. The 10 participants read each of the sentences three times, and all together 120 utterances were recorded.

The goal of the statistical analyses in this experiment was to determine whether the values of each parameter (pitch, duration, intensity, vowel quality and spectral tilt) in the rightmost syllables of pre-nuclear accented words and those of nuclear accented words are significantly different from each other, to determine if these parameters can cue the distinction between the two types of accent in Persian. Each parameter was analyzed separately: a univariate Analysis of Variance (ANOVA) and a least significant difference (LSD) post hoc test were performed to determine if the parameter value differs significantly in pre-nuclear accented target words and nuclear accented target words. In each test, the related parameter (for example the f0 maxima values in the analysis of pitch) was set as the dependent variable, and INTERVAL (first, second, third, fourth) and ACCENT TYPE (pre-nuclear vs. nuclear) were set as fixed factors. In each ANOVA test, an average value was first calculated for the three repetitions of each speaker. Two different statistical tests were conducted for each parameter. The aim of the first test was to investigate if there are significant differences between parameter values of final and penultimate intervals or not. For instance, in case of f0, the aim was to determine if there is an INTERVAL effect or not. The second tests were conducted on cases in which there were significant differences between parameter values of the final and the penultimate intervals. The goal of the second

tests was to determine if the differences of values between final and penultimate intervals in nuclear accented target words differ significantly with those in pre-nuclear accented words or not. For example, in case of f0, the aim of these tests was to investigate if there is a significant ($\text{INTERVAL}_{\text{final-penult}}$)*ACCENTTYPE interaction or not. In case of pitch and vowel formants, the analysis was carried out separately for each sex. Data were analyzed using SPSS version 20 for Windows.

5.2.2 Results

In this section the results of statistical analyses will be reported separately for each parameter. For each parameter, first the statistical analysis of the closed syllable target word (*dandandan*) will be presented, then the analysis of the open syllable target word (*mâmâmâmâ*) will be reported.

Pitch

The analysis of f0 maxima values revealed that in both pre-nuclear accented and nuclear accented words, f0 gradually falls from the leftmost interval to the right, but rises abruptly on the final interval. The final interval is a part of the word with which the accent is associated. However, interestingly the pitch rise on the final interval on pre-nuclear accented words is significantly higher than the rise on the final interval of nuclear accented words. This seems to be in contradiction with the fact that nuclear accented words are perceived more prominently than pre-nuclear accented words. Since all downtrend effects have been controlled in the carrier sentences embedding the two kinds of target words, this difference cannot be attributed to downtrend effects such as declination or downstep. Following sections will provide the details of the results.

Closed syllable target word

In closed syllable target words, in both pre-nuclear accented cases and nuclear accented ones, the pitch falls on the second interval, but rises abruptly on the third interval which is associated with accent. Table 2 shows the descriptive analysis of f0 maxima in each interval of the word *dandandan* measured for the two types of accent for female and male speakers separately. The f0 values are in Hz and the standard deviations are given in parentheses.

Table 2 Descriptive analysis of f0 maxima for the closed syllable target word

Interval	Pre-nuclear		Nuclear	
	Female	Male	Female	Male
1	278.91 (20.49)	135.63 (12.39)	286.64 (8.623)	147.04 (14.11)
2	253.01 (10.44)	118.54 (10.23)	238.56 (13.57)	117.23 (12.32)
3	306.01 (9.42)	141.92 (13.03)	273.25 (23.06)	132.66 (14.69)
Total N	18*3=54	12*3=36	18*3=54	12*3=36

The graph in figure 4 is made based on the data in Table 2 and shows the mean f0 maxima values in the intervals of the target word *dandandan* uttered by female speakers. As can be seen, in both pre-nuclear and nuclear accented words, the f0 falls on the second interval, but rises substantially on the third interval. However, this rise is much more in case of pre-nuclear accented words.

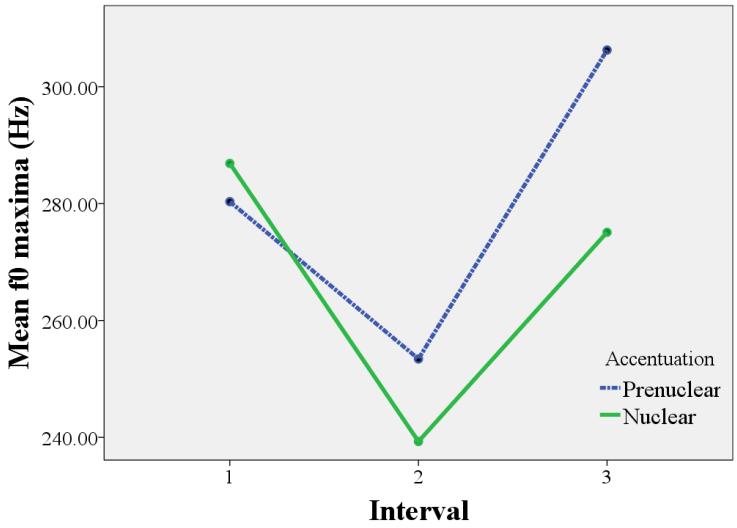


Figure 4 Mean f0 maxima values in the measurement intervals of the target word *dandandan* uttered by female speakers

Similarly, the graph in figure 5 demonstrates the mean f0 maxima values in the intervals of the target word *dandandan* uttered by male speakers. Again, as is apparent, in both types of accents, the f0 falls on the second interval, and rises on the third one, however, this rise seems to be higher in case of pre-nuclear accented words.

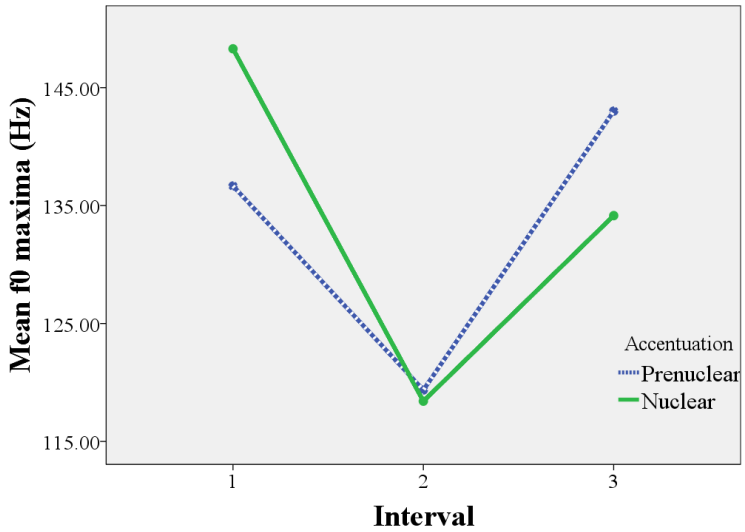


Figure 5 mean f0 maxima values in the measurement intervals of the target word *dandandan* uttered by male speakers

First an ANOVA test was conducted to investigate whether there is an interaction between the accent type (ACCENTTYPE) and the gender of the participants (SEX). The f0 of the final intervals was set as dependent variable and ACCENTTYPE and SEX as fixed factors. The results showed no significant ACCENTTYPE*SEX interaction [F(1,16)=2.024, p=0.174]. Therefore, statistical analyses were conducted separately for the results of each gender.

An ANOVA test results revealed that there was a significant INTERVAL effect in both pre-nuclear and nuclear accented words and for both sexes. A summary of ANOVA test results is given in Table 3.

Table 3 ANOVA test results for mean f0 maxima on the intervals of the target word *dandandan* in pre-nuclear and nuclear accented words for both sexes

Accent	sex	ANOVA result
Pre-nuclear	F	$F(2,10)= 24.51, p<0.001$
	M	$F(2, 6)= 3.99 , p=0.048$
Nuclear	F	$F(2,10)= 19.35, p<0.001$
	M	$F(2, 6)= 4.50, p=0.048$

An LSD post hoc test revealed that in both pre-nuclear and nuclear accented words, and for both sexes, the mean f0 maxima is significantly lower on interval 2 than that on intervals 1 and 3. A summary of LSD test results is given in table 4.

Table 4 LSD test results for mean f0 maxima on the intervals of the target word *dandandan* in pre-nuclear and nuclear accented words for both sexes

Accent	sex	Interval 1 and Interval 2	Interval 2 and Interval 3
Pre-nuclear	F	2 lower than 1 ($p=0.037$)	3 higher than 2 ($p<0.001$)
	M	2 lower than 1 ($p=0.049$)	3 higher than 2 ($p<0.001$)
Nuclear	F	2 lower than 1 ($p<0.001$)	3 higher than 2 ($p<0.001$)
	M	2 lower than 1 ($p=0.011$)	3 higher than 2 ($p=0.018$)

As mentioned above, the f0 rise on the third intervals seems to be higher in pre-nuclear accented words than in nuclear accented ones. Therefore, an ANOVA test was conducted to examine whether the f0 rise between the second and the third intervals in pre-nuclear accented words is significantly higher than that in nuclear accented words (the ($\text{INTERVAL}_{\text{final-penult}}$)*ACCENTTYPE interaction). The ANOVA test on the female data showed a significant effect of maximum f0 value rise on the last interval, [$F(1,5)= 22.07$, $p<0.001$]. The test on the male data also showed a significant effect [$F(1,3)= 13.72$, $p=0.005$]. This shows that the pitch rise between penultimate and final intervals is significantly higher on pre-nuclear accented words than that on nuclear accented words. The comparison results are shown in the graph in figure 6.

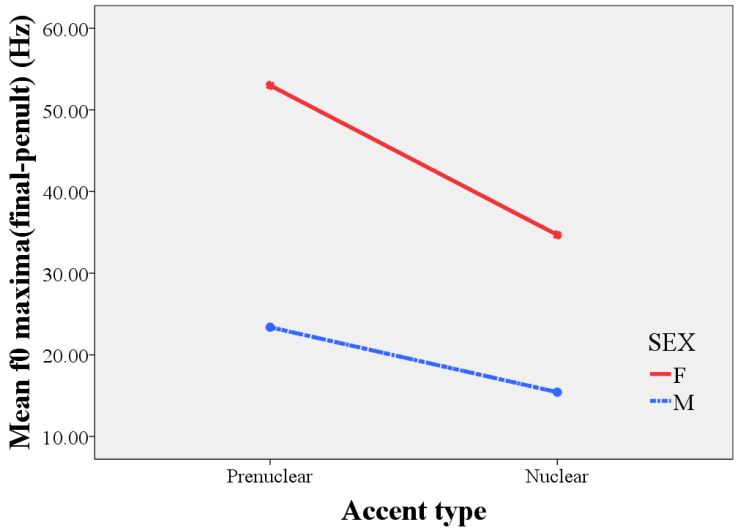


Figure 6 The difference between f0 on final interval and that on penultimate interval, measured in both pre-nuclear and nuclear accented words for both sexes

Open syllable target word

The results for the open syllable target word (*mâmâmâmâ*) were very similar to the results of the closed syllable target word mentioned above. In both pre-nuclear and nuclear accented words, pitch falls on the second and third intervals, but rises abruptly on the fourth interval which is associated with accent. Table 5 shows the descriptive analysis of f0 maxima in each interval of the word *mâmâmâmâ* measured for the two types of accents for female and male speakers separately.

Table 5 Descriptive analysis of f0 maxima for the open syllable target word *mâmâmâmâ*

Interval	Pre-nuclear		Nuclear	
	Female	Male	Female	Male
1	287.57 (16.12)	132.36 (14.50)	284.59 (14.32)	124.97 (20.16)
2	252.67 (6.33)	112.09 (6.54)	249.44 (11.68)	113.21 (17.38)
3	247.85 (7.261)	108.30 (6.83)	248.69 (10.53)	110.13 (14.57)
4	294.85 (18.34)	140.58 (12.72)	291.28 (18.02)	137.66 (9.86)
Total N	18*4=72	12*4=48	18*4=72	12*4=48

Figure 7 is made based on the data in Table 5 and shows the mean f0 maxima values in the intervals of the target word *mâmâmâmâ*, in the utterance of female speakers. In both pre-nuclear and nuclear accented words, the f0 falls on the second and the third interval, but rises substantially on the fourth interval, however, this rise is higher in case of pre-nuclear accented words.

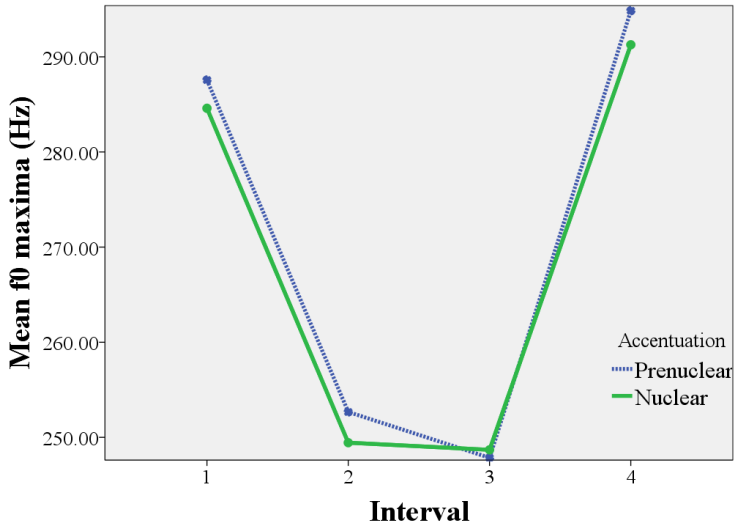


Figure 7 Mean f0 maxima values in the measurement intervals of the target word *mâmâmâmâ* uttered by female speakers

The graph in figure 8 demonstrates the mean f0 maxima values in the intervals of the target word *mâmâmâmâ* uttered by male speakers. Again, in both types of accents, the f0 falls on the second and the third intervals, and rises on the fourth one. However, again, this rise seems to be higher in case of pre-nuclear accented words.

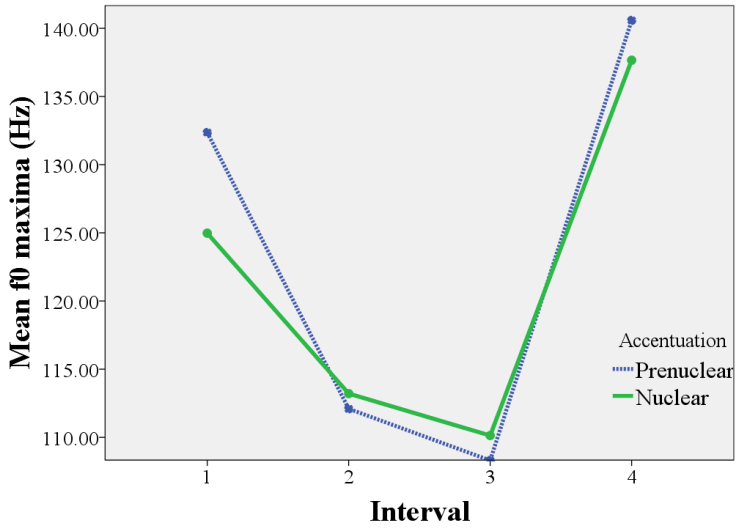


Figure 8 Mean f0 maxima values in the measurement intervals of the target word *mâmâmâmâ* uttered by male speakers

Again similar to closed syllable words, an ANOVA test was conducted to investigate whether there is an interaction between ACCENTTYPE and SEX. The results showed no significant interaction [F(1,16)=0.127, p=0.727]. Therefore, statistical analyses were conducted separately for the results of each gender.

The ANOVA test results showed that there was a significant INTERVAL effect in both pre-nuclear and nuclear accented words, and for both sexes. A summary of ANOVA test results is given in Table 6.

Table 6 ANOVA test results for mean f0 maxima on the intervals of the target word *mâmâmâmâ*, in pre-nuclear and nuclear accented words for both sexes

Accent	sex	ANOVA result
Pre-nuclear	F	F(3,15)= 22.52, p<0.001
	M	F(3,9)= 8.17, p=0.02
Nuclear	F	F(3,15)= 17.16, p<0.001
	M	F(3,9)= 3.29, p=0.043

An LSD post hoc test revealed that in both pre-nuclear and nuclear accented words, and for both sexes, the mean f0 maxima are significantly higher on interval 4 than on interval 3. The f0 on the second interval is significantly lower than the f0 on the first interval in all cases, except in the nuclear accented words uttered by male speakers. Finally, there was no significant differences between the f0 maxima values on the second and that on the third intervals. A summary of LSD test results is given in table 8. Shaded cells in the table indicate no significant difference.

Table 7 LSD test results for mean f0 maxima on the intervals of the target word *mâmâmâmâ*, in pre-nuclear and nuclear accented words for both sexes

Accent	sex	Interval 1 and Interval 2	Int. 2 and Int. 3	Interval 3 and Interval 4
Pre-nuclear	F	2 lower than 1 (p=0.031)	no sig. dif. (p=0.72)	4 higher than 3 (p<0.001)
	M	2 lower than 1 (p<0.001)	no sig. dif. (p=0.37)	4 higher than 3 (p<0.001)
Nuclear	F	2 lower than 1 (p<0.001)	no sig. dif. (p=0.96)	4 higher than 3 (p<0.001)
	M	no sig. dif. (p=0.37)	no sig. dif. (p=0.88)	4 higher than 3 (p=0.018)

An ANOVA test was done to examine whether the f0 rise between the third and the fourth interval in pre-nuclear accented words is significantly higher than that in nuclear accented words. The ANOVA test on the female data showed a significant effect of maximum f0 value rise on the last interval [$F(1,5)=31.27$, $p<0.001$]. The test on the male data also showed a significant effect [$F(1,3)=22.38$, $p=0.001$]. This shows that, again similar to the closed syllables, the pitch rise between penultimate and final intervals is significantly higher in pre-nuclear accented words than that in nuclear accented words. This is shown in the graph in figure 9.

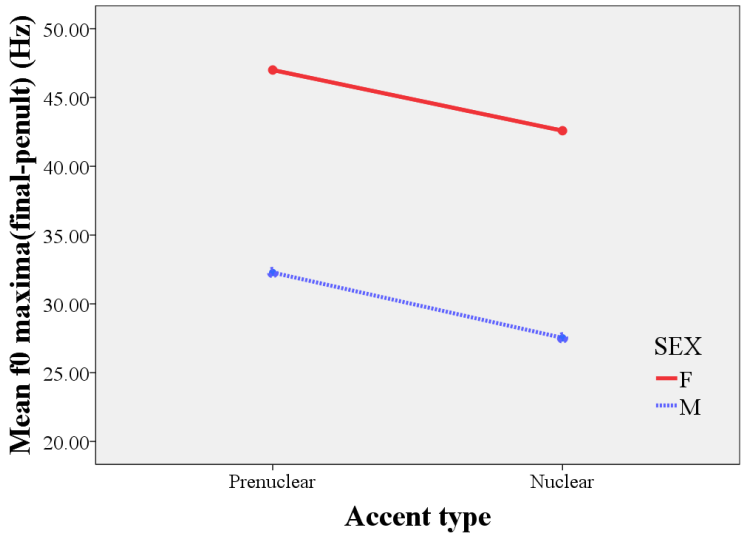


Figure 9 The difference between f0 on the final interval and that on the penultimate interval, measured in both pre-nuclear and nuclear accented words for both sexes

The analyses of f0 maxima in closed and open syllable target words show that the pitch rise on the rightmost syllable of a pre-nuclear accent bearing word is significantly higher than the pitch rise on the rightmost syllable of a nuclear accent bearing word. If we assume that higher pitch will lead to more prominence, these results seem to be in contrast with the fact that nuclear accent bearing words are perceived more prominently than pre-nuclear accent bearing ones.

Duration

The duration measured in this experiment was the relative duration of the interval to the duration of the entire target word. The results of measurements were substantially different in pre-nuclear accented words and nuclear accented words. While the duration does not increase in the final intervals of pre-nuclear words, the duration of final intervals in nuclear accent bearing words increase significantly. This suggests that duration increase is an acoustic correlate of nuclear accent, while it does not cue a pre-nuclear accent. In this

section first the results for the closed syllable target word will be reported, and then the results for the open syllable target word will be presented.

Closed syllable target word

In pre-nuclear accented words, the relative duration of the intervals decrease gradually from the first interval to the third one, however, in nuclear accented word, the relative duration of the second interval is slightly shorter than that of the first interval, but it is substantially longer in the third interval in comparison with both the first and the second intervals. The data is presented in table 8 in which values are in milliseconds.

Table 8 Descriptive analysis of relative duration in each interval of the word *dandandan* measured for the two accent positions

Interval	Pre-nuclear	Nuclear
1	34.09 (1.69)	31.33 (0.86)
2	32.12 (2.20)	30.23 (1.05)
3	28.28 (2.66)	34.43 (1.67)
Total N	30*3=90	30*3=90

The graph in figure 10 depicts the data in table 8. As can be seen, the nuclear accent bearing word differs from the pre-nuclear accent bearing one in that the final interval is longer than the other intervals in the former.

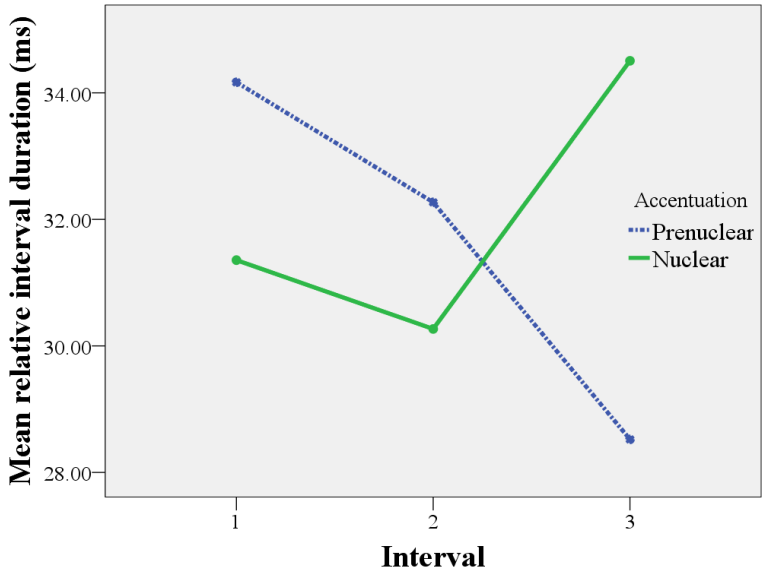


Figure 10 Mean relative interval duration values in the measurement intervals of the target word *dandandan*

The ANOVA test results revealed that there was a significant effect of relative duration in pre-nuclear accented words [F(2,18)= 23.42, p<0.001] and in nuclear accented words [F(2,18)= 33.70, p<0.001]. An LSD post hoc test also showed that in pre-nuclear accented words, the second interval is significantly shorter than the first one, and the third interval is shorter than the second one. However, in nuclear accented words, although the second interval is significantly shorter than the first one, the third interval is significantly longer than the two other intervals. A summary of the LSD test results is given in table 10.

Table 9 LSD test results for mean relative durations of the intervals of the target word *dandandan* in pre-nuclear and nuclear accented words

Accent	Interval 1 and Interval 2	Interval 2 and Interval 3
Pre-nuclear	2 lower than 1, p=049	3 lower than 2, p<0.001
Nuclear	2 lower than 1, p=042	3 higher than 2, p<0.001

Finally, an ANOVA test was done to examine whether the difference between duration of the final and the penultimate intervals in pre-nuclear accented words is significantly shorter than that in nuclear accented words. The test showed a significant effect of relative interval duration increase [$F(1,9)= 33.08$, $p<0.001$].

This shows that in the closed syllable target words, the duration of the final interval increases significantly in nuclear accented words, while in pre-nuclear accented words, the final interval is shorter than the penultimate one. This is shown in the graph in figure 11.

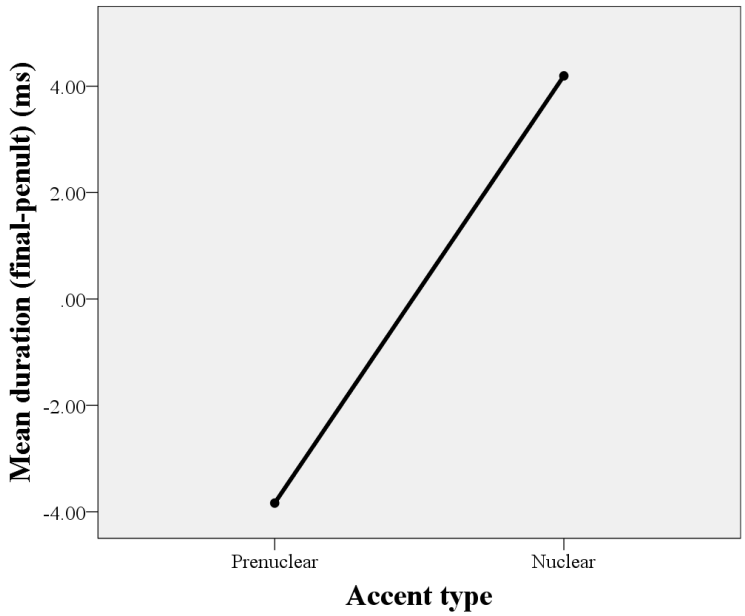


Figure 11 The difference between duration of the final interval and that of the penultimate interval, measured in both pre-nuclear and nuclear accented words

Open syllable target word

In pre-nuclear accented words, the relative duration of the intervals decreases gradually from the first interval to the fourth one, however, in nuclear accented words, the second interval is slightly shorter than the first interval, and the third interval is slightly shorter

than the second one, but the duration is substantially longer in the fourth interval in comparison with all the other intervals. The data is presented in table 10.

Table 10 Descriptive analysis of relative duration in each interval of the word *mâmâmâmâ* measured for the two accent positions

Interval	Pre-nuclear	Nuclear
1	15.35 (1.81)	14.73 (2.10)
2	14.26 (1.98)	13.30 (1.77)
3	14.20 (2.58)	13.88 (2.10)
4	14.77 (1.78)	18.25 (2.40)
Total N	30*4=120	30*4=120

The graph in figure 12 depicts the data in table 10. As can be seen, the nuclear accent bearing word differs with the pre-nuclear accent bearing one in that the final interval is substantially longer than the other intervals in the former.

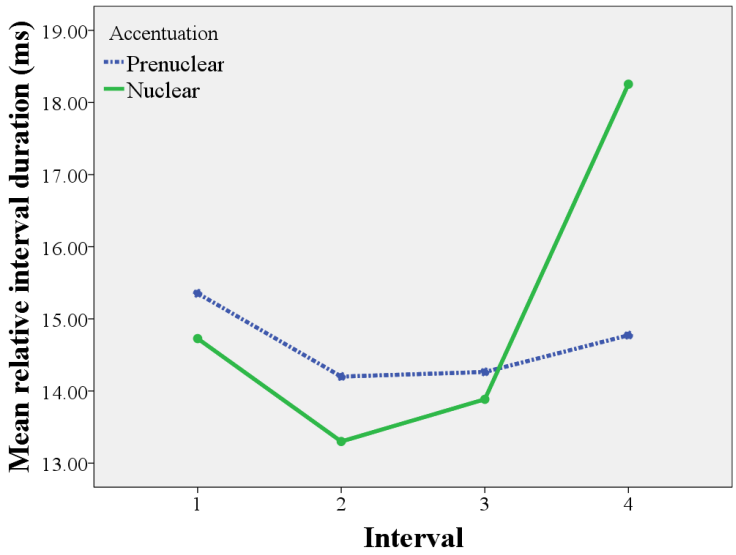


Figure 12 Mean relative interval duration values in the measurement intervals of the target word *mâmâmâmâ*

The ANOVA test results revealed that there was no significant effect of relative duration values in pre-nuclear accented words [F(3,27)=1.97, p=0.134], but there was a significant effect of relative duration values in nuclear accented words [F(3,27)= 31.56, p<0.001]. An LSD post hoc test also showed that in nuclear accented words, the second interval is significantly shorter than the first one. There are no significant differences between the duration of the second and the third intervals, but the fourth interval is significantly longer than the three other intervals. A summary of the LSD test results is given in table 11.

Table 11 LSD test results for mean relative durations of the intervals of the target word *mâmâmâmâ* in pre-nuclear and nuclear accented words

Accent	Interval 1 and Interval 2	Int. 2 and Int. 3	Interval 3 and Interval 4
Nuclear	2 is lower than 1, p=0.07	No sig. dif., p=0.75	4 is higher than 3, p<0.001

Finally, an ANOVA test was done to examine whether the difference between durations of the final and the penultimate intervals in pre-nuclear accented words is significantly shorter

than that in nuclear accented words. The test showed a significant effect of relative interval duration increase [$F(1,9)= 50.67$, $p<0.001$].

This shows that in open syllable target words, similar to closed syllable target words, the duration of the final interval increases significantly in nuclear accented words, while in pre-nuclear accented words, there is no significant difference between the durations of the final interval and the penultimate one. This is shown in the graph in figure 13.

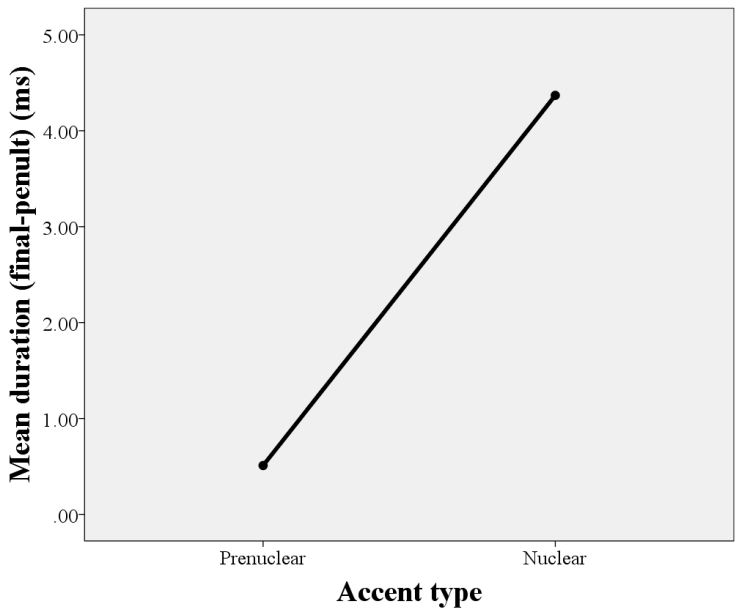


Figure 13 The difference between duration of the final interval and that of the penultimate interval, measured in both pre-nuclear and nuclear accented words

The results of relative duration analysis reveal that only in nuclear accented words the duration of final syllable increases significantly. In pre-nuclear accented words in this experiment no duration increase was observed in the final intervals.

Overall intensity

Mean intensity increased slightly on the final interval in both pre-nuclear and nuclear accented words, but the amount of increase was not statistically significant in either type. The data of closed syllable target word and open syllable target word will be reported separately in this section.

Closed syllable target word

In both pre-nuclear and nuclear accented words, the mean intensity falls on the second interval, but rises on the third one. The data is presented in table 12 and the related graph is demonstrated in figure 14. The intensity values are in dB.

Table 12 Descriptive analysis of overall intensity in each interval of the word *dandandan* measured for the two accent positions

Interval	Pre-nuclear	Nuclear
1	79.37 (6.57)	79.84 (6.23)
2	78.27 (7.11)	78.14 (6.65)
3	79.56 (6.86)	78.89 (6.75)
Total N	30*3=90	30*3=90

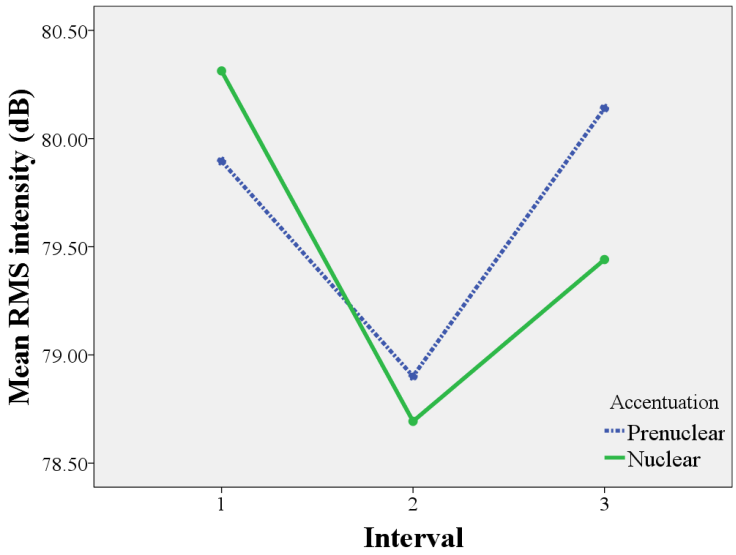


Figure 14 Mean interval intensity values in the measurement intervals of the target word *dandandan*

However, an ANOVA test revealed that the differences among mean intensities of the three intervals are significant neither in pre-nuclear accented words [F(2,18)=0.073, p=0.99], nor in nuclear accented words [F(2,18)= 0.088, p=0.95].

Open syllable target word

In Open syllable target words as well, in both pre-nuclear and nuclear accented words the mean intensity falls on the second interval and rises on the third one. The data is presented in table 13 and the related graph is given in figure 15.

Table 13 Descriptive analysis of overall intensity in each interval of the word *mâmâmâmâ* measured for the two accent positions

Interval	Pre-nuclear	Nuclear
1	81.57 (5.71)	83.97 (4.82)
2	80.19 (6.99)	83.16 (5.38)
3	79.06 (6.96)	82.50 (5.29)
4	80.68 (6.63)	84.76 (4.13)
Total N	30*4=120	30*4=120

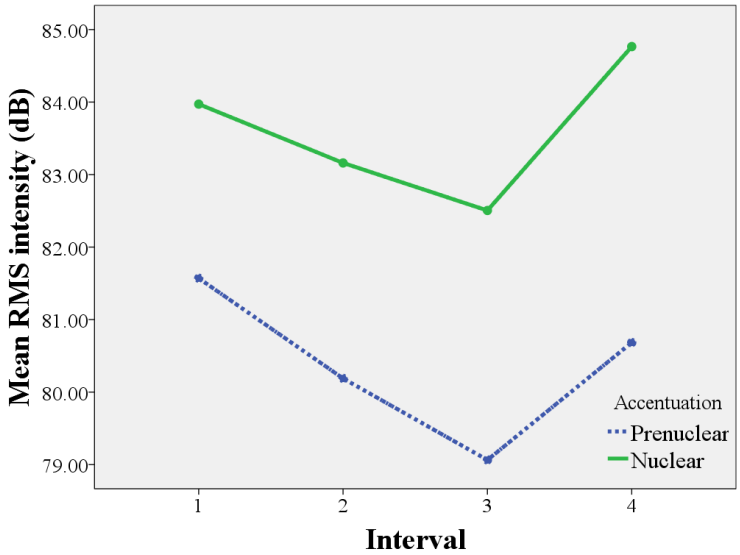


Figure 15 Mean interval intensity values in the measurement intervals of the target word *mâmâmâmâ*

Again, an ANOVA test revealed that the differences among mean intensities of the three intervals are significant neither in pre-nuclear accented words [F(3,27)=0.282, p=0.843], nor in nuclear accented words [F(3,27)= 0.232, p=0.878].

Spectral tilt

There was significant differences between spectral measures of any of the intervals neither in closed syllable target word, nor in open syllable one. The final syllable had lower values than the penultimate interval in most cases, but the difference was of no statistical significance.

Closed syllable target word

The ANOVA test results for closed syllable target words are given in table 14. As can be seen, there was no significant differences between the spectral measures of intervals, therefore, pre-nuclear accented words and nuclear accented words were not compared together.

Table 14 ANOVA test results for spectral measures on the intervals of the target word *dandandan* in pre-nuclear and nuclear accented words

Parameter	Accent type	ANOVA result
H1-H2	Pre-nuclear	F(2,18)= 2.21, p=0.10
	Nuclear	F(2,18)= 0.38, p=0.70
H1-A1	Pre-nuclear	F(2,18)= 0.64, p=0.54
	Nuclear	F(2,18)= 0.18, p=0.85
H1-A2	Pre-nuclear	F(2,18)= 0.97, p=0.40
	Nuclear	F(2,18)= 1.37, p=0.26
H1-A3	Pre-nuclear	F(2,18)= 0.53, p=0.61
	Nuclear	F(2,18)= 0.12, p=0.91

Open syllable target word

In open syllable target words, again there was no significant differences between the spectral measures of intervals. The ANOVA test results are given in table 15.

Table 15 ANOVA test results for spectral measures on the intervals of the target word *mâmâmâmâ* in pre-nuclear and nuclear accented words

Accent	sex	ANOVA result
H1-H2	Pre-nuclear	F(3, 27)= 0.13, p=0.95
	Nuclear	F(3, 27)= 0.43, p=0.98
H1-A1	Pre-nuclear	F(3, 27)= 0.65, p=0.57
	Nuclear	F(3, 27)= 0.64, p=0.60
H1-A2	Pre-nuclear	F(3, 27)= 3.07, p=0.39
	Nuclear	F(3, 36)= 2.44, p=0.78
H1-A3	Pre-nuclear	F(3, 27)= 0.28, p=0.84
	Nuclear	F(3, 27)= 1.32, p=0.28

Vowel quality

No statistically significant difference was observed among the vowel formant values of the intervals either in closed syllable target words, or in open syllable target words. There was a tendency for F1 to be slightly higher on the final vowel of both pre-nuclear and nuclear accented words, and for F2 to be slightly lower on the final vowel of both pre-nuclear and nuclear accented words. However, none of these differences were statistically significant. The ANOVA test results of closed syllable target words are given in table 16, and the results of open syllable target words are given in table 17. Since there were no significant differences among the intervals, Pre-nuclear and nuclear accented words have not been compared.

Closed syllable target word

Table 16 ANOVA test results for vowel formants in the middle of intervals of the target word *dandandan* in pre-nuclear and nuclear accented words for both sexes

Formant	Accent	sex	ANOVA result
F1	Pre-nuclear	F	F(2,5)=2.49, p=0.09
		M	F(2, 3)=0.24, p=0.90
	Nuclear	F	F(2,5)=2.14, p=0.12
		M	F(2, 3)=3.09, p=0.07
F2	Pre-nuclear	F	F(2,5)=0.65, p=0.58
		M	F(2, 3)=0.08, p=0.97
	Nuclear	F	F(2,5)=2.04, p=0.13
		M	F(2, 3)=1.11, p=0.08
F3	Pre-nuclear	F	F(2,5)=0.91, p=0.42
		M	F(2, 3)=0.28, p=0.99
	Nuclear	F	F(2,5)=1.80, p=0.17
		M	F(2, 3)=1.45, p=0.25

Open syllable target word

Table 17 ANOVA test results for vowel formants in the middle of intervals of the target word *mâmâmâmâ* in pre-nuclear and nuclear accented words for both sexes

Formant	Accent	sex	ANOVA result
F1	Pre-nuclear	F	F(3,15)=0.41, p=0.76
		M	F(3,9)=1.01, p=0.41
	Nuclear	F	F(3, 15)=0.85, p=0.48
		M	F(3,9)=0.04, p=0.99
F2	Pre-nuclear	F	F(3, 15)=0.34, p=0.81
		M	F(3,9)=1.13, p=0.36

	Nuclear	F	F(3, 15)=0.22, p=0.89
		M	F(3,9)=0.01, p=0.99
F3	Pre-nuclear	F	F(3, 15)=2.49, p=0.08
		M	F(3,9)=0.46, p=0.99
	Nuclear	F	F(3, 15)=0.47, p=0.71
		M	F(3,9)=0.10, p=0.99

5.2.3 Discussion

The research question of this experiment was whether nuclear accent differs from pre-nuclear accent phonetically, or it is perceived more prominently only because it is the final accent of the IPhrase. The answer this experiment provides is that the two types of accents are in fact phonetically different in Persian. A syllable which is associated with a pre-nuclear accent tends to have a higher pitch, while a syllable associated with a nuclear accent tends to have a longer duration. Both types of accents use f0 rise to cue prominence, while only nuclear accent uses duration as an acoustic correlate of accent. Since nuclear accent is always the final accent of an IPhrase, and occurs after all pre-nuclear accents, one may argue that nuclear accented syllables have a lower pitch due to downtrends effects. However, as explained in the methods section, in this experiment, the distance between a target word and the beginning of the carrier sentence embedding it was exactly the same in terms of number of syllables and number of accents (f0 peaks) in all cases. In other words, the target words located in a pre-nuclear bearing position had the same distance from the initial point of the utterance as the target words placed in a nuclear accent bearing position. Therefore, it is not possible to attribute the lower pitch of nuclear accented syllables to declination and downstep.

The previous literature on Persian compared the duration of accented syllables with the duration of deaccented (post-focal) syllables, and compared the duration of final syllables in focused words with that of non-focused ones, and there are no researches comparing the

duration of pre-nuclear accented syllables and that of nuclear accented syllables. As we saw in the results, the nuclear accented syllable is significantly longer than the pre-nuclear accented syllable. One possible objection here may be that the nuclear accented words have been produced with some kind of narrow (distinctive) focus on them, and the substantial duration increase in the final syllables of these words is due to narrow focus rather than the nuclear accent. This objection cannot be accepted for two reasons. First, all carrier sentences were planned to be uttered as informationally all-new utterances, without narrow focus on any of the words, and the utterances which were judged to be produced with narrow focus were eliminated from the analyses. Second, as mentioned in Sadat-Tehrani (2007), Taheri-Ardali and Xu (2012), Abolhasanizadeh et al. (2012) and Hosseini (2012a, 2013a), the integral acoustic correlate of narrow (contrastive) focus in Persian is the abrupt pitch rise on the final syllable of the focused word. However, in the nuclear accented words in this experiment, the pitch rise on the final syllables is less than the pitch rise on the final syllables in pre-nuclear accented words. Please note that the interval duration in this experiment is the relative duration of the final interval to the duration of the whole target word. Thus, the duration increase in the final intervals of nuclear accented words cannot be attributed to the increase of the length of the whole target word which is common in narrow focused words according to Taheri-Ardali and Xu (2012).

Sluijter (1995) shows that in Dutch and English, metrically strong syllables have longer durations than the other syllables, even when they are not associated with an accent. Thus, she suggests that duration is only an acoustic correlate of stress, not that of accent. However, the results of this experiment showed that in Persian, unlike Dutch and English, duration is an acoustic correlate of a specific type of accent.

The syllable duration is known to increase at the right edge of prosodic constituents such as PWords, PPhrases and IPhrases. This is known as pre-boundary lengthening (final lengthening) effect. As discussed in Prieto et al. (1995), in general, major phrase boundaries such as IPhrase boundaries tend to be associated with larger pre-boundary lengthening effects than minor phrase boundaries such as PWords. The same phenomenon

has been shown and discussed in other studies such as Liberman and Pierrehumbert (1984) and Silverman and Pierrehumbert (1990).

As we saw, nuclear accents are associated with IPhrase heads, while pre-nuclear accents are associated with PPhrase heads, and a nuclear accent is the final accent in the IPhrase. One may argue that the difference between the lengths of syllables associated with pre-nuclear accents with those associated with nuclear accents can be attributed to pre-boundary lengthening effects. However, this cannot be acceptable, since although nuclear accents are the rightmost accents in an IPhrase, they are not necessarily associated with the rightmost syllable of the IPhrase. Final lengthening is an edge sensitive phenomenon, while nuclear accents do not always occur at the edges of IPhrases. For example, the sentence in (10i), ends with a prefixed verb, which forms a VP that maps onto a PPhrase as shown in (10ii). This PPhrase contains four PWords (*bâl*)$_{\omega}$ (*dar*)$_{\omega}$ (*mi*)$_{\omega}$ (*âvord-am*)$_{\omega}$, among which the first PWord (*bâl*) is the PPhrase head. The nuclear accent associates with this final PPhrase, since it is the IPhrase head. Therefore, *bâl* bears the nuclear accent of this IPhrase.

(10)

i. dâšt-am az xoshâli bâl dar-mi-âvord-am
havePAST-1SG from happiness wing out-INDI-bringPAST-1SG
'I was growing wings (walking on air) with happiness.'

ii. ((*dâšt-am*)$_{\varphi}$ (*az xoshâli*)$_{\varphi}$ (*bâl dar-mi-âvord-am*)$_{\varphi}$)$_{\iota}$

Examples like (10) show clearly that the duration increase in nuclear accented syllables in comparison with pre-nuclear accented ones, cannot be attributed to edge sensitive phenomena such as pre-boundary lengthening effect. In the utterances used in experiments of this study, nuclear accent bearing syllables never occurred at the right edge of a major prosodic phrase where edge-sensitive phenomena such as final lengthening may target.

The overall intensity was higher in the final intervals comparing to that in penultimate intervals in both pre-nuclear accented words and nuclear accented words, but the difference

was not statistically significant. This is in accordance with the findings in Abolhasanizadeh et al. (2012) in which accented syllables have higher intensity than unaccented (post-focal) words but the difference between the two is not significant.

Since the intensity rise on the final intervals of pre-nuclear accented words and nuclear accented words is not significant, it is pointless to compare the intensity rise in the two types of accents.

In spectral measures, no significant difference was observed among the final interval and the other intervals, although in both nuclear and pre-nuclear accented words, there was a tendency for the spectral measures to be lower in the final interval. It seems that unlike English and Dutch, Persian does not use spectral tilt as a strong cue to prosodic prominence.

There were no statistically significant differences in the F1, F2 and F3 values of final and penultimate intervals, neither in pre-nuclear accented words, nor in nuclear accented words. In both types of accented words, there was a tendency for F1 to rise on the final interval and for F2 to fall on the final interval, but these tendencies were statistically insignificant.

Finally, it is worth noting that most production experiments revealed a significant difference of acoustic parameter values between the first and the second intervals as well. First syllables have a higher f0, a longer duration and also a relatively higher intensity. This may need further investigations, but it is in accordance with the acoustic experiment in chapter 3 that presumes the fact that word-initial syllables are strong in Persian.

5.3 Perceptual correlates of pre-nuclear and nuclear accents

This section reports two perception experiments conducted in the present study. First, the goal and general background of the experiment will be outlined, and then the two tests will be explained in detail.

It was already known from Silverman and Pierrehumbert (1990) that nuclear accents are timed with the final (stressed) syllable of the words they are associated with, while pre-nuclear accents tend to appear slightly later than the final syllables of the words. The results of the production experiments in this study also suggested that in pre-nuclear accented words, the pitch rises on the final syllable more than it rises on the final syllables of nuclear accented words, and also in nuclear accented words, the relative duration of the final syllable to the duration of the entire word is much more than the same value in pre-nuclear accented words.

The goal of the present experiment is to determine whether manipulations of the f0 values and the duration of the final syllable in one of these two types of accented words, can indeed lead the native listeners to perceive it as the other type or not. The f0 was manipulated in two ways: first, the maximum point of f0 was raised in nuclear accented syllables, second the f0 on the right edge of the nuclear accented syllables was raised. In order to conduct the experiments, two target words were embedded in two carrier sentences in pre-nuclear and nuclear accent bearing positions. The sentences were read by a native speaker, and the data were recorded, manipulated and were used as stimuli in perceptual experiments. The amount of manipulations was determined as the maximum differences between the pre-nuclear and nuclear values in the speech of the same speaker. For example, if in the production experiments, the maximum difference between pitch values of pre-nuclear accented syllables and nuclear accented syllables was 25 Hz in one step, in perception experiment, the pitch value of the nuclear accented syllable was raised 25 Hz. The perception experiment contained two related tests. In the first test, the values were manipulated abruptly at the maximum point. For instance the f0 was raised 25 Hz. In the second test, the parameters whose abrupt manipulation had a significant effect on perception, were manipulated gradually in several steps, to examine how much manipulation of the parameter is needed to render the stimuli to be perceived as the opposite accent type. The gradual manipulations of parameters were done separately also and in combination with each other. For example, f0 manipulations were done both alone, and in combination with duration manipulations.

5.3.1 Methods

Perception tests were conducted using carrier sentences containing target words, and masked portions. The carrier sentences with their verbs being masked were minimal pairs differing only in one pre-nuclear accent as opposed to a nuclear accent. Recorded utterances were manipulated and used as stimuli in a Praat experiment.

Recording Materials

Since in informationally neutral sentences in Persian, always the rightmost PPhrase is associated with nuclear accent and the word bearing this type of accent is always fixed and determined by the syntactic structure, it is not possible to find minimal pair sentences that differ in only one pre-nuclear accent as opposed to a nuclear one. Also since in Persian sentences, there is always one and only one nuclear accent bearing word which follows all pre-nuclear accent bearing words (if any), if we find two segmentally identical sentences in which the same word can be read with either a pre-nuclear or nuclear accent, the number of accents (f0 peaks) will be different in these two versions, i.e., if the word is produced with a pre-nuclear accent, the utterance will have one more f0 peak comparing when the word is produced with a nuclear accent. Therefore, the number of peaks can cue the difference between the two sentences and such pair of sentences cannot be regarded as a true minimal pair.

In order to overcome this challenge, carrier sentences containing a target word and a masked portion were used in this experiment, and the participants were asked to choose the correct word for the masked portion between two options. The target words were placed immediately before the verb of the carrier sentences. The verb was either a copula or a transitive verb such as ‘laughed’ or ‘danced’. The target word preceding a copula takes the nuclear accent, while since the transitive verb takes the nuclear accent, the target word preceding a transitive verb will be associated with a pre-nuclear accent. The verbs were

entirely masked by a beep sound and the participants were asked to fill in the masked portion with a copula or a given transitive verb. The first research question of this experiment was whether native listeners can discriminate the two types of accents (making a correct guess about the masked verb) without actually hearing the verb. In the next step, the f0 and the duration of the final syllables of the nuclear accented target words were manipulated to see if these manipulations can lead the participants to perceive the word as a pre-nuclear accent bearing word. The standards for manipulations, as explained above, were determined based on the previously done production experiments and on the utterances of the same speaker.

There were two pairs of carrier sentences in this experiment which are given in (11) and (12). Target words are underlined and the masked portion is shaded. In (11i), *mahtâb*[1] being the rightmost PPhrase of the IPhrase, is associated with the nuclear accent and the copula verb *bud* bears no accent. In (11ii), the transitive verb *xandid* has the nuclear accent, and the noun *mahtâb* bears a pre-nuclear accent. Thus, if the verbs in (11i) and (11ii) are masked by a beep sound, the remaining will form a minimal pair in which the utterances differ only in a nuclear accent on *mahtâb* in (11i) as opposed to a pre-nuclear accent on *mahtâb* in (11i).

(11) i. *emšáb-am* *mahtâb* *bud*
tonight-too moonlight bePAST3SG
'Tonight as well was a moonlight.'

ii. *emšáb-am* *mahtâb* *xand-íd*
probably PR laugh-PAST3SG
'Mahtab laughed tonight as well.'

[1] The word *mahtâb* 'moonlight' can also be a female proper name in Persian.

The word *bahman*[2] in (12i) is the final PPhrase and takes the nuclear accent, while *bahman* in (12ii) takes a pre-nuclear accent, and the intransitive verb *raqsid* takes the nuclear accent. Like example (11), the aim of this test was to examine if listeners are able to discriminate pre-nuclear and nuclear accented versions of the word *bahman*, when the final verb is masked by a beep sound.

(12) i. *arusí-šun* *bahmán* *bud*
wedding-POSS3PL Bahman be-PAST3SG
'Their wedding was in [month] Bahman.'

ii. *arusí-šun* *bahmán* *raqs-íd*
wedding-POSS3PL PR dance-PAST3SG
'Bahman danced in their wedding.'

Production experiments revealed that the pitch rise on the final syllable of nuclear accented words is less than that on the final syllable of pre-nuclear accented words. At the same time, the duration increase in the final syllable of nuclear accented words is more than the duration rise in the final syllable of pre-nuclear accented words. Therefore, the participants listened to both sentence pairs in (11) and (12) with their verbs being masked, and in each case they had to choose the appropriate verb for the masked portion between two given choices. Also, the nuclear accented target words in (11i) and (12i) were manipulated in different ways to examine whether the listeners will perceive them as the pre-nuclear accented word in (12ii) and (12ii). The main purpose of manipulations was to investigate what kind of, and how much manipulation of which parameter(s) will lead the nuclear accented word to be perceived as a pre-nuclear accented word.

[2] The word *bahman* is a male proper name and also is the name of the 11th month in Iranian calendar (January 21－February 19)

Recordings

The four sentences in (11) and (12) were read by a 30 year old male monolingual speaker, born and raised in Tehran. The recorded utterances were checked carefully to be produced with informationally neutral accentuation and with no narrow focus on any element. The recordings were done in a soundproof booth in the Sound Lab. at Komaba Campus, The University of Tokyo. The stimuli sounds were recorded by a SONY ECM-MS975 microphone at a *sampling* rate of 44100 Hz using the software Praat, and were directly saved to the hard disk of a laptop computer. The mouth-to-microphone distance was kept at 5 cm during the recordings.

Manipulations

All manipulations were done automatically using the software Praat, after identifying and annotating syllable and word boundaries. This section will explain the manipulation methods briefly.

Pitch

In order to manipulate the pitch, first the sonorant part of the syllables which were the subject of manipulations were identified and annotated as the manipulation intervals. There were two different types of pitch manipulations in this experiment. In the first type, pitch was manipulated at the maximum f0 point of the intervals. The maximum f0 on the final syllable of the nuclear accented word was raised to see if the listeners will perceive the manipulated stimulus as a pre-nuclear accented word. In the second type, the f0 on the right edge of the final syllable in the nuclear accented word was manipulated and raised. As discussed at the beginning of this chapter, the delay of f0 peaks in pre-nuclear accents, leads to a difference between the shapes of f0 curves on nuclear accented words and pre-nuclear accented words. The f0 peak is at the middle of the syllable in nuclear accented

syllables, and at the right edge in pre-nuclear accented syllables. As a result, in pre-nuclear accented words, the f0 on the right edge of the final syllable is rising, while in nuclear accented words, the f0 on the right edge of the final syllable is falling. Therefore, the f0 on the right edge of the final syllable of the nuclear accented word was manipulated and raised to examine if the listeners will perceive this word as a pre-nuclear accented word.

In order to manipulate the f0 in both manipulation types, first the utterance files were opened as a 'manipulation object' in Praat and the manipulation intervals were selected as shown in figure (16).

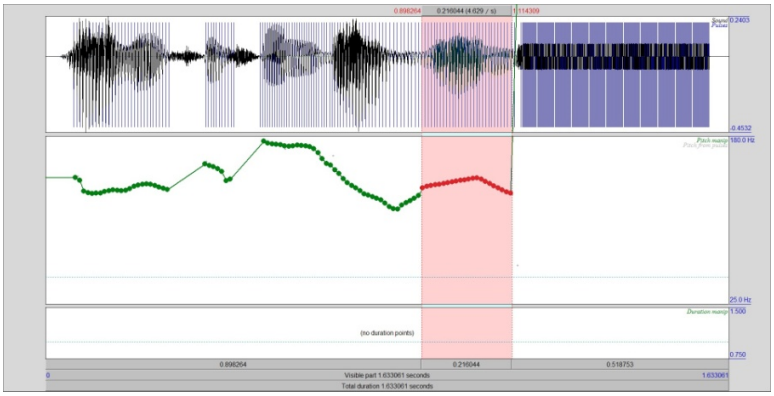


Figure 16 Selecting the manipulation interval in a Praat manipulation object

Then, all the pitch points (thick dots) in the interval, except for the leftmost, the rightmost, and the maximum one were removed as shown in figure (17).

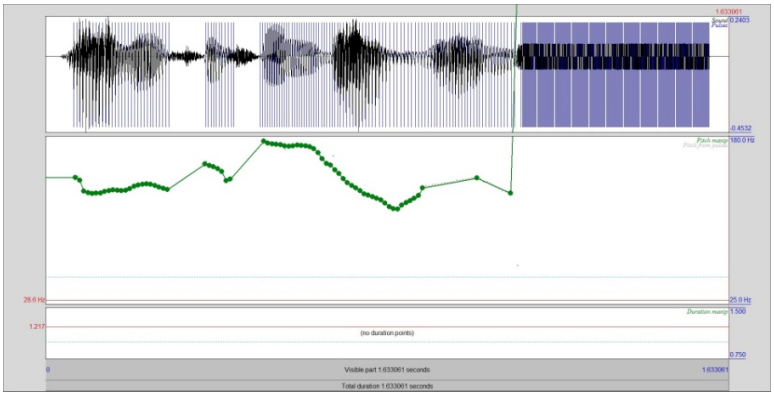


Figure 17 Removing all pitch points in the interval except for the rightmost one

In order to manipulate the f0 maximum point, the maximum point was selected and the f0 at that point was manipulated automatically using the 'shift pitch frequencies' function as shown in figure (18).

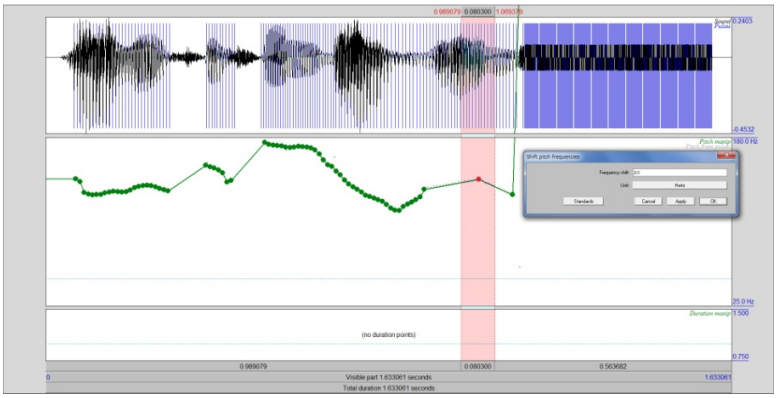


Figure 18 Manipulation of the maximum f0 using the function 'shift pitch frequencies'

For manipulating the rightmost point, this point was selected and the f0 at this point was raised using the 'shift pitch frequencies' function. Finally the manipulated sounds were published as a resynthesized object, and were saved as a distinct .wav file.

Figure (19) shows an original pitch curve (bottom) and a manipulated pitch curve (top) in which the maximum point of the interval has been raised 20 Hz.

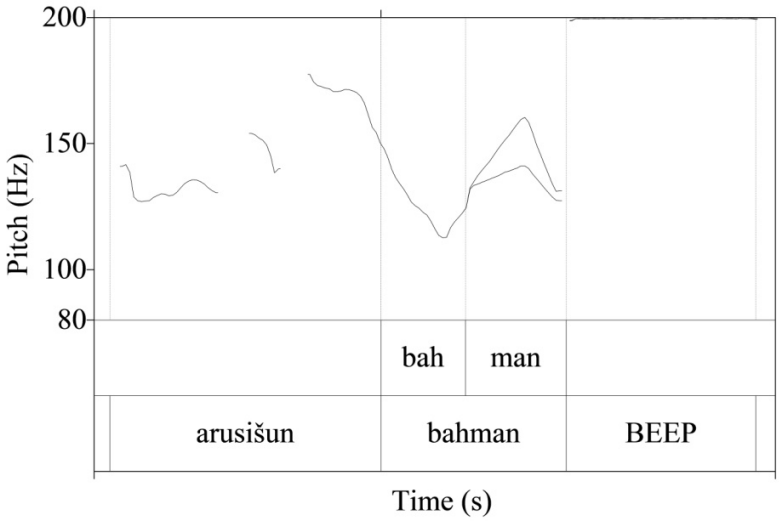


Figure 19 Original and manipulated f0 curves. The maximum f0 point of the interval has been raised 20 Hz

In figure (20), the f0 at the right edge of the manipulation interval has been manipulated and increased 20 Hz. In the original curve (the lower one), the f0 peak is at the middle of the interval and the right edge has a falling f0, however, in the manipulated version (the one at the top), the peak is at the right edge of the interval, where the f0 has a rising slope.

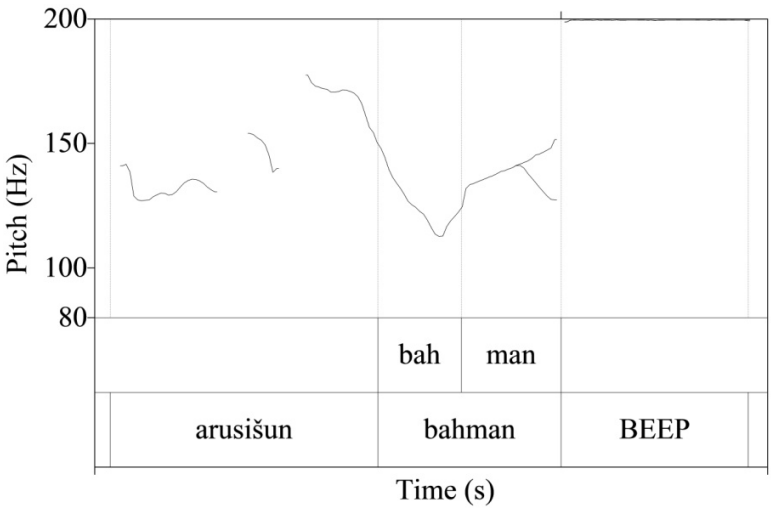


Figure 20 Original and manipulated f0 curves. The rightmost f0 point of the interval has been raised 25 Hz

Duration

To manipulate duration automatically using Praat, the sound files were opened as manipulation objects, and the manipulation intervals were selected. Next, two duration points were added at the left and right edges of the intervals as shown in figure (21).

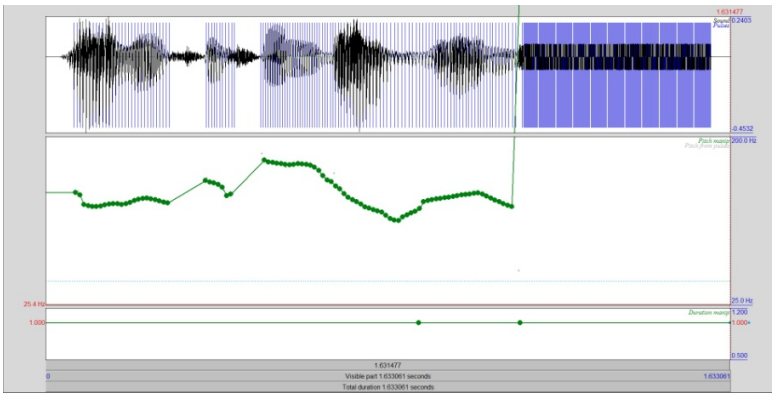


Figure 21 Selecting the manipulation interval in Praat manipulation object and adding duration points at its right and left edge

Next, the closest possible point to the left edge in the interval was selected, and the duration at this point was set at a value by which the duration should be multiplied as demonstrated in figure (22).

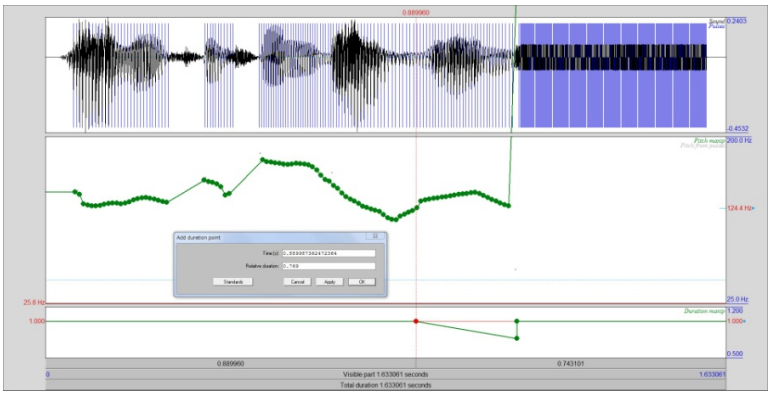


Figure 22 Adding duration point at the closest possible point to the left edge in the interval

Similarly, for the right edge, the closest possible point to the edge in the interval was selected, and a duration point was added at this point with the same value as the additional point at the left edge as shown in figure (23).

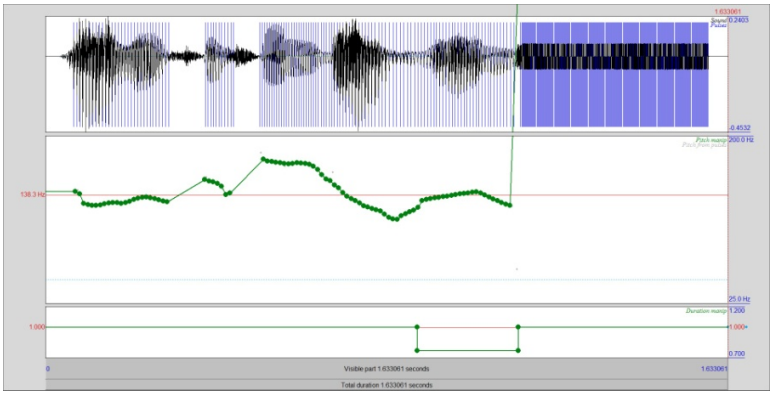


Figure 23 Shortening the duration of the manipulation interval by adding duration points at its two edges

Finally, the manipulated sound was published as a resynthesized object and was saved as a distinct .wav file. In the examples in figures (21-23), the value of duration points are 0.769,

thus, the duration of the manipulation interval will be multiplied by 0.769. For instance, if the duration of the manipulation interval in the original sound is 241ms, the duration of the interval in the manipulated sound will be around 185ms.

Masking

The beep sound used to mask words in this experiment was a bleeping sound commonly used in the media to censor profanity or classified information. The frequency of the beep sound was 250 Hz and the intensity was modified and set at 72dB which was the average intensity of the speaker's voice in the center of syllables. The length of a beep sound was set as the average length of the two words (the copula and the transitive verb) that were masked. The masked words were selected and cut thoroughly, and were replaced by the prepared beep sound.

Praat Experiments

All the stimuli were presented to the participants using a Praat identification experiment. Participants listened to the stimuli played on a laptop computer using a headphone in a soundproof booth. For each stimuli set, first the participants encountered a page on the computer screen with the directions explaining how the experiment is done. After clicking on the screen, the first stimulus was played and at the same time, a page appeared with two reply buttons at the center of the screen. A copula was written on one of the reply buttons in Persian alphabet, and a transitive verb on the other, and the participants had to to click and choose the appropriate word for the masked portion of the utterances they heard. Each stimulus was played two times, and the stimuli were played randomly using the randomizing strategy 'Permute balanced' that ensures that the same stimulus is never applied twice in a row. Figure (24) illustrates the cover page of the experiment.

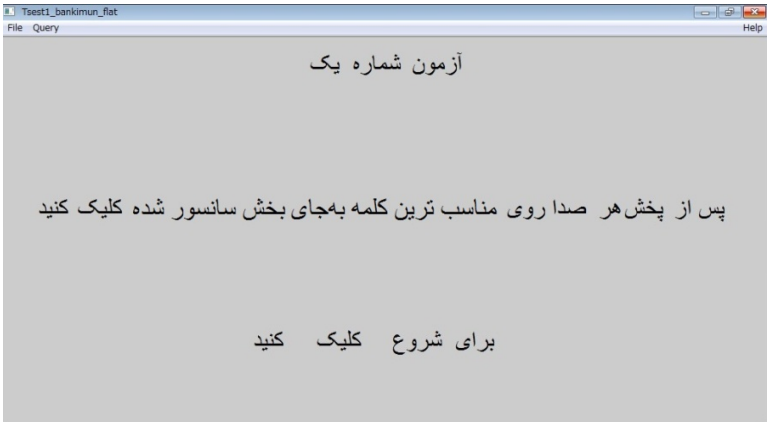


Figure 24 The cover page of experiments

Figure (25) shows the page which appeared on the breaks between the stimuli. By clicking on this page the participants could hear the next stimulus.

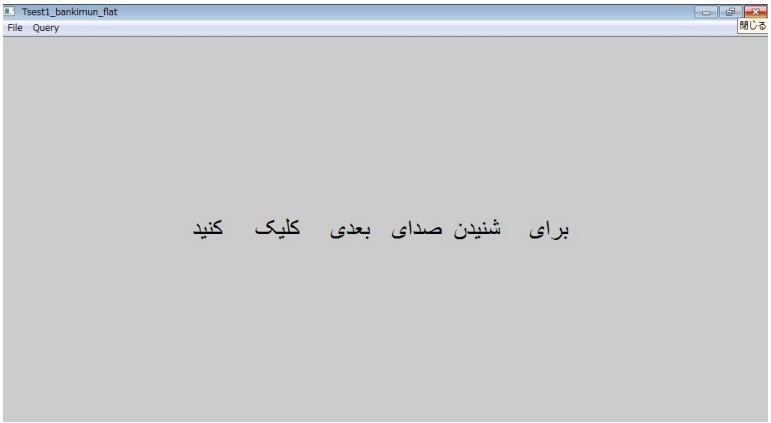


Figure 25 The break page between the stimuli

The illustration in figure (26) shows the page which appeared simultaneously with a stimulus. The participants had to click on one of the reply buttons in yellow.

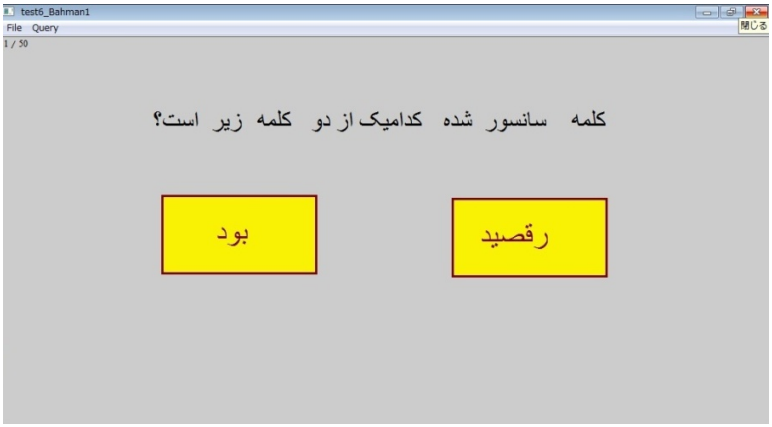


Figure 26 An example of a page in which participants responded to the stimuli

Participants

Participants were 10 native Persian speakers, five female and five male. Their age ranged from 18 to 37, with an average of 31.01 years. All were monolinguals born and raised in Iran, but some of them had studied English and Japanese in their adulthood. They were all unaware of the purpose of the study.

5.3.2 Test A

In this test, the maximum f0 and the f0 on the final point of the manipulation intervals of the nuclear accented words were increased and decreased 25 Hz. This value was determined based on the production experiments of the speaker who read the original stimuli. The duration of the manipulation intervals of the nuclear accented words were also increased (multiplied by 1.30) and decreased (multiplied by 0.769). These manipulated stimuli together with the original ones were presented to the participants to examine if these manipulations will lead the participants to perceive the nuclear accented words as pre-nuclear accented ones. There were eight types of stimuli in this test, and the two sentences

in (11i) and (12i) were used to make the stimuli. Each stimulus was repeated two times in the test, and 10 participants took the perception test. Therefore, all together there were 48 stimuli and 480 responses were obtained. Table (18) presents a complete list of the stimulus types used in test A. For each stimulus type in table (18), 40 responses were obtained and analyzed. Stimulus number 1 is the original (non-manipulated) utterance with the verb being masked by a beep sound and the target word having a nuclear accent. Stimulus number 2 is similar to the stimulus 1, except that in stimulus 2 the target word carries a pre-nuclear accent. Stimuli 3 and 4 are manipulated versions of stimulus 1, in which the duration of the final syllable of the target word has been multiplied by 0.769 and 1.30 respectively. Stimuli 5 and 6 are also manipulated versions of stimulus 1 in which the maximum f0 point in the final syllable of the target word has been raised and lowered 25 Hertz respectively. Stimuli 7 and 8 are very similar to stimuli 5 and 6, except that in 7 and 8 the f0 of the rightmost point of the final syllable of the target words has been raised and lowered 25 Hertz respectively.

Table 18 An inclusive list of the stimulus types used in test A

No.	Stimulus	No.	Stimulus
1	Original nuclear	5	Nuclear, max. f0 plus 25 Hz
2	Original pre-nuclear	6	Nuclear, max. f0 minus 25 Hz
3	Nuclear, dur. multiplied by 0.769	7	Nuclear, final f0 plus 25 Hz
4	Nuclear, dur. multiplied by 1.3	8	Nuclear, final f0 minus 25 Hz

When the participants chose the transitive verb as their answer, the response was counted as a “pre-nuclear” response, because if the masked portion is a transitive verb, the target word would bear a pre-nuclear accent. When a participant clicked on the copula, the response was recorded as a “nuclear” response, since if the masked portion is a copula, the target word which immediately precedes the masked portion will be associated with a nuclear accent.

Results and discussions

The first important issue in the results of this test was the fact that the participants succeeded in distinguishing the two types of accents without listening to the post-accentual portion of the utterance. When the verbs were masked by a beep sound, 31 out of 40 (77.5 %) of the responses to the stimuli containing pre-nuclear accented word were correctly "pre-nuclear", while there were only three (7.5%) "pre-nuclear" responses to the stimuli containing non-manipulated nuclear accented word. This shows that native listeners of Persian make use of the phonetic differences between nuclear and pre-nuclear accents to distinguish the two types of accents. The only manipulation that could eventually lead the participants to perceive the nuclear accent as a pre-nuclear one was rising the f0 on the right edge of the final syllable in the target word, which lead 75% of the responses to this type of stimulus to be "pre-nuclear". The results of this test are shown in figure 27.

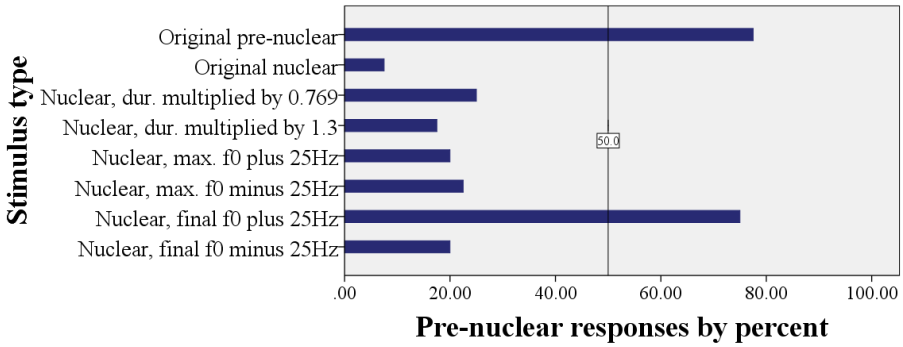


Figure 27 The percentage of "pre-nuclear" responses to the original and manipulated stimuli

Since the data obtained in this experiment were either "nuclear" or "pre-nuclear" responses to a set of stimuli, they should be regarded as nominal binary data. The aim of the statistical analyses of these data is to investigate whether the proportion (or percentage) of a specific response to a particular stimulus is significantly different from the same responses to the other stimulus or not. In statistical analyses, in similar cases, often a Pearson's Chi-squared

test is performed. However, according to Landau and Everit (2004) among others, when there are small sized samples, Chi-squared test is not valid anymore and should be replaced by exact tests. Some of the samples in the results of this experiment were very small in size. For example, there were only three "pre-nuclear" responses to the original nuclear stimulus. Therefore, a pairwise Fisher's exact test was performed to investigate if any of the manipulations leads to a significant difference in the responses. The null hypothesis of this test is that the responses to the original nuclear stimuli and the responses to the manipulated versions of the same stimuli are independent, i.e., there are no significant differences between the responses to the original stimuli and the responses to manipulated ones. Therefore, in this pairwise test, the responses to each manipulated stimulus were compared with the responses to the original non-manipulated nuclear stimulus. If the p-value in any of these tests is smaller than the significance level (0.05), then there is a significant difference between the responses to the stimuli in that test and the non-manipulated original stimuli. Table (19) demonstrates the results of the Fisher's exact tests. In order to control the familywise error rate, the tests were conducted using Bonferroni adjusted alpha levels of 0.00714 per test (0.05/7).

Table 19 Results of pairwise Fisher's exact tests for test A

Pairwise comparisons in the Fisher's Exact test	**Two sided p-value**
Original nuclear & Original pre-nuclear	$p<0.0001$
Original nuclear & Nuclear, dur. multiplied by 0.769	0.066
Original nuclear & Nuclear, dur. multiplied by 1.3	0.31
Original nuclear & Nuclear, max. f0 plus 25 Hz	0.1927
Original nuclear & Nuclear, max. f0 minus 25 Hz	0.1149
Original nuclear & Nuclear, final f0 plus 25 Hz	$p<0.0001$
Original nuclear & Nuclear, final f0 minus 25 Hz	0.1927

As could be predicted from the descriptive analysis, the only group that showed a significant effect was the stimuli in which the f0 on the right edge of the manipulation interval was raised 25 Hz. The majority of responses to this manipulated nuclear accented

stimulus were “pre-nuclear”, showing that raising the f0 at the right edge of the final syllable in a nuclear accented word will lead the listeners to perceive it as a pre-nuclear accented one. These results also show that manipulating the maximum f0 point of the final syllable, or the duration of the final syllable does not lead to a significant change in listeners’ perception. As discussed at the beginning of this chapter, the delay of f0 peaks in pre-nuclear accented words, and the lack of this delay on nuclear accented words, leads to a difference between the shapes of the pitch curve on the final syllables of the words bearing theses two types of accents. The f0 is falling at the right edge of the final syllable of a nuclear accented word, while it is rising at the right edge of the final syllable of a pre-nuclear accented word. Raising the f0 at the right edge of the final syllable of a nuclear accented word (25 Hz for this speaker) alone will lead the listeners to perceive the nuclear accented word as a pre-nuclear accented word. Therefore, it seems that the most important phonetic difference between a nuclear accent and a pre-nuclear one is the shape of the pitch curves on the words which are associated with these two types of accents. However, two questions remain: first, precisely how much rise of the f0 on the right edge is enough to render a nuclear accented word to be perceived as a pre-nuclear accented one? Second, the production experiments showed that unlike intensity, spectral tilt or vowel quality, duration of the final syllable was significantly different in nuclear and pre-nuclear accented words. Nuclear accented words had substantially longer final syllables than pre-nuclear accented ones. However, this perception test revealed that shortening the final syllable of nuclear accented words does not render them to be perceived as pre-nuclear accented words. The question is whether the duration is completely redundant, or does the change in the duration can have some affect on the perception, if it is accompanied by the change in the f0 of the right edge. Test B is conducted to answer these two questions.

5.3.3 Test B

In this test, two sets of stimuli were presented to the participants using a Praat experiment. The first set consisted of utterances containing a target word associated with a nuclear accent and the post-target verb being masked by a beep sound. The f0 on the right edge of the final syllables of the target words were manipulated and raised gradually in five steps. The original pitch curve and the five manipulated versions of it can be seen in figure (28). On the final syllable of the target word *bahman*, the lowest f0 curve is the original one, and the five curves above it are the manipulated ones. In the original curve, the f0 peak of the syllable is in the middle of it, while in the manipulated curves, the peak gradually moves to the right edge of the syllable, as is common in pre-nuclear accented words.

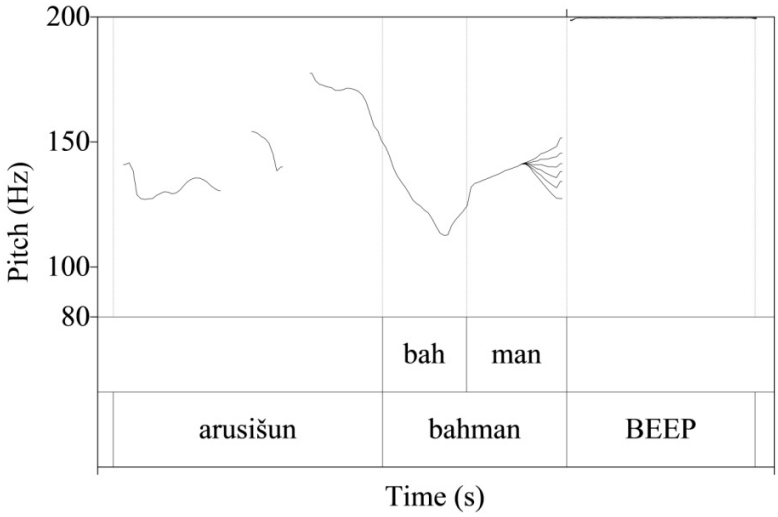


Figure 28 Original and gradually manipulated f0 curves at the right edge of the syllable

The second set was very similar to the first one, except that in the second set, first the duration of the final syllable of the target word was manipulated and shortened (multiplied by 0.769) and then, the right edge of the final syllables were raised gradually in five steps.

Table (20) shows the list of stimuli used in this experiment. In this test, 480 responses were obtained, and there were 40 responses for each of the stimuli types in table (20).

Table 19 An inclusive list of the stimulus types used in test B

No.	Stimulus
1	Original Nuclear accented
2	Nuclear, final f0 plus 5 Hz
3	Nuclear, final f0 plus 10 Hz
4	Nuclear, final f0 plus 15 Hz
5	Nuclear, final f0 plus 20 Hz
6	Nuclear, final f0 plus 25 Hz
7	Nuclear, duration multiplied by 0.769
8	Nuclear, duration multiplied by 0.769 and final f0 plus 05 Hz
9	Nuclear, duration multiplied by 0.769 and final f0 plus 10 Hz
10	Nuclear, duration multiplied by 0.769 and final f0 plus 15 Hz
11	Nuclear, duration multiplied by 0.769 and final f0 plus 20 Hz
12	Nuclear, duration multiplied by 0.769 and final f0 plus 25 Hz

Results and discussions

The most important finding of test B was the fact that duration does affect the perception of nuclear and pre-nuclear accents. Manipulated nuclear accented words are perceived as pre-nuclear accented ones much easier, if the pitch manipulations are accompanied by shortening of the final syllable. This can be observed in the graph in figure (29).

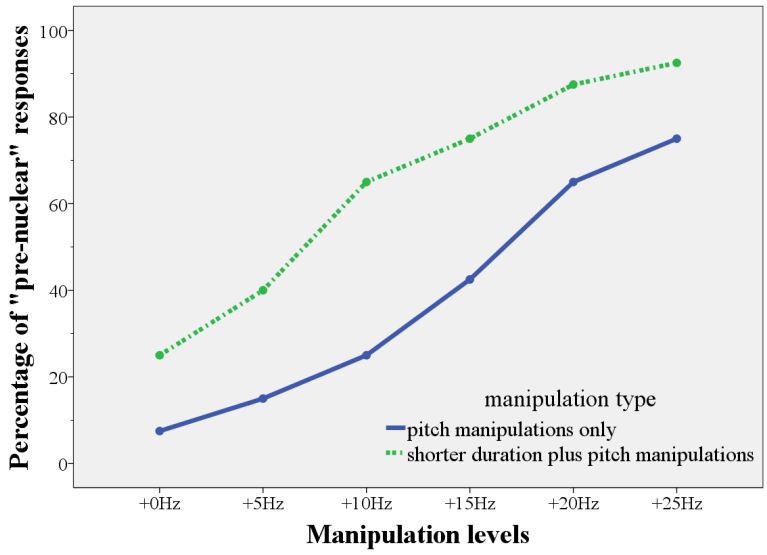


Figure 29 The effect of gradual increase of pitch at the right edge of the stressed syllable of a nuclear accented word, with and without manipulating the duration

The blue solid line shows the percentage of "pre-nuclear" responses to the nuclear accented stimuli in which only the f0 is manipulated in five levels. The green dotted line indicates the percentage of "pre-nuclear" responses to the manipulated nuclear stimuli in which first the final syllable of the target word is shortened, and then the f0 is manipulated in five levels. As can be seen, the green curve is located above the blue one, which suggests that the difference in duration, when accompanied by the difference in the shape of the pitch curve, can help the listeners discriminate a nuclear accent from a pre-nuclear one more effectively. This is consistent with the findings in production experiments that the duration increase in the stressed syllable of nuclear accented words is significantly more than that in pre-nuclear accented ones. The results of test B are given in table (21).

Table 20 The descriptive results of test B

F0 manipulation levels	Pitch manipulations only		Shortened duration with pitch manipulations	
	"pre-nuclear" responses	"nuclear" responses	"pre-nuclear" responses	"nuclear" responses
No manipulation of f0	3	37	10	30
Right edge f0 plus 05 Hz	6	34	16	24
Right edge f0 plus 10 Hz	10	30	26	14
Right edge f0 plus 15 Hz	17	23	30	10
Right edge f0 plus 20 Hz	26	14	35	5
Right edge f0 plus 25 Hz	30	10	37	3

Fisher's exact tests were conducted to examine if the difference between the responses to the two types of stimuli at the same level of f0 manipulation are statistically significant or not. For instance, the number of "nuclear" and "pre-nuclear" responses to the stimuli in which duration was shortened and f0 was raised 5 Hz at the right edge was compared with the number of "nuclear" and "pre-nuclear" responses to the stimuli in which only f0 was raised 5 Hz at the right edge using a pairwise Fisher's exact test, and the same comparison was done in all f0 manipulation levels. Table (22) shows the results of the exact test.

Table 21 Results of pairwise Fisher's exact tests for test B

F0 manipulation level	Two sided p-value
No manipulation of f0	0.066090
Right edge f0 plus 05 Hz	0.022990
Right edge f0 plus 10 Hz	0.000648
Right edge f0 plus 15 Hz	0.006023
Right edge f0 plus 20 Hz	0.033890
Right edge f0 plus 25 Hz	0.066090

With a p-value threshold of 0.05, responses to the two types of stimuli differ significantly in four out of six levels of f0 manipulations. When there was no manipulation of f0, and when f0 was raised 25 Hz, (shown in gray shaded cells in table (22), the number of "pre-nuclear" responses to the two types of stimuli are not significantly different.

Further statistical tests were conducted to investigate at which level the responses change from "nuclear" to "pre-nuclear". If we concentrate on the stimuli with f0 manipulations only, when the pitch is raised 5 Hz, no significant difference is observed in responses to this stimulus, comparing when there was no manipulations ($p=0.4814$). When the f0 is raised 10 Hz, still no significant difference is observed ($p=0.06609$). However when the f0 is raised 15 Hz, there is a significant difference in responses to this manipulated stimulus, comparing when there was no manipulations ($p<0.001$). The same tests on the stimuli in which the duration of final syllables were shortened had different results. Although when the pitch is raised 5 Hz, no significant difference is observed in responses to this stimulus, comparing when there was no manipulations of f0 ($p=0.2324$), when the f0 is raised 10 Hz, there is a significant difference in responses to this manipulated stimulus, comparing when there was no manipulations of f0 ($p<0.001$). shows that for the voice of this speaker, if only f0 at the right edge of the final syllable is manipulated, the listeners will start to perceive the nuclear accented word as pre-nuclear accented one, when the f0 has been raised around 15 Hz. However, if the final syllables is shortened first, the listeners will start to perceive the nuclear accented word as pre-nuclear accented one, when the f0 has been raised only around 10 Hz. This proves that the duration of the syllable which is associated with the accent can have a role in discriminating the type of accent.

5.4 General discussions and summary

The main research question of this chapter was whether nuclear accented and pre-nuclear accented words differ phonetically, or what makes the nuclear accent to be perceived more prominently is merely the fact that it is the final accent in an IPhrase as suggested by

Bolinger (1986) and Ladd (2008). The perception experiments in this study showed that native listeners can indeed discriminate the types of accents without having access to the portion of utterance that occurs after the accented word. This shows that there are phonetic differences between the two types of accents that can help listeners distinguish them from each other.

Silverman and Pierrehumbert (1990) have shown that nuclear accented words and pre-nuclear accented words in English differ phonetically in that, in pre-nuclear accented words f0 peaks occur after the syllable with which they are intuitively associated. In Persian, the f0 peaks in pre-nuclear accented words occur at the enclitics of the word if there are any, at the right edge of the word, or at the left edge of the following word. In nuclear accented words however, f0 peaks are timed with the metrically strong (stressed) syllable of the word and tend to occur in the middle of the stressed syllables.

Sadat-Tehrani (2007) suggests that non-final Accentual Phrases (accent bearing units) are right-aligned with a high boundary tone, while the final AP is aligned with a low boundary tone. The present study suggests that the high or low AP boundary tones which are observable in the pitch curve of utterances can be interpreted as peak delays on pre-nuclear accented words and lack of delay on nuclear accented words. Figure (30) shows the pitch curve of an utterance of the non-word *dandandan*, uttered in a pre-nuclear accent bearing position by a female native speaker. This utterance is taken from the recorded data of the production experiment reported earlier in this chapter. As can be seen, the f0 on the final syllable which is metrically the strongest syllable rises abruptly, and the maximum point of the f0 curve on this syllable is at the right edge of the syllable.

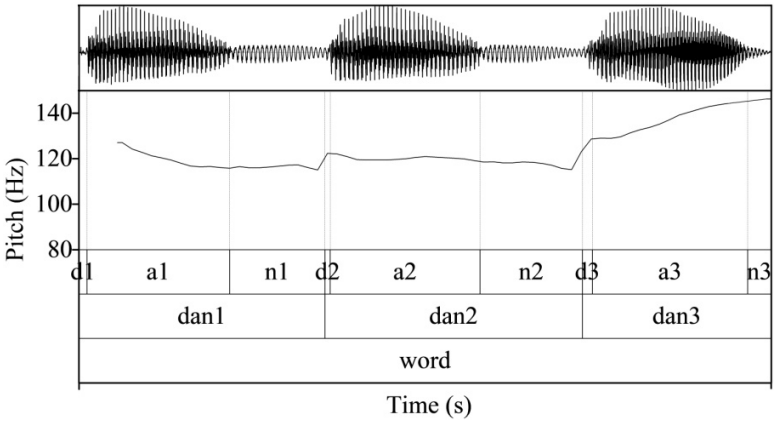


Figure 30 The non-word *dandandan* uttered by a female speaker in a pre-nuclear accent bearing position

Figure (31) on the other hand, shows the pitch curve of an utterance of the same non-word uttered in a nuclear accent bearing position by the same speaker. As can be seen, the maximum point of the f0 on the final syllable is in the middle of the syllable, and the right edge of the curve has a falling slope.

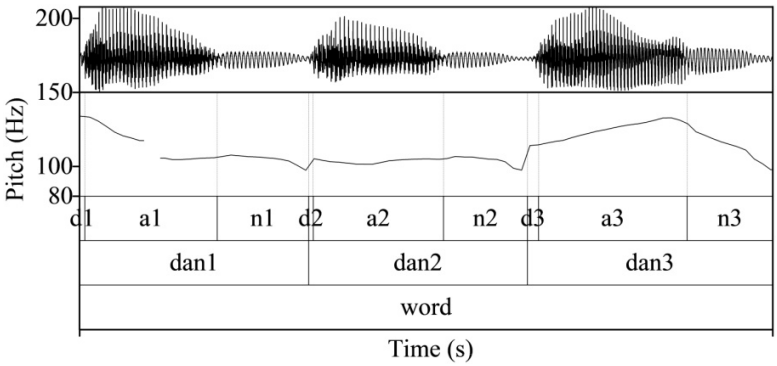


Figure 31 The non-word *dandandan* uttered by a female speaker in a nuclear accent bearing position

This seems to be the most important difference between a nuclear and a pre-nuclear accent. The perception tests reported in this chapter revealed that if in a nuclear accented word similar to the one in figure (31), we manipulate the pitch curve on the final syllable, and

raise the right edge of the curve, so that the f0 peak occurs at the right edge, the native listeners will perceive this word as a pre-nuclear accented word.

Please note that as discussed in Mahjani (2003) and Sadat-Tehrani (2007), the general type of accent in Persian is an accent-lending rise which consists of an f0 peak preceded by a valley (transcribed as L+H* in ToBI system). The fact that the right edge of the f0 curve has a falling slope in nuclear accent, but a rising slope in pre-nuclear ones, should not be confused with the distinction between accent-lending rise (rising pitch accent) and accent-lending fall (falling pitch accent) as argued in Caspers and van Heuven (1993) and other works, because both pre-nuclear and nuclear accents in Persian are accent-lending rises which are composed of a peak preceded by a valley. However, while the peak is aligned with the accent bearing syllable in nuclear accents, it may occur with a delay in pre-nuclear accents.

Prieto et al. (1995) in an experiment on Spanish language with read-aloud tasks show that f0 peaks before prosodic boundaries (PWord, Intermediate Phrase and IPhrase) tend to be placed earlier in their syllables in this language. Arvaniti (1998) in an experiment on Greek demonstrates that in pre-nuclear accents, the alignment of the high tone may be affected by "tonal crowding", when the accented syllable is close to the end of the word. There are also a number of other studies pointing out the relation between alignment or anchoring of f0 peaks and prosodic boundaries. These studies suggest that further distance between an f0 peak and the following prosodic boundary may trigger delays in f0 peaks. However, unlike these languages, it seems impossible to regard the peak delay to be a consequence of the distance between the peak and the following prosodic boundary, because both nuclear and pre-nuclear accents can equally be boundary adjacent or detached from the boundary. At PWord level, both nuclear accent bearing syllables and pre-nuclear accent bearing ones are PWord-final in non-cliticized words, but non-PWord-final in cliticized words. At PPhrase level, both types of accents can occur PPhrase-finally only if the PPhrase consists of a single non-cliticized PWord. Only nuclear accents can appear IPhrase-finally, only when the rightmost PPhrase in the IPhrase consists of a single non-cliticized PWord. Therefore,

since none of the two accent types necessarily occur close to boundaries, it is not possible to postulate any relations between the type of accent and boundary adjacency.

It also seems impossible to attribute the difference between pre-nuclear and nuclear accents which as we saw, arises from miss-alignment between pre-nuclear accent peaks and the syllables with which they are associated, to other tones or phrasal accents, because no such tones or phrasal accents have been shown to exist in Persian.

Production experiments in this study showed that the f0 maxima on the final syllables of pre-nuclear accented words are significantly higher than the f0 maxima on the final syllables of nuclear accented words. This seems to be in contrast with the fact that nuclear accented words are perceived more prominently than pre-nuclear accented ones. However, perception experiments revealed that raising the f0 maxima of the final syllable of a nuclear accented word will not render it to be perceived as a pre-nuclear accented word. Liberman and Pierrehumbert (1984) have established an important generalization in respect of the phonetic course of sequences of downstepping pitch-accents. According to this generalization, in such a course, the final accent is considerably lower than expected. Liberman and Pierrehumbert (1984) call this effect 'final lowering', and model it as an effect of lowering overlaid in final position. In the production experiments of this study, the target words were controlled for downstep effects (The number of accents preceding the target word, and the syllable based distance between the target word and the initial point of the utterance was exactly the same in all utterances), therefore, it would be possible to ascribe the low peaks on nuclear accents, not to downstep, but to final lowering as suggested in Liberman and Pierrehumbert (1984). The f0 peaks on nuclear accented words are lower than those on pre-nuclear accented words, because nuclear accents are always the final peaks in the utterance.

Production experiment also showed that the duration of a syllable associated with a nuclear accent is significantly longer than the duration of a syllable associated with a pre-nuclear accent. The perception tests revealed that shortening the duration of the final syllable of a

nuclear accented word alone cannot render the listeners to perceive it as a pre-nuclear accented word. However, if this duration shortening is accompanied by the rise of f0 at the right edge of the final syllable, listeners will perceive the word as a pre-nuclear accented one more efficiently. In other words, duration can cue the difference between the two types of accents, only when it is accompanied by a change in the shape of the pitch curve on the final syllable. The perception tests also proved that moving the f0 peak on the final syllable of a nuclear accented word from the middle of the syllable to its right edge, will more effectively make the word to be perceived as a pre-nuclear accented word, if it is accompanied by a decrease in the length of the syllable.

The results of production experiments made it clear that there are no differences between a syllable associated to a nuclear accent, and a syllable associated to a pre-nuclear accent in terms of overall intensity, spectral tilt and vowel quality.

To summarize, the production and perception experiments reported in this chapter revealed that in nuclear accented and pre-nuclear accented words, the shape of the pitch curve on the final syllable differs substantially, and this can cue the difference between the two accents. These experiments also showed that the f0 peak on pre-nuclear accented words is higher than the f0 peak on nuclear accented words, but this difference does not seem to have any effect on the perception. Another finding of this experiment was that syllables associated with nuclear accent are significantly longer than syllables associated with pre-nuclear accents and this difference in duration, when accompanied by the change in the shape of f0 curve can help the listeners distinguish the type of the accent.

In conclusion, nuclear and pre-nuclear accents are predictable by prosodic structure which itself reflects the syntactic structure, in information-neutral sentences. Pre-nuclear accents associate to PPhrase heads and nuclear accents associate to IPhrase heads. The f0 peaks of nuclear accents are timed with accent bearing syllables which are comparatively long, while pre-nuclear peaks appear with a delay on relatively shorter syllables. However, this delay or the difference in length cannot be predicted by boundary adjacent effects or other

tones or phrasal accents. In perception experiments there were minimal pairs that differed in one nuclear accent as opposed to a pre-nuclear one. In such cases, if the phonetic characteristics of nuclear accent are manipulated to resemble those of pre-nuclear ones, the nuclear accent bearing word will be perceived as pre-nuclear accent bearing one. All these considerations lead us to regard the difference between the two types of accents in Persian to be a phonological one. In other words, nuclear accents and pre-nuclear accents are essentially different types of accents, and one cannot reduce the differences between the two types to the fact that nuclear accents occur after the pre-nuclear ones. It seems more appropriate to postulate two different constraints that require association between prosodic categories and accents. One constraint (ASSOC-PRENUCAC-φ) requires PPhrase heads to be associated with pre-nuclear accents, and another constraint (ASSOC-NUCAC-ι) requires IPhrase heads to be associated with nuclear accents.

Chapter 6 Conclusions

6.1 Summary of proposals and concluding remarks

This dissertation began with foot structure and suggested that in languages like Persian which has non-iterative weight-insensitive edgemost stress, a single foot which is edge-aligned with the minimal PWord best explains the prosodic pattern at word level. Similar foot structures have been proposed in the literature for similar language, like Baković's (2004) analysis of Uzbek. This study addressed the prosodic differences between lexical words and weak function words and showed that any attempt to describe the prosodic structure of Persian without taking this difference into consideration, will not be able to explain for a wide range of phenomena. Prosodic clitics have been the subject of many studies on Persian, however, all of these studies focus on enclitics and the prosodic structure of proclitics has not been specifically explored. This study pointed out theoretical and empirical problems with existing accounts on clitics in Persian, specifically Amini (1997) and Kahnemuyipour (2003), and concentrated on the asymmetrical behavior of proclitics and enclitics and based on phonological arguments and phonetic evidence showed that enclitics prosodize as affixal clitics, while proclitics are free clitics. The present study next dealt with XP-external function words and showed that the direction of phrasing in these clitics is decided by phonological well-formedness constraints. It also addressed the issue of clitic clusters in Persian which was not dealt with in the previous works. Another contribution of this study was its reclassification of exceptionally initial-stressed words.

This dissertation addressed the longstanding problem of the prosodic structure of Ezafe constructions. Kahnemuyipour (2003) had suggested that the entire Ezafe construction maps onto a single PWord. The present study argued that since each lexical word present in

the Ezafe construction has a distinct prosodic prominence, and since prominence is culminative in Persian at word level, this account cannot be accepted. Instead, based on syntactic analyses in Mahootian (1993) and Moinzadeh (2005), it proposed that each lexical word in an Ezafe construction maps onto a PPhrase, and the Ezafe morpheme phrases with its preceding material to satisfy the phonological well-formedness constraint ONSET. The prosodic structure of XP-external clitics was explained by adopting a syntax-prosody interface constraint, namely, MAP-XP that bans two sister XPs inside a single PPhrase. The prosodic structure of all types of syntactic phrases in Persian was explained by ranking this constraint among other syntax-phonology interface constraints and prosodic markedness constraints. Finally, it was proposed that a constraint requires PPhrase heads to be associated with audible accents. Table 1 shows prosodic categories that correspond to some morpho-syntactic category, and their characteristics by which these categories can be identified. PWords in Persian can have recursive structures, and some characteristics of maximal PWords differ from those of minimal PWords.

Table 1 Prosodic categories that correspond with morpho-syntactic categories, and their unique characteristics

IPhrase	Is the domain of downstep. Corresponds to sentences (IP/CP). Associates with one and only one nuclear accent. Is right-headed.	
PPhrase	Corresponds to syntactic phrases (XPs). Associates with one and only one audible accent. Is left-headed.	
PWord	max	Corresponds to a lexical word and its enclitics. Syllabification is obligatory and metrical stress is obligatory and culminative.
	min	Corresponds to lexical words (X^0's). Metrical stress is obligatory and culminative. Right-aligns with a foot.

The Phonetic part of this dissertation dealt with the question whether pre-nuclear and nuclear accents differ phonetically or not. Previous works such as Bolinger (1986) and Ladd (1996) had suggested that these two types of accents are not phonetically different in English, while Silverman and Pierrehumbert (1990) showed that nuclear accents are

phonetically timed with the syllables they are associated with, but f0 peaks in pre-nuclear accents appear later than the syllable they are associated with. This study suggested that the difference between the shapes of f0 curve on syllables associated with nuclear and pre-nuclear accents that were studied and attributed to AP boundary tones in Sadat-Tehrani (2007), can be reanalyzed as an effect of peak delay on pre-nuclear accented syllables.

The results of production experiments revealed that nuclear accented syllables have a lower f0 range, but a longer duration in comparison with pre-nuclear accented syllables. Other parameters such as overall intensity, spectral tilt and vowel quality do not differ significantly in the two types of accents. Perception experiments revealed that native listeners can indeed distinguish the two types of accents without having access to the portion of the utterance that follows the accent. This proves that the two accents are phonetically different. Perception tests also showed that the difference between the shapes of f0 is the main acoustic parameter that helps the listeners distinguish the two types of accents. In pre-nuclear accented words, as can be seen in the schematic idealized curve in figure 1, the f0 peak is at the right edge of the syllable, and the curve has a rising slope at this point. In pre-nuclear accents, the peak may occur on the initial syllable of the following word. In the syllables associated with nuclear accents as is shown in the schematic idealized curve in figure 2, the f0 peak is inside the syllable, and the curve has a falling slope at the right edge of the syllable.

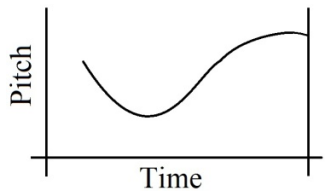


Figure 1 An idealized schematic f0 curve on a pre-nuclear accented syllable

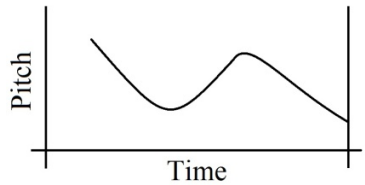


Figure 2 An idealized schematic f0 curve on a nuclear accented syllable

When the f0 at the right edge of nuclear accented syllables was manipulated and raised to move the f0 peak to the right edge, the participants of the perception experiment perceived

the word containing this syllable as a pre-nuclear accented word. It also became clear that duration alone cannot cue the difference between the two types of accents, however when accompanied by f0 changes, it can help the listeners distinguish the two accents more easily and more efficiently.

6.2 Suggestions for further studies

This study only dealt with simple sentences of Persian in which the sentence is mapped onto an IPhrase which is an utterance at the same time. Further investigations are needed to determine the prosodic structure of compound sentences within the framework of Prosodic Phonology.

The particle *ke* in some of its usages seems to phrase freely with both its preceding and following material, and behave like a floating clitic. This particle was not dealt with specifically in this study, and it seems that further studies should be done on classification of different types of *ke*, and their syntactic and prosodic structures.

This study adopted the OT constraint MAP-XP, to explain the prosodic structure of syntactic phrases. It seems that typological work is needed to examine whether this constraint can meet the similar needs in other languages, and whether existence of such constraint is typologically grounded or not.

The phonetic part of this dissertation showed that the difference between the shapes of f0 curves on syllables associated with nuclear and pre-nuclear accents can cue the difference between the two accents, and that duration changes when accompanied by the appropriate changes in f0, can have a role in distinguishing the two accents. Further investigations on various types of languages are needed to see whether these phonetic differences between final and non-final accents are universal, or limited to specific languages.

The delay of f0 peaks on nuclear accented words should be investigated in detail in future studies. It should be made clear when the pre-nuclear accent peak occurs on the right edge of the syllable, and when it appears on the first syllable of the following word. The relation between speech rate and the delay of pre-nuclear peaks should also be studied.

References

Abolghassemi, Mohsen (Abolqâsemi, Mohsen) (1387 [2008–2009]). *Dastur-e târixi-ye moxtasar-e zabân-e fârsi [A Short Historical Grammar of the Persian Language]* (6th ed.). Tehran: SAMT Publications.

Abolhasanizadeh, Vahideh, Bijankhan, Mahmood, and Gussenhoven, Carlos (2012). The Persian Pitch Accent and its Retention after the Focus. *Lingua*, *122*(13), 1380–1394.

Abolhasanizadeh, Vahideh, Gussenhoven, Carlos, and Bijankhan, Mahmood (2010). The Position of Clitics in Persian Intonational Structure. In *Proceedings of the Speech Prosody Conference*. (pp. 1-4). Chicago.

Amini, Afsaneh (1997). On Stress in Persian. *Toronto Working Papers in Linguistics*, *16*(1), 1–20.

Anderson, Stephen (2005). *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.

Arbisi-Kelm, Timothy (2007). An Analysis of the Intonation of Complex Sentences in Farsi. *UCLA Working Papers in Phonetics*, *105*, 35–50.

Arvaniti, Amalia, Ladd, Robert, & Mennen, Ineke (1998). Stability of Tonal Alignment: The Case of Greek Prenuclear Accents. *Journal of Phonetics*, *26*(1), 3-25.

Atai, Parvin (1964). *A Contrastive Study of English and Persian Question Signals*. (Doctoral dissertation, University of Michigan).

Bader, Christopher B. (2001). *Givenness, Focus, and Prosody* (Doctoral dissertation, Massachusetts Institute of Technology).

Bakovic, Eric (2004). Unbounded Stress and Factorial Typology. In John McCarthy (Ed.), *Optimality Theory in Phonology: A Reader* (pp. 202–214). London: Blackwell.

Balusu, Rahul (2010). Acoustic Correlates of Stress and Accent in Telugu. 21st South Asian Languages Analysis Roundtable, University of Konstanz. Available online at: https://files.nyu.edu/rb964/public/correlates.pdf

Bateni, Mohammad Reza (1348 [1969–1970]). *Tosif-e sâxtemân-e dasturi-ye zabân-e fârsi [A Description of Persian Syntactic Structure]*. Tehran: Amir Kabir.

Bauer, Laurie (2003). *Introducing Linguistic Morphology* (2nd ed.). Washington DC: Georgetown University Press.

Baumann, Stefan, and Grice, Martine (2006). The Intonation of Accessibility. *Journal of Pragmatics*, *38*(10), 1636-1657.

Baumann, Stefan, and Hadelich Kerstin (2003). Accent Type and Givenness: An Experiment with Auditory and Visual Priming. In *Proceedings of 15th ICPhS.* Barcelona, 1811–1814.

Beckman, Mary E. (1986). *Stress and Non-stress Accent* (Vol. 7). Dordrecht: Foris Publications.

Beckman, Mary E., and Pierrehumbert, Janet B. (1986). Intonational Structure in Japanese and English. *Phonology Yearbook*, *3*. 255-309.

Berinstein, Ava E. (1979). A Cross-Linguistic Study on the Perception and Production of Stress. *Working Papers in Phonetics*, 47, 13–32.

Boersma, Paul, and Weenink, David (2013). Praat: Doing Phonetics by Computer [Computer program]. Version 5.3.53, retrieved 9 July 2013 from http://www.praat.org/

Bögel, Tina, and Butt, Miriam (2013). Possessive Clitics and Ezafe in Urdu. In Kersti Börjars, David Denison and Alan Scott (eds.), *Morpho-Syntactic Categories and the Expression of Possession* (pp. 291–322). Amsterdam: John Benjamins.

Bolinger, Dwight M. (1986). *Intonation and Its Parts: Melody in Spoken English*. Stanford: Stanford University Press.

Bolinger, Dwight M. (1989). *Intonation and Its Uses: Melody in Grammar and Discourse*. Stanford: Stanford University Press.

Booij, Geert (1995). *The Phonology of Dutch*. Oxford University Press. Oxford.

Booij, Geert (1996). Cliticization as Prosodic Integration: The Case of Dutch. *The Linguistic Review*, *13*, 219–42.

Brown, Gillian (1983). Prosodic Structure and the Given/New Distinction. In Ann Cutler and Robert D. Ladd (Eds.), *Prosody: Models and Measurements* (pp. 67–77). *Berlin: Springer*.

Browning, Marguerite A. and Karimi, Ezzat (1994). Scrambling to Object Position in Persian, in N. Corver and H. van Riemsdijk (Eds.), *Studies in Scrambling*, (pp. 61–

100). Berlin: Mouton de Gruyter.

Büring, Daneil (2001). Let's Phrase It! – Focus, Word Order, and Prosodic Phrasing in German Double Object Constructions. In Gereon Müller and Wolfgang Sternefeld (Eds.), *Competition in Syntax* (pp. 69–106). Berlin & New York: Mouton de Gruyter.

Butt, Miriam, Bögel, Tina, and Sulger, Sebastian (2008). Urdu Ezafe and the Morphology-Syntax Interface. In *Proceedings of LFG08*. Stanford: CSLI Publications, 129–149.

Caspers, Johanneke and van Heuven, Vincent J. (1993). Effects of Time Pressure on the Phonetic Realization of The Dutch Accent-lending Pitch Rise and Fall. *Phonetica*, *50*(3), 161-171.

Chodzko, Alexander (1852). *Grammaire Persane ou Principes de l'Iranien Moderne, Accompagnées de Facsimilés pour Servir de Modèle d'Écriture et de Style pour la Correspondence Diplomatique et Familière*. Paris: Maisonneuve & Cie.

Connell, Bruce (2001). Downdrift, Downstep, and Declination. Typology of African. Paper presented in Typology of African Prosodic Systems Workshop. Bielefeld University, Germany. May 18–20.

Cooper, André (1991). *An Articulatory Account of Aspiration in English*. (Doctoral dissertation, Yale University).

Cooper, André (1994). *Aspiration in English: How Should it be Defined and Where Should it be Described?*. Unpublished manuscript.

Cruttenden, Alan (2006). The Deaccenting of Given Information: a Cognitive Universal? In Giuliano Bernini and Marcia L. Schwartz (Eds.), *The Pragmatic Organization of Discourse in the Languages of Europe* (pp. 311–356). Berlin: Mouton de Gruyter.

Dabir-Moghaddam, Mohammad (1982). *Syntax and Semantics of Causative Constructions in Persian*. (Doctoral dissertation, University of Illinois, Urbana).

Dabir-Moghaddam, Mohammad (1997). Compound Verbs in Persian. *Studies in the Linguistic Sciences*, *27*(2), 25–59.

Dabir-Moghaddam, Mohammad (Dabirmoqaddam Mohammad) (1384 [2005–2006]). Pirâmnun-e 'ra' dar zabân-e fârsi [About '*ra*' in Persian Language]. In *Pajuheshâ-*

ye zabânšenâxti-ye fârsi majmmuʔe maqâlât [Studies in Persian Linguistics: Selected Articles] (pp. 81–148). Tehran: markaz-e našr-e dânešgâhi [Iran University Press].

de Lacy, Paul (1997). *Prosodic Categorisation* (MA Thesis, University of Auckland).

Diesing, Molly (1992). *Indefinites*. Cambridge, Massachusetts: MIT Press.

Downing, Laura (2006). *Canonical Forms in Prosodic Morphology*. Oxford: Oxford University Press.

Elfner, Emily (2009). The Interaction of Linearization and Prosody: Pronoun Postposing in Irish. Paper presented at Formal Approaches to Celtic Linguistics, University of Arizona. March 27.

Elfner, Emily (2012). *Syntax-Prosody Interactions in Irish* (Doctoral dissertation, University of Massachusetts, Amherst).

Eslami, Moharram (1379 [2000-2001]). *Šenâxt-e navâ-ye goftâr-e zabân-e fârsi va kârbord-e ân dar bâzsâzi va bâzšenâsi-ye râyâneʔi-ye goftâr [The Prosody of the Persian Language and Its Application in Computer-Aided Speech Recognition]*. (Doctoral dissertation, University of Tehran).

Eslami, Moharram (1382 [2003–2004]). Nahv va vâjšenâsi: yek sath-e moštarak [Syntax and Phonology: An Interface]. In Omid Tabibzadeh (Ed.), *Jašnnâme-ye doktor ali ašraf-e sâdeqi, pajuhešha-ye zabânšenasi-ye irâni (1) [Festschrift for Prof. Dr. A. Sadeghi, Iranian Linguistic Studies (1)]* (pp. 141–176). Tehran: Hermes Publications.

Eslami, Moharram, and Mahmoud Bijankhan (2002). Nezâm-e âhæng-e zæbân-e fârsi [Persian Intonation System]. *Iranian Journal of Linguistics*, *34*, 36-61.

Esposito, Christina M. and Patrick Barjam (2007). The Intonation of Questions in Farsi. *UCLA Working Papers in Phonetics*, *105*, 1–18.

Fagyal, Zsuzsanna, Kibbee, Douglas, and Jenkins, Frederic (2006). *French: A Linguistic Introduction*. Cambridge: Cambridge University Press.

Ferguson, Charles (1957). Word Stress in Persian. *Language*, *33*(2):123-135.

Féry, Caroline (2010). Recursion in Prosodic Structure. *Gengo Kenkyu* [*Phonological Studies]*, *13*, 51–60.

Féry, Caroline, and Ishihara, Shinichiro (2009). How Focus and Givenness Shape

Prosody . In Malte Zimmermann (Ed.), *Information Structure : Theoretical , Typological , and Experimental Perspectives* (1st Ed.), (pp. 36–63). Oxford: Oxford University Press.

Foadi, Hossein (1312 [1933-1934]).Âhang-e zabân-e fârsi [The Intonation of Persian Language]. *Mehr*, *1*, 964–968.

Fry, Dennis B. (1955). Duration and Intensity as Physical Correlates of Linguistic Stress. *The Journal of the Acoustical Society of America*, *27*, 765–769.

Fry, Dennis. B. (1958). Experiments in the Perception of Stress. *Language and Speech*, *1*, 126–152.

Gerlach, Birgit (2002). *Clitics between Syntax and Lexicon*. Amsterdam: John Benjamins.

Ghomeshi, Jila (1996). Non-Projecting Nouns and the Ezafe Construction in Persian. *Natural Language & Linguistic Theory*, *15*(4), 729–788.

Ghomeshi, Jila (2001). Control and Thematic Agreement, *Canadian Journal of Linguistics*, *46*(1/2), 9–40.

Greenberg, Joseph (1963). Some Universals of Grammar with Particular Reference to the order of Meaningful Elements. In: Joseph Greenberg (Ed.), *Universals of Language* (pp. 73–113). Cambridge & Massachusetts: MIT Press.

Gussenhoven, Carlos (1984) *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris Publications.

Gussenhoven, Carlos (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos (2005). Transcription of Dutch Intonation. In Sun-Ah Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, (pp. 118–145). Oxford: Oxford University Press.

Haghshenas, Alimohammad (Haqšenâs, Alimohammad) (1380 [2001–2002]). *Âvâšenâsi [Phonetics]* (7th ed.). Tehran: Entešârât-e âgâh [Agah Publications].

Hayati, Majid (1998). A Contrastive Analysis of English and Persian Intonation. *Papers and Studies in Contrastive Linguistics (Poznan, Poland)*, *34*, 53–72.

Hayes, Bruce (1982). Extrametricality and English Stress. *Linguistic Inquiry*, *13*(2), 227–276.

Hayes, Bruce (1991) *Metrical Stress Theory: Principles and Case Studies*. Unpublished manuscript, UCLA.

Hayes, Bruce (1995). *Metrical Stress Theory*. Chicago: University of Chicago Press.

Hayes, Bruce, and Lahiri, Aditi (1991). Bengali Intonational Phonology. *Natural Language & Linguistic Theory*, *9*(1), 47–96.

Hodge, Carleton T. (1957). Some Aspects of Persian Style. *Language*, *33*, 355–369.

Hosseini, Ayat (2012a). Acoustic Correlates of Stress in Persian. An unpublished talk given in 7th Phonology Festa at Biwako. Saga, Japan.

Hosseini, Ayat (2012b). The Prosodization of Function Words in Persian. *Orientalia Suecana, 61,* 126–137.

Hosseini, Ayat (2013a). L1 Interference in L2 Prosody: Contrastive Focus in Japanese and Persian. *Language and Information Sciences*, *11*, 55–67.

Hosseini, Ayat (2013b). *Phonetic Differences between Pre-nuclear and Nuclear Accents in Persian*. Poster presented at the 3rd International Conference on Phonetics and Phonology (3rd ICPP). National Institute for Japanese Language and Linguistics (NINJAL) December 20. Tokyo, Japan.

Hosseini, Ayat (2014a). The Ezafe Morpheme in Persian: An XP-external Clitic. *Phonological Studies*, *17*, to appear.

Hosseini, Ayat (2014b). Review of the Book *Prosody Matters: Essays in Honor of Elisabeth Selkirk*, ed. by: Toni Borowsky, Shigeto Kawahara, Takahito Shinya, Mariko Sugahara. *English Linguistics*, *31*, to appear.

Hyde, Brett (2001). *Metrical and Prosodic Structure in Optimality Theory* (Doctoral dissertation, Rutgers University).

Hyman, Larry M. (1977). On the Nature of Linguistic Stress. In Larry Hyman (Ed.), *Studies in Stress and Accent* (pp. 37–82). SCOPIL 4. Los Angeles: University of Southern California.

Hyman, Larry M. (2006). Word-prosodic Typology. *Phonology*, *23*, 225–257.

Hyman, Larry M. (2009). How (Not) to Do Phonological Typology: The Case of Pitch-accent. *Language Sciences*, *31*, 213–238.

IBM Corp. (2011). SPSS Statistics for Windows, [Computer software] Version 20.0 Chicago, IL: SPSS Inc.

Igarashi, Yosuke (2012). Prosodic Typology in Japanese Dialects from a Cross-linguistic Perspective. *Lingua*, *122*(13), 1441–1453.

Inkelas, Sharon (1989). *Prosodic Constituency in the Lexicon* (Doctoral dissertation). Stanford University.

Ishihara, Shinichiro (2003). *Intonation and Interface Conditions* (Doctoral dissertation, Massachusetts Institute of Technology).

Ishihara, Shinichiro (2007). Major Phrase, Focus Intonation, Multiple Spell-Out (MaP, FI, MSO). *The Linguistic Review*, *24*(2–3), 137–167.

Ito, Junko, and Mester, Armin (1992) [2003]. Weak Layering and Word Binarity. in Takeru Honma, Masao Okazaki, Toshiyuki Tabata and Shin-ichi Tanaka, (Eds.), *A New Century of Phonology and Phonological Theory. A Festschrift for Professor Shosuke Haraguchi on the Occasion of His Sixtieth Birthday* (pp. 26–65). Tokyo: Kaitakusha.

Ito, Junko, and Mester, Armin (1999). Realignment. In René Kager, Harry van der Hulst & Wim Zonneveld (Eds.), *The Prosody-morphology Interface* (pp. 188–217). Cambridge: Cambridge University Press.

Ito, Junko, and Mester, Armin (2006). Prosodic Adjunction in Japanese Compounds. In Yoichi Miyamoto & Masao Ochi (eds.), *The Proceedings of Formal Approaches to Japanese Linguistics* 4, (pp. 7–112). Cambridge, MA: MITWPL 55.

Ito, Junko, and Mester, Armin (2009). The Extended Prosodic Word. In Janet Grijzenhout and Kabak, Bariş (Eds.), *Phonological Domains: Universals and Deviations* (pp. 135–194). Berlin: Mouton de Gruyter.

Ito, Junko, and Mester, Armin (2012). Recursive Prosodic Phrasing in Japanese. In Toni Borowsky, Shigeto Kawahara, Takahito Shinya and Mariko Sugahara (Eds.), *Prosody Matters: Essays in Honor of Elisabeth Selkirk* (p. 280–303). London: Equinox Publishers.

Jahani, Carina (2005). The Glottal Plosive: A Phoneme in Spoken Modern Persian or Not. In Éva Ágnes Csató, Bo Isaksson and Carina Jahani (Eds.), *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*, (pp. 79–96). London and New York: Routledge.

Jazayeri, Mohammad-Ali, and Paper, Herbert H. (1961). *A Reference Grammar of*

Modern Persian. Ann Arbor: The University of Michigan Press.

Kabak, Bariş, and Reviathiadou, Anthi (2007). Prosodic Structure above the Phonological Word. Paper presented at the Workshop on Phonological Domains; Universals and Deviations, DGfS-29, Siegen University.

Kabak, Bariş, and Reviathiadou, Anthi (2009). An Interface Approach to Prosodic and Word Recursion. In Janet Grijzenhout and Bariş Kabak (Eds.), *Phonological Domains: Universals and Deviations* (pp. 105–134). Berlin: Mouton de Gruyter.

Kabak, Bariş, and Vogel, Irene (2001). The phonological Word and Stress Assignment in Turkish. *Phonology*, *18*, 315–360.

Kager, René (1993). Alternatives to the Iambic-trochaic Law. *Natural Language and Linguistic Theory*, *11*(2), 381–432.

Kager, René W. J. (1989). *A Metrical Theory of Stress and Destressing in English and Dutch*. Dordrecht: ICG.

Kager, René W. J. (2007). Feet and Metrical Stress. In Paul de Lacy (Ed.), *The Cambridge Handbook of Phonology* (pp. 195–227). Cambridge, UK: Cambridge University Press.

Kager, René W. J. (2010). The Typology of Stress Languages; with Special Reference to Stress Windows. A talk given in International Symposium on Accent and Tone. National Institute for Japanese Language and Linguistics. Tokyo.

Kahnemuyipour, Arsalan (2000a). On the Derivationality of Some Inflectional Affixes in Persian. Paper presented at the annual meeting of the Linguistic Society of America, Chicago.

Kahnemuyipour, Arsalan (2000b). Persian Ezafe Construction Revisited: Evidence for Modifier Phrase. Ottawa Working Papers in Linguistics. In *Proceedings of the Canadian Linguistic Association Conference*. Edmonton. Canada. 173–185.

Kahnemuyipour, Arsalan (2001). On WH-questions in Persian. *Canadian Journal of Linguistics*, *46*(1–2), 41–61.

Kahnemuyipour, Arsalan (2003). Syntactic Categories and Persian Stress. *Natural Language & Linguistic Theory*, *21*(2), 333–379.

Kahnemuyipour, Arsalan (2004). *The Syntax of Sentential Stress*. (Doctoral dissertation, University of Toronto).

Kahnemuyipour, Arsalan (2005). Towards a Phase-Based Theory of Sentential Stress. In Perspectives on Phases. In *Proceedings of the Workshop on EPP and* Phases (pp. 125–149). MIT Working Papers in Linguistics Series.

Kahnemuyipour, Arsalan (2006). Persian Ezafe Construction: Case, Agreement or Something Else. In *Proceedings of the 2nd Workshop on Persian Language and Computer* (pp. 3–16). Tehran University, Tehran, Iran.

Kahnemuyipour, Arsalan (2009). *The Syntax of Sentential Stress*. Oxford: Oxford University Press.

Karimi, Simin (1989). *Aspects of Persian Syntax, Specificity, and the Theory of Grammar*. (Doctoral dissertation, University of Washington).

Karimi, Simin (1996). Case and Specificity: Persian *ra* Revisited. *Linguistic Analysis*, *26*(3/4), 173–194.

Karimi, Simin (1997). Persian Complex Verbs: Idiomatic or Compositional. *Lexicology*, *3*(2), 273–318.

Karimi, Simin (2001). Persian Complex DPs: How Mysterious Are They? *Canadian Journal of Linguistics*, *46*, 63–96.

Karimi, Simin (2003). On Object Positions, Specificity and Scrambling in Persian. In Simin Karimi (Ed.), *Word Order and Scrambling* (pp. 91–124). Oxford & Berlin: Blackwell Publishers.

Karimi, Simin (2005). A *Minimalist Approach to Scrambling. Evidence from Persian*. Berlin & New York: Mouton de Gruyter.

Karimi-Doostan, Gholamhossein (1997). *Light Verb Constructions in Persian*. (Doctoral dissertation, University of Essex).

Khazayi-Far, Ali (1383 [2004-2005]). Tekiye-ye jomle dar jomle-ye sâde-ye fârsi [Sentential Stress in Persian Simple Sentences]. *Nâme-ye Farhangestân*, *6*(4), 30–47.

Klavans, Judith L. (1985). The Independence of Syntax and Phonology in Cliticization. *Language*, *61*, 95–120.

Koopman, Hilda, and Sportiche, Dominique (1991). The Position of Subjects, *Lingua*, *85*, 211–258.

Kornai, András, and Pullum, Geoffrey (1990). The X-Bar Theory of Phrase Structure.

Language, 66, 24–50.

Ladd, Robert D. (1986) Intonational Phrasing: The Case for Recursive Prosodic Structure. *Phonology* 3, 311–340.

Ladd, Robert D. (1992) *Compound Prosodic Domains*. Occasional Paper, Linguistics Department, University of Edinburgh.

Ladd, Robert D. (1996/2008). *Intonational Phonology*. Cambridge: Cambridge University Press.

Ladefoged, Peter (1996). *Elements of Acoustic Phonetics* 2nd ed. Chicago: University of Chicago Press.

Ladefoged, Peter (2006). *A Course in Phonetics* 5th ed. Boston: Thomson Wadsworth.

Laeufer, Christiane (1995). Morphology and Syllabification Domains. *Lingua*, *97*, 101–121.

Lahiri, Aditi, and van der Hulst, Harry (1988). On Foot Typology. In *Proceedings of the North Eastern Linguistics Society*, *18*(2), 286–299.

Lai, Yuwen. 2008. *Acoustic Realization and Perception of English Lexical Stress by Mandarin Learners*. (Doctoral dissertation, University of Kansas).

Lambton, Ann K. (1957). *Persian Grammar*. Cambridge: Cambridge University Press.

Landau, Sabine, and Everitt, Brian (2004). *A Handbook of Statistical Analyses Using SPSS* (Vol. 1). Boca Raton, FL: Chapman & Hall/CRC.

Larson, Richard, and Yamakido, Hiroko (2005). Ezafe and the Deep Position of Nominal Modifiers. Paper presented in Barcelona Workshop on Adjectives and Adverbs. Barcelona.

Lawrence. Raphael J., Borden, Gloria J., and Harris, Katherine S. (2011). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech* (6th Ed.). Baltimore, MD: Lippincott, Williams, & Wilkins.

Lazard, Gilbert (1992). *A Grammar of Contemporary Persian*. English translation. Costa Mesa, California: Mazda Publishers. (Translated from French by Shirley Lyons; first published in 1957 as Grammaire du Persan Contemporain. Paris: Klinksieck).

Liberman, Mark and Pierrehumbert, Janet. (1984). Intonational Invariance under Changes in Pitch Range and Length. In Mark Aronoff and Richard. J. Oehrle (Eds.),

Language Sound Structure (pp. 157–233). Cambridge & Mass: MIT Press.

Liberman, Mark. Y., and Prince, Alan (1977). On Stress and Linguistic Rhythm. *Linguistic Inquiry*, *8*(2), 249–336.

Liberman, Mark. Y., and Streeter, Lynn A. (1978). Use of Nonsense Syllable Mimicry in the Study of Prosodic Phenomena. *The Journal of the Acoustical Society of America*, *63*(1), 231–233.

Lovestrand, Joseph (2009). *Prosodically Motivated Focus in Hausa: an Optimality Theory Account*. Occasional Papers in Applied Linguistics.

Lyons, Christopher (1999). *Definiteness*. Cambridge: Cambridge University Press.

Mahajan, Anoop (1990). *The A/A-bar Distinction and Movement Theory*, unpublished (Doctoral dissertation, Massachusetts Institute of Technology).

Mahjani, Behzad (2003). *An Instrumental Study of Prosodic Features and Intonation in Modern Farsi (Persian)* (MS thesis).

Mahootian, Shahrzad (1993*). A Null Theory of Codeswitching*. (Doctoral dissertation, Northwestern University).

Mahootian, Shahrzad (1997). *Persian*. London: Routledge.

Mahootian, Shahrzad and Santorini, Beatrice (1996). Code Switching and the Complement/Adjunct Distinction. *Linguistic Inquiry*, *27*(3), 464–479.

McCarthy, John J. (1979). On Stress and Syllabification. *Linguistic Inquiry*, *10*(3), 443–66.

McCarthy, John J. (2003). OT Constraints are Categorical. *Phonology*, *20*, 75–138.

McCarthy, John J. and Alan Prince (1991) Prosodic Minimality. Lecture presented at University of Illinois Conference The Organization of Phonology.

McCarthy, John J. and Prince, Alan (1993a) *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Cambridge & Mass: MIT Press.

McCarthy, John J. and Prince, Alan (1993b) Generalized Alignment. In Geerl Booij and Jaap van Marie (Eds.), *Yearbook of Morphology*, (pp. 79–153). Dordrecht: Springer.

McCarthy, John J. and Prince, Alan (1994). The Emergence of the Unmarked: Optimality in Prosodic Morphology. In Mercè Gonzàlez (Ed.), *Proceedings of the North East Linguistics Society* 24 (pp. 333–379). GLSA, Amherst, MA.

Moinzadeh, Ahmad (2005). A New Approach to the Ezafe Phrase in Persian. *Cahiers*

Linguistiques d'Ottawa, *33*(1), 35–61.

Moinzadeh, Ahmad (2006). The Ezafe Phrase in Persian: How Complements are added to N°s and A°s. *Journal of Social Sciences and Humanities of Shiraz University*, *23*(1): 45–57.

Moorcroft, Regina (1995). *Clause-level Functional Categories in Germanic V2 Languages*. (Doctoral dissertation, University of Toronto).

Morton, John, and Jassem, Wiktor (1965). Acoustic Correlates of Stress. *Language and Speech*, *8*, 148–158.

Müller, Gereon (1999) Optionality in Optimality-Theoretic Syntax. *GLOT International*, *4*(5), 3–8.

Naderi, Navid, and van Oostendorp, Marc (2011). Reducing the Number of Farsi Epenthetic Consonants. In Agnes Korn, Geoffrey Haig, Simin Karimi, and Pollet Samvelian (Eds.), *Topics in Iranian Linguistics* (pp. 153–166). Wiesbaden: Reichert.

Nespor, Marina, and Vogel, Irene (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.

Nespor, Marina, and Vogel, Irene (1989). On Clashes and Lapses. *Phonology*, *6*(1), 69–116.

Nye, Gertrude E. (1955). *The Phonemes and Morphemes of Modern Persian: A Descriptive Study.* (Doctoral dissertation, University of Michigan).

Oboki, Anthony (2006). *Acoustic Correlates of Word Stress in American English*. (Doctoral dissertation, Massachusetts Institute of Technology).

Odden, David (1987). Kimatuumbi Phrasal Phonology. *Phonology Yearbook*, *4*: 13–26.

Pantcheva, Marina (2006a). Persian Preposition Classes. In Peter Svenonius (Ed.), *Nordlyd: Tromsø Working Papers in Linguistics* 33(1), (pp. 1–25). Tromsø.

Pantcheva, Marina (2006b). Categorial Status of Persian Class 2 Prepositions. In Anna Asbury, Jakub Dotlačil, Berit Gehrke and Rick Nouwen (Eds.), *The Syntax and Semantics of Spatial Prepositions*. Amsterdam: John Benjamins.

Parmoon, Yadollah (1385[2006-2007]). Yek algoritm-e aruzi barâ-ye tekiye-ye pišro-ye kalame dar fârsi-ye emruz [A Prosodic Algorithm for Progressive Lexical Stress in Modern Persian]. In *Proceedings of the 2nd Workshop on the Persian Language and Computer* (pp. 262–284), June 2006, Tehran.

Pierrehumbert, Janet B. (1980). *The Phonology and Phonetics of English Intonation* (Doctoral dissertation, Massachusetts Institute of Technology).

Pierrehumbert, Janet B. and Hirschberg, Julia (1990). The Meaning of Intonational Contours in the Interpretation of Discourse. In Philip R. Cohen, Jerry Morgan, Martha E. Pollack (Eds.), *Intentions in Communications*. Cambridge (pp. 271–311), Mass: MIT Press.

Pierrehumbert, Janet B., and Beckman, Marry E. (1988). *Japanese Tone Structure*. Cambridge & Mass: MIT Press.

Plag, Ingo, Kunter, Gero, and Schramm, Mareile (2011). Acoustic Correlates of Primary and Secondary Stress in North American English. *Journal of Phonetics*, *39*, 362–374.

Prieto, Pilar, Van Santen, Jan and Hirschberg, Julia (1995). Tonal Alignment Patterns in Spanish. *Journal of Phonetics*, *23*(4), 429-451.

Prince, Alan (1976). *'Applying' Stress*. Unpublished manuscript, University of Massachusetts, Amherst. Downloadable from: http://ruccs.rutgers.edu/~prince/hold/app_str.pdf

Prince, Alan (1980). A Metrical Theory of Estonian Quantity. *Linguistic Inquiry*, *11*(3), 511–562.

Prince, Alan (1985). Improving Tree Theory. *Berkeley Linguistic Society*, *11*, 471–490.

Prince, Alan, and Smolensky, Paul (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University & University of Colorado at Boulder. Published 2004, Oxford: Blackwell.

Sadat-Tehrani, Nima (2007). *The Intonational Grammar of Persian* (Doctoral dissertation, University of Manitoba).

Sadeghi, Ali Ashraf (Sâdeqi, Ali Ašraf) (1366 [1987–1988]). *Avaz šodan-e jâ-ye tekiye dar baʔi kalamât-e fârsi [Change in the Place of Stress in Some Persian Words]*. Majalle-ye zabânšenasi [Journal of Linguistics], The University of Tehran, *6* (1), 23–28.

Sadeghi, Vahid (1391 [2012-2013]). Dark-e šenidâri-ye tekiye-ye vâjegâni dar zabân-e fârsi [Auditory Perception of Lexical Stress in Persian Language]. *Faslnâme-ye Pajuheshâ-ye zabân va adabiyyât-e tatbiqi [Quarterly Journal of Comparative*

Researches in language and literature], *10*, 153–176.

Sadeghi, Vahid (2011). Acoustic Correlates of Lexical Stress in Persian. Paper presented in The International Congress on Phonetic Sciences XVII, Hong Kong, pp. 1738–1741.

Samareh, Yadollah (Samare, Yadollâh) (1378 [1999–2000]). *Âvâšenâsi-ye zabân-e fârsi [Phonetics of the Persian Language]*, 5th ed. Tehran: Markaz-e našr-e dânešgâhi.

Same'i, Hossein (1374 [1996-1997]). Tekye-ye feʔl dar zabân-e fârsi: yek barrasi-ye mojaddad [Verb Stress in Persian: A Re-examination]. *Nâme-ye Farhangestân*, *1*(4), 6–21.

Samiian, Vida (1983). *Structure of Phrasal Categories in Persian, an X-bar Analysis.* (Doctoral dissertation, University of California Los Angeles).

Samiian, Vida (1994). The Ezafe Construction: Some Implications for the Theory of X-bar Syntax, in Mehdi Marashi (Ed.), *Persian Studies in North America* (pp. 17–41) Bethesda: Iranbooks.

Samvelian, Pollet (2006). *When Morphology does Better than Syntax: The Ezafe Construction in Persian*. Unpublished manuscript, Université de Paris 3.

Samvelian, Pollet (2007). A (phrasal) Affix Analysis of the Persian Ezafe. *Journal of Linguistics, 43*, 605–645.

Samvelian, Pollet and Tseng, Jesse (2010). Persian Object Clitics and the Syntax-Morphology Interface. In *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar* (pp. 212–232). Université Paris Diderot, Paris 7, France.

Scarborough, Rebecca (2007). The Intonation of Focus in Farsi. *UCLA Working Papers in Phonetics*, *105*, 19–34.

Scott, Charles T. (1964). Syllable Structure in Tehran Persian. *Anthropological Linguistics*, *6*(1), 27–31.

Selkirk, Elisabeth O. (1978). On Prosodic Structure and Its Relation to Syntactic Structure. In Thorstein Fretheim (Ed.), *Nordic Prosody II: Papers from a Symposium* (pp. 111–140). Trondheim: TAPIR.

Selkirk, Elisabeth O. (1981). The Nature of Phonological Representation. In John Anderson, John Laver and Terry Myers (Eds.), *The Cognitive Representation of*

Speech (pp. 379–388). Dordrecht: North Holland Publishing Company.

Selkirk, Elisabeth O. (1984). *Phonology and Syntax: The Relation between Sound and Structure.* Cambridge: MIT Press.

Selkirk, Elisabeth O. (1986). On Derived Domains in Sentence Phonology. *Phonology Yearbook*, *3*, 371–405.

Selkirk, Elisabeth O. (1995). The Prosodic Structure of Function Words. In Jill Beckman, Laura Dickey and Suzanne Urbanczyk (Eds.), *Papers in Optimality Theory. University of Massachusetts Occasional Papers* 18 (pp. 439–469). Amherst & Mass.: GLSA.

Selkirk, Elisabeth O. (2000). The Interaction of Constraints on Prosodic Phrasing. In Merle Horne (Ed.), *Prosody: Theory and Experiment* (pp. 231–262), Dordrecht: Kluwer.

Selkirk, Elisabeth O. (2001). The Syntax-Phonology Interface. In N.J. Smelser and Paul B. Baltes (Eds.), *International Encyclopedia of the Social and Behavioral Sciences*, (pp. 15407–15412). Oxford: Pergamon.

Selkirk, Elisabeth O. (2005). Comments on Intonational Phrasing in English. In Frota, Sónia, Marina Vigário & Maria João Freitas (Eds.), *Prosodies*, (pp. 11–58). Berlin: Mouton de Gruyter.

Selkirk, Elisabeth O. (2009). On Clause and Intonational Phrase in Japanese: The Syntactic Grounding of Prosodic Constituent Structure. *Gengo Kenkyu [Phonological studies]*, *136,* 35–73.

Selkirk, Elisabeth O. (2011). The Syntax-Phonology Interface. In John Goldsmith, Jason Riggle, and Alan Yu (Eds.), *The Handbook of Phonological Theory* (2nd ed.) (pp. 435–484). London: Blackwell Publishing.

Selkirk, Elisabeth O. and Tong, Shen (1990). Prosodic Domains in Shanghai Chinese. In Sharon Inkelas and Draga Zec (Eds.), *The Syntax/Phonology Connection* (pp. 313–338). Chicago: University of Chicago Press.

Selkirk, Elisabeth O., and Tateishi, Koichi (1988). Minor Phrase Formation in Japanese. *CLS*, *24*, 316-336.

Selkirk, Elisabeth O., and Tateishi, Koichi (1991). Syntax and Downstep in Japanese. in Carol P. Georgopoulos and Roberta L. Ishihara (Eds.), *Interdisciplinary Approaches*

to Language: Essays in Honor of Shige-Yuki Kuroda, (pp. 519-544). Dordrecht: Kluwer.

Sepanta, Sasan (1356 [1977-1978]). Barrasihâ-ye jadid dar mowred-e tekye-ye zabân-e fârsi [New Studies of Word Stress in Persian]. *Našriye-ye dâneškade-ye olum-e ensâni-ye dânešgâh-e tehrân [Journal of Faculty of Humanities, The University of Tehran]*, *7*(2), 301–308.

Sepanta, Sasan (1374[1995–1996]). Âhang-e goftâr va musiqi [Speech Intonation and Music]. In *Proceedings of the Third Tehran Linguistics Conference* (pp. 39–53), Tehran, Iran.

Silverman, Kim and Pierrehumbert, Janet B. (1990). The Timing of Prenuclear High Accents in English. In John Kingston and Marry Beckman (Eds.), *Papers in Laboratory Phonology I* (pp. 71-106). Cambridge: Cambridge University Press.

Sluijter, Agatha M. C. (1995). *Phonetic Correlates of Stress and Accent* (Vol. 15). The Hague: Holland Academic Graphics.

Sluijter, Agatha M. C. and van Heuven, V. J. (1996a). Acoustic Correlates of Linguistic Stress and Accent in Dutch and American English. In *Proceedings of the 4th International Conference on Spoken Language Processing 3* (pp. 630–633) Philadelphia.

Sluijter, Agatha M. C. and van Heuven, Vincent J. (1996b). Spectral Balance as an Acoustic Correlate of Linguistic Stress. *Journal of the Acoustiacal Society of America*. *100*(4), 2471–2485.

Sluijter, Agatha M. C., Shattuck-Hufnagel, S., Stevens, K. N., and van Heuven, Vincent J. (1996). Supralaryngeal Resonance and Glottal Pulse Shape as Correlate of Stress and Accent in English, *in Proceedings of ICPhS 95* (pp. 630–633). Department of Speech Communication and Music Acoustics, KTH and Department of Linguistics, Stockholm University, Stockholm.

Sluijter, Agatha M. C., van Heuven, Vincent J., and Pacilly Jos J. A. (1997). Spectral Balance as a Cue in the Perception of Linguistic Stress. *Journal of the Acoustical Society of America*, *101*(1), 203–513.

Solhju, Ali, Tabatabayi, Ala'eddin, Tabibzadeh, Omid, Shojayi, Mohsen, Eslami, Moharram, and Naqzguy-Kohan, Mehrdad (interviewer) and Sadeghi, Ali Ashraf

(interviewee) (1382 [2003–2004]). Gofto-gu bâ doktor ali ašraf sâdeqi [An interview with Professor Ali Ashraf Sadeghi]. In Omid Tabibzadeh (ed.) *Jašnnâme-ye doktor ali ašraf-e sâdeqi, pajuheshâ-ye zabânšenasi-ye irâni (1) [Festschrift for Prof. Dr. A. Sadeghi, Iranian Linguistic Studies (1)]* (pp. 3–140). Tehran: Hermes Publications.

Steedman, Mark (1991). Structure and Intonation. *Language*, *67*(2), 260–296.

Taheri-Ardali, Mortaza, and Xu, Yi (2012). Phonetic Realization of Prosodic Focus in Persian. In *Proceedings of Speech Prosody* 2012, Shanghai, 326–329.

Towhidi, Jalil(1974). *Studies in the Phonetics of Modern Persian. Intonation and Related Features* (=Forum Phoneticum 2). Hamburg: Helmut Buske [originally dissertation, London 1973].

Truckenbrodt, Hubert (1995) *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. (Doctoral dissertation, Massachusetts Institute of Technology).

Truckenbrodt, Hubert (1999). On the Relation between Syntactic Phrases and Phonological Phrases. *Linguistic Inquiry*, *30*(2), 219–255.

Uhmann, Susanne (1991) *Fokusphonologie*. Tübingen: Niemeyer.

Vahedi-Langrudi, Mohammad-Mehdi (1996). *The Syntax, Semantics and Argument Structure of Complex Predicates in Modern Farsi*. (Doctoral dissertation, University of Ottawa).

Vahidian-Kamyar, Taghi (1379 [2000–2001]). *Navâ-ye goftâr dar fârsi [Speech Prosody in Persian]*. Mashhad: Ferdowsi University Press.

van Der Hulst, Harry (2000). Issues in Foot Typology. In Mike Davenport and S.J. Hannahs (Eds.), *Issues in Phonological Structure* (pp. 95–127). Amsterdam: John Benjamins.

van Katwijk, Albert (1974). *Accentuation in Dutch: an Experimental Linguistic Study* (Doctoral dissertation, University of Utrecht).

Venditti, Jennifer. V., Maekawa, Kikuo, and Beckman, Marry E. (2008). Prominence Marking In the Japanese Intonation System. In Shigeru Miyagawa and Mamoru Saito (Eds.), *Handbook of Japanese Linguistics* (p: 456–512). Oxford: Oxford University Press.

Versteegh, Kees (2001). Linguistic Contacts between Arabic and Other Languages.

Arabica, *48*(4), 470–508.

Wagner, Michael (2005). *Prosody and Recursion* (Doctoral dissertation, Massachusetts Institute of Technology).

Walker, Rachel (1997). *Mongolian Stress, Licensing, and Factorial Typology.* Unpublished Manuscript, University of California at Santa Cruz. Downloadable from Rutgers Optimality Archive, http://roa.rutgers.edu, ROA: 172.

Wang, Qian (2008). *Perception of English Stress by Mandarin Chinese Learners of English: An Acoustic Study*. (Doctoral dissertation, University of Victoria).

Watson, Janet C. E. (2011). Word Stress in Arabic. In Marc Oostendorp, Colin Ewen, Elizabeth Hume, Keren Rice (Eds.), *The Blackwell Companion to Phonology 5* (pp. 2990–3018). Oxford: Wiley-Blackwell.

Werle, Adam (2009). *Word, Phrase, and Clitic Prosody in Bosnian, Serbian, and Croatian*. (Doctoral dissertation, University of Massachusetts, Amherst).

Wheeler, Deirdre (1979). *A Metrical Analysis of Stress and Related Processes in Southern Paiute and Tübatulabal*. University of Massachusetts Occasional Papers in Linguistics, Vol. 5, 145–175.

Wicka Somers, Katerina (2009). *From Phonology to Syntax: Pronominal Cliticization in Otfrid's Evangelienbuch* (Vol. 530). Tübingen: Max Niemeyer.

Wiese, Richard (1996). *The Phonology of German*. New York: Oxford University Press.

Windfuhr, Gernot L. (1979). *Persian Grammar: History and State of Its Study* (Vol. 12). Walter de Gruyter.

Windfuhr, Gernot L. (1987). Persian. In Bernard Comrie (Ed.), *The World's Major Languages* (523–546). New York: Oxford University Press.

Windfuhr, Gernot L., and Perry, John R. (2009). Persian and Tajik. In: Gernot Windfuhr (Ed.), *The Iranian Languages* (pp. 416–544). London and New York: Routledge.

Yarmohammadi, Lotfollah (1964). *A Contrastive Study of Modern English and Modern Persian*. (Doctoral dissertation, Indiana University).

Zec, Draga (2005). Prosodic Differences among Function Words. *Phonology*, *22*, 77–112.

Zepter, Alexandra (2003). *Phrase Structure Directionality: Having a Few Choices* (Doctoral dissertation, Rutgers, The State University of New Jersey).

Zimmermann, Malte (2006). Focus in Western Chadic: A Unified OT-account. In *Proceedings of NELS 36*. UMASS, Amherst.

Zubizarreta, Maria L. (1998). *Prosody, Focus and Word Order.* Cambridge, MA: The MIT Press.