An Instrumental Study of Prosodic Features and Intonation in Modern Farsi (Persian)

Behzad Mahjani

Supervisor: Robert Ladd

Master of Science

in

Speech and Language Processing

Department of Theoretical and Applied Linguistics

School of Philosophy, Psychology, Linguistics and Social Sciences

University of Edinburgh

September 2003

Abstract

This work was an attempt in studying Farsi prosodic and intonation system within the framework of Autosegmental Metrical Phonology. Throughout the dissertation I tried to define the set of basic intonation units in Farsi, and further analyzed their distributions throughout Farsi prosodic phrases. I also attempted to describe various intonation patterns of Farsi utterances in terms of the relevant prosodic features and properties such as intonational contour, prominence, intonational phrasing, and boundary tones.

The study is within Pierrehumbert approach (1980), i.e. a model of intonation originally developed by her and used in the investigation of English intonation. The same two level primitive tones (namely high and low) were used for phonological representation of Farsi utterances. These level tones are then employed into tonal events (pitch accents and boundary tones), which occur within utterances. It is the behaviors of these tonal events throughout various utterances, which were the focus of this work. In fact such behaviors contribute toward distinguishing between various sentences types, or within the same sentence type determine nuisances of meaning and points of emphasis.

For this study complete texts as well as individual sample sentences were selected and read by four native speakers of Farsi. They include all types of sentences namely declaratives, interrogatives, imperatives and exclamatory sentences. Analysis of data revealed that in Farsi the prosodic structure consists of three prosodic units, namely Intonation Phrases, Intermediate Phrases, and Accentual Phrases. It was also found that the default pattern for Farsi pitch accents in all sentence types was L+H*. So in Farsi sentence types are mostly distinguished by their boundary tones.

Further in the dissertation, each sentence type was investigated for its particular pattern and arrangements of its tonal events. Among them declaratives and interrogatives were paid special attention, as they appeared to have the most basic intonational patterns. Considering the space of an MSc dissertation, some preliminary investigations were also carried out on focus within declaratives and interrogatives.

Declaration

I declare that this MSc. dissertation was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text. This work has not been submitted for any other degree or professional qualification except as specified.

(Behzad Mahjani)

Acknowledgments

To Homa Poorbarat

Contents

List of Figures

Introduction

Speaking, and reading, quite inevitably is accompanied by a musical tone. This melody, its rhythm, and its intensity constitute intonation and gives way to emergence of prosodic features. For a long time, linguists have mostly relegated prosody and intonation to the study of emotions and social attitudes, and investigated them in very limited domains e.g. indication of the declarative or interrogative modality of the sentence. Although instrumental analyses have long been introduced in the study of phonetics and phonology of various languages, due to this overlook, experimental investigations of prosody and intonation were not so common. Only more recently such experimental research has been taken more seriously. As part these instrumental investigations, the study of prosody and intonation of both tone and stress (intonation languages) has been carried out through the careful measurement and precise investigation of pitch, which is mainly realized in the fundamental frequency of the vocal folds (also known as F0). During the last two decades of the 20th century, such experimental investigations have been made for many languages both European and Non-European ones. This work is an attempt in providing a preliminary investigation on Farsi prosodic and intonational characteristics mostly through investigation of pitch realization within the framework of Autosegmental Metrical theory. In the absence of apparently any prior serious experimental studies of Farsi prosodic and intonation system, this project can be helpful and practical. It is hoped that this work can provide at least a partial foundation for further more detailed research in this field.

Aims of this Survey

The aim of this dissertation is a description of Farsi prosodic and intonation system within the framework of AM (Autosegmental Metrical Phonology). It is hoped that such investigation will contribute at least partially to the description of Farsi prosodic features and intonational lexicon. The study is based on two kinds of recorded data, firstly

sentences read in isolation and secondly sentences read within a larger pragmatic context (two short narrations). As this preliminary work can be considered one of the very first surveys on the Farsi prosodic and intonation system, the results can offer a foundation for more serious future studies. Throughout this work the author tries firstly to define the set of basic intonation units for Farsi, and analyze their distributions so as to obtain the Farsi prosodic groupings. Secondly he will attempt to describe to a certain degree the possible meanings of prosodic contours within the prosodic and intonation system of Farsi language. However as a preliminary work, this survey mostly concerns overall formal structure of intonational patterns, and leaves aside for now very detailed study of specific tune meanings. Something, which the author feels obliged to mention here, is that this dissertation by no means presents an exustive and comprehensive survey on Farsi intonation and prosodic features. In a short time Msc. project like this many other important aspects of Farsi prosodic and intonational features could not be included; among them one can name a few; such as the role of syntax in intonation groupings and nucleus placement in Farsi, or a detailed investigation of the semantics and pragmatics of the set of tones in this language. So it is obvious that investigation of such phenomena will be left for future research.

Importance of this Survey

The importance of such study can be multifold. From an engineering point of view investigation of Farsi prosodic structures and intonation patterns can be useful both in speech synthesis and in speech recognition. Such studies can offer a solid platform for developing such systems. Using prosodic features enables us to increase both the intelligibility and naturalness of our systems within speech synthesis domain and to raise the efficacy of recognition in our SR systems[1]. On the other hand as a descriptive work; this survey can be quite valuable for native Farsi phonologists and those linguists who are interested in knowing facts about other languages. Also from a pure theoretical point of view, such studies can be of great value. Extracting universal and language-specific parameters across the dimensions of prosody and intonation of a specific language and

---

[1] In recent years, there have been some attempt to design and implement prosody dependent SR systems. For a hesitant reader, a simple search in google will provide interesting projects and essays.

its theoretical implications can be used for verification of some widely assumed theories of prosody and intonation. Such studies contribute to our understanding of various interactions, which exist between prosodic structures at one side and other linguistic/paralinguistic elements at the other (such as syntactic constituents, semantic and discourse related particles, pragmatic interpretation, etc.). Thus languages of various prosodic types, with different intonation patterns may widen our horizon to a better understanding of such phenomena and eventually lead us to devising more efficient theories.

Structure of Dissertation

This work consists of 4 chapters, an introduction and a conclusion. Chapter 1 provides an introduction to Farsi language. There will be a brief look at Farsi grammar, phoneme inventory, lexical stress rules, as well as a short review of the related literature. Chapter 2 is devoted particularly to a review on the theory of intonational representation, covering firstly some general aspects of prosody and intonation and secondly the phonological theory of prosody and intonation. Here a model of intonation representation (the one which originally was proposed by Pierrehumbert) has been explained. It is on the basis of this model that in later chapters various intonational patterns and prosodic features such as pitch accents, boundary tones, etc are represented. Chapter 3 provides an overview of the principal intonational tunes of Farsi and its common tonal events. Here the prosodic constituent structure of Farsi is described, and there is a discussion about the corpus of texts and the speakers chosen for preparation of the recorded data. Chapter 4 is the principal experimental chapter of this work. In this chapter there is a precise analysis of various Farsi sentence types[1] and the way they are tonally marked (in terms of their boundary tone types and probably relevant pitch accent patterns.). Finally Conclusion run through the whole discussions in brief and concludes the dissertation.

---

[1] Sentence types are namely simple declaratives, interrogatives including both wh-questions and yes/no questions, imperatives, and some typical exclamatory sentences.

Chapter 1

A Quick Look at Farsi Language

Farsi (also known as Persian) is a widely spoken language. It is the official language of Iran. Also it is spoken in countries like Afghanistan, Tajikistan, parts of Uzbekistan and Pakistan as well as the Pamir Mountain region. In addition there are other minority groups of native speakers in many other places of the world including Europe and North America. There is an estimation of over 35 million Farsi native speakers in the world[1]. (Persian Profile)

Modern Farsi had developed by the 9th century. The Early Modern period of the language (ninth to thirteenth centuries), preserved in the literature of the Empire, is known as Classical Persian, due to the eminence and distinction of poets such as Rudaki, Firdowsi, and Khayyam. During this period, Farsi was adopted as the lingua franca of the eastern Islamic nations. Even until recent centuries it was culturally and historically one of the most prominent languages of the Middle East and regions beyond. For example, it was an important language during the reign of the Moguls in India where knowledge of Persian was cultivated and encouraged. This led to compilation of numerous annals, chronicles, and court volumes of poetry outside Iran; the use of Farsi in the courts of Mogul India ended in 1837, banned by British officials of the East Indian Company. Persian scholars were prominent in both Turkish and Indian courts during the fifteenth to eighteenth centuries in composing dictionaries and grammatical works. In fact a Persian Indian vernacular developed and many colonial British officers learned their Persian from Indian scribes.

---

[1] In Iran the language is generally referred to as Farsi, but in Afghanistan as Dari.

Classical Persian remained essentially unchanged until the nineteenth century, when the dialect of Tehran rose in prominence, having been chosen as the capital of Persia by the Qajar dynasty in 1787. This Modern Persian dialect became the basis of what is now called Contemporary Standard Persian. Although it still contains a large number of Arab[1] terms, most borrowings have been nativized, with a much lower percentage of Arabic words in colloquial forms of the language.

1.1 Farsi Grammar

Farsi is classified as an SOV language, i.e. sentences are made in the order Subject-Object-Verb. Modifiers follow the nouns they modify and the language has prepositions. However in normal daily conversations, it appears as a highly free word order language, i.e. the sentential constituents can be moved around in the clause, especially the adverbial and prepositional phrases. Also the subject, if not completely omitted (because Farsi is a pro-drop language), can move around almost anywhere in the sentence. Such deviations from the canonical word order are often used for focused or topicalized readings. But we should take into consideration that in formal language style especially in written form, although most elements may appear in relatively free word order, the sentences often remain verb-final. Adverbs and prepositional phrases, however, can appear in various positions quite freely. The language relies on an affixal system that makes use of both prefixes and suffixes. However, much of the complex nominal and verbal inflection of Old Persian has been lost in modern Persian, including the inflectional distinction of case, and gender. However human and non-human types are distinguishable within the pronominal system. Also person and number distinctions are maintained, and specified objects of transitive verbs are marked by a marker. Farsi has no articles. An unmarked noun refers to a class of objects rather than a single thing. For example, the phrase "*man ketâb mixânam.*" (I book read.) means "I read books." i.e. the singular unmarked noun *"ketâb"* represents the whole class of "books" in general rather than any specific known book. The suffix "-i" is added as an indefinite marker, e.g. "*ketâb-i*" refers to "a book", so the phrase "*man ketâb-i mixânam.*" means "I read a

---

[1] A large number of Arabic words were added to Farsi vocabulary as the result of the conquest of the Persians by the Muslim Arabs in the 7th cent. A.D. In fact, many writers of Classical Persian used Arabic terms freely either for literary effect or to display erudition.

book", but the book here is indefinite, i.e. the listener doesn't know which book I read. In Farsi there is no equivalent of the English definite marker "the". Nouns are marked for specificity. There is one marker in the singular and two in the plural. The two markers of specific plurality in Farsi are either the Farsi plural markers "-an" and "-ha" or the Arabic "broken plural" forms. However, the Arabic broken plural may only be applied to Arabic loan words, and is not productive in Persian, i.e. it cannot be added to newly formed Persian words. On the contrary Arabic words can also change into plural form using Farsi plural suffixes (so they can have two different plural forms.). In formal language, the Farsi plural marker "-an" is used for humans and "-ha" for inanimate objects and animals. But recently, "-ha" is used indiscriminately. As mentioned above most of the time nominal modifiers follow the nouns they modify (though demonstrative adjectives and numerals precede nouns). Head nouns are connected to the modifiers that follow them by what is called ezafe[1], represented by "-e" sometimes realized as "-y" (phonetically determined).

There are two ways of indicating possession in Farsi. The clitic "-e" or the enclitic pronoun "-am" may be used to mark possessed nouns. For example, "*ketâb*" is "book"; "*ketâb-e man*" and "*ketâb-am*" both mean "my book".

The suffix "*-râ*" is used to mark specific direct objects in Persian. On the other hand indirect objects and adverbial phrases are marked by prepositions.

The Persian interrogative pronouns are "*kî*", and "*če*", which function like "who" and "what".

The particle "*ke*" is used to introduce relative clauses. It functions like both "who" and "which" in English. The clitic "*-i*" is added to a noun modified by a restrictive relative clause. The head noun may be represented pronominally within the relative clause. For example:

---

[1] Ezafe is an unstressed vowel "-e" ("-ye" after vowels) which appears on various positions in Farsi such as (i) a noun before another noun (attributive), (ii) a noun before an adjective, (iii) a noun before a possessor (noun or pronoun), (iv) an adjective before another adjective, etc.

- mard-i   ke   ?az   ?u   gereft-àm-aš (The man from whom I got it.)

man- Indef that from him   got-1SG-it.

Verbs are formed using one of two basic stems, present and past; aspect is as important as tense: all verbs are marked as perfective and imperfective. The latter is marked by means of prefixation. Both perfective and imperfective verb forms appear in three tenses: present, past and inferential past. While all forms may be used in a future context, future is not marked. Verbs usually agree with the subject in person and number. Something important about Farsi verbs is that they are normally compounds, i.e. consisting of a noun and a verb (light verbs). The nominal often can be an Arabic loan word, For example, the verb "*montazer budan/šodan*", meaning "to wait", literally means "to be/become waiting". The verbs "*budan*", and "*šodan*" ("to be" or "to become" respectively) are particularly productive function verbs in this respect. Finally negation of sentences is brought about by adding the prefix "*na-*" to the main verb. (World languages knowledge fair, Persian) (Karine Megerdoomian 1999)

1.2 Farsi Phonemes

The sound system of standard Farsi is quite symmetric. The phonemic system consists of 29 phonemes. There are six vowel sounds in Farsi, 3 long vowels and 3 short vowels. The three long vowels are /i/, /u/, and /â/; the three short vowels are /a/, /e/, and /o/. There are also two diphthongs: /ei/ and /ou/.

Tongue Height	Part of Tongue		
	Front	Center	Back
High	i		u
Mid	e		o
Low	a		â

Table 1-2: Farsi Vowels

Consonants consists of 23 phonemes among which the velar fricatives /x/ and /q/ do not have any counterparts in English[1].

		Bilabial	Labiodental	Alveolar	Palatal	Velar	Uvular	Glottal
Stops	Voiceless	p		t		k		?
	Voiced	b		d		g	q	
Fricatives	Voiceless		f	s	š	x		h
	Voiced		v	z	ž			
Affricates	Voiceless				č			
	Voiced				j			
Nasals		m		n				
Liquids				r, l				
Glides					y			

Table 1-1: Farsi Consonants

[1] Here I am not using IPA symbols precisely; rather I use symbols, which are frequently used by Farsi linguists for transliteration. Appendix A represents some of the IPA equivalents. Also appendix B presents an IPA presentation of Farsi consonants. It is a little bit different from the one I have used here (but roughly they are the same.).

1.3   A Short Review of Farsi Lexical Stress

Farsi syllables always take one of these patterns, CV, CVC, or CVCC. So in Farsi the syllable structure always contains an onset, i.e. it is not like English in that we can have a syllable, with only a rhyme, containing only a nucleus and a coda. On the other hand as the phonological restrictions in Farsi do not allow the occurrence of two vowels in the same syllable, syllabification is a trivial task in Farsi. The number of syllables almost always is equal to the number of vowels. (Samareh 1986)

A closer look at Farsi lexical stress system reveals that it is a weight-insensitive language (Windfuhr 1990) that is it assigns stress to a fixed syllable, e.g. the final syllable, or the stem, regardless of the internal makeup of syllables in the word. It is quite unlike the weight-sensitive stress systems in which there are certain syllable types that tend to attract stress based on their relatively greater weight (like Latin).
Many scholars accept the general rule, that in Farsi, word stress is progressive; therefore, it falls on the final syllable of a word (Windfuhr, World's Major Languages). However the apparent diversity of lexical stress patterns in Farsi has led linguists to suggest a split between some major lexical categories. One category consists of nouns and adjectives in which almost always the main stress goes on the final syllable of the word. The other category contains verbs for which the stress pattern is not as clear. Whereas in nouns and adjectives, the main stress is on the final syllable, in verbs there are a variety of possibilities for stress occurrence. Sometimes main stress occurs on the penultimate syllable, sometimes on the initial (or antepenultimate) syllable. As a result of such superficial differences in this category, scholars have proposed different stress rules for nouns and adjectives on the one hand and verbs on the other. (Arsalan Kahnemuyipour 2001)

Chodzko (1852) was the first one who identified the basic rule of final syllable stress pattern in Farsi. He recognizes that stress is word final in simple, derived, and compound nouns and adjectives. For verbal stress, he proposes different rules for different tenses. Other scholars such as Ferguson (1957), and Lazard (1992) too, distinguished verbal stress from the other categories. They considered non-verbal words having the stress on the last syllable and the verbs having "recessive stress". Therefore, verbs tend to be stressed on the initial syllable. For example, the compound noun "*bâz-*

kon", which means "opener", is stressed on the final syllable, while the verb phrase represented by " *bâz kon*", which means "open" is stressed on the initial syllable. Recently Mahootian (1997) asserts that stress is word-final in simple nouns, derived nouns, compound nouns, simple adjectives, derived adjectives, infinitives, and the comparative and superlative forms of adjectives as well as in nouns with plural suffixes, and mentions verbal stress as one of the exceptions to this rule. For the latter category Amini (1997) proposes two different word-layer construction rules, End Rule Left and End Rule Right, which are sensitive to lexical categories. She uses the first rule for prefixed verbs and the second one for all other categories. (Arsalan Kahnemuyipour 2001)

Some scholars by focusing on morphological processes carried on verb roots and the way they receive derivational and inflectional affixes, claim that depending on the type of morpheme attaching to roots, the stress of the verbs can shift. This shift of stress obeys numerous rules and principles, which are mostly referred to in a merely descriptive way in many grammar books written for Farsi. Family (2001) claims that there are some hints that show such principles at least partially are motivated by constraints in the grammar of the language, that is positions of stress is somehow in accordance with certain morphological structures, specifically 'heads ', and that Farsi phonology is not blind to morphology. Because of such complexities in Farsi many books mention relatively numerous rules, which determine the exact position of stress in words. In General we can claim that in Farsi the stress pattern of words is similar to that of Turkish that is except in some exceptional cases we always stress the final syllable of the word, regardless of how long it is. There is however some other stress patterns in Farsi i.e. having the stress on the first or penultimate syllable of the word. However such deviations from the dominant stress pattern (word finally stressed) is mostly controlled through the grammar of the language rather than phonologically enforced. So in most Farsi words as mentioned above the last syllable is stressed. Names, most adverbs, adjectives, most pronouns, main numbers, and the first part in multi-word combinational verbs are in this group. But in some Farsi words the first syllable is stressed (this group can mostly be defined grammatically.) for example all negative verbs, imperative verbs, simple present verbs and some adverbs and some pronouns are stressed on their initial syllables. Yet in some Farsi words like simple past verbs, the penultimate syllable is stressed (however within phrases and sentences, some words are never stressed such

as most single-word prepositions and conjunctions.). To make this discussion more concrete there follows some examples:

● In simple nouns or adjectives, stress is mostly on last syllable:

"pedár" (father, Noun)
"bozórg" (large, Adj)

● Certain polysyllabic conjunctions or adverbs carry the stress on the first syllable:

"bálke" (rather)
"máhâzâ" (nonetheless)
"xéyli" (very)

● In compound nouns, the main stress regularly falls on the final syllable of the last word:

"?âtaš -parást" (fire-worshipper)
"tâze-vâréd" (new-comer)

● In personal verb forms the stress is on non-final syllable. If a verb form is non-negative, the stress is placed on the last syllable of the stem of the verb form (It should not begin with inflectional prefixes such as "mi" (prefix showing progressive mood) or "be" (prefix showing subjunctive)).

"geréftam" (I took, geréft is the stem)

- If it begins with prefixes mentioned above then the prefix would be stressed:

"mígereftam" (I was taking, geréft is the stem)
"béporsand" (they may ask, pors is the stem)

- If a verb form is negative, then the stress is placed on the negative morpheme.

"nágereftam" (I did not take, na is the negation prefix)

"némigereftam" (I was not taking, ne is the negation prefix)

While many scholars have pointed out such superficial diversities in Farsi lexical stress system, their surveys are mostly descriptive in nature in the sense that they have only tried to classify stress behavior of different subgroups especially within verbal category without any systematic justification (similar to above mentioned examples). Kahnemuyipour (2001) on the other hand has tried to investigate closely the surface inconsistency between the uniformity in stress placement in nouns and the variability in verbs in his paper "Unifying Categories: Persian Stress Revisited". Through his arguments, he has tried to provide a unified (i.e. independent of lexical categories) account of Farsi stress, showing that by differentiating word and phrase level stress rules, one can account for the superficial differences in Farsi lexical stress patterns. He has done his survey within the framework of Prosodic Phonology (Selkirk 1980, 1981, 1984, 1986; Nespor & Vogel 1982, 1986), in which prosodic domains are defined based on syntactic structures and different phonological rules apply within these domains. His preliminary investigations (though focusing on syntactic phrases) show that in Farsi the prosodic structures are apparently very important in placement of stress at word level. (That is due to the fact that prosodic structures usually tend to map precisely into syntactic phrases). That is in Farsi, phrasal stress rules tend to monitor at least partially the behaviors of stress at the level of words (lexicon), therefore this leads us to an account of Persian stress, which is not merely dependent on lexical categories. Such findings reinforce drastically the importance of serious and precise investigations of Farsi prosodic and intonation system.

I will finish this part by just mentioning another function of lexical stress in Farsi. In this language like English sometimes stress can have a contrastive role as well, that is two phonetically similar words can contrast through stress. It can contribute to distinguishing grammatical functions like "mard-í" (manliness, man-Nom) and "márd-i" (a man, man-Indef), or "dárgozašt" (died, simple past tense verb) and "dargozášt" (death, Noun), or it can lead to totally different words, like "váli" (but), and "valí" (guardian/parent).

1.4   Previous Works on the Farsi Prosodic and Intonation system

In its traditional sense, prosody is considered as the study of the grammar of rhythm and meter in poetry. It deals with those patterns, which have been developed in specific literature traditions in order to enable (and constrain) poetic composition. This involves identification of those conventionalized patterns as meters i.e. the perception of beats into regular patterns. Having a rich history of classical poetry and verse, Persian writers started writing about prosody and meter quite early back in history, among them one can name people like Moulana Yusef of Nishapur (10th century), Bahrami of Sarakhs author of the khojaste-name and Bozarjmehr Qommi, Shamsoddin Mohammadebn-e Qeis Razi (early 13th Century) and Khajeh Nasiroddine Tusi (13th Century). Besides these native scholars, numerous surveys on versification system and meters of Persian classical poetry can be found among European works. Such works mostly reproduce the theories of the oriental scholars without significant change; among the famous scholars one can name Gladwin (18th Century), Rückert and Garcin de Tassy (early 19th century), and Blochmann (late 19th century). (Elwell_Sutton 1976)

Apart from such classical works, and as far as the modern concept of prosody is concerned, there has been very limited research on the study of intonation and prosodic features in Farsi. Most of the materials in this respect only cover some basic study of Farsi lexical stress patterns as well as some general aspects of Farsi intonation. One of the very first people who did instrumental experiments for investigation of Farsi phonemes and lexical stress patterns was P. N. Khanlari. The same scholar introduced some shallow concepts of intonation and suprasegmental features to the study of Farsi language.

In addition to this common trend of surveys, one can only find a few works, which have touched to a certain extent on the role intonation plays in Farsi, focusing briefly on various meanings each major intonation pattern (e.g. various reading of a single statement) convey to the listener. One of the earliest works on this subject is T. Vahidiyan Kamyar's MA dissertation titled "Suprasegmental Events (1972)", Kamyar firstly investigates Farsi lexical stress patterns covering issues such as placement of stress in various grammatical categories such as nouns, adjectives, pronouns,

prepositions, etc. He pays special attention to various lexical stress changes within different sentence types, and investigates relevant phenomena such as stress weakening and strengthening, stress elimination, and stress displacement and shifting. He includes in his survey some discussions about the probable effects of Arabic lexical stress patterns on Farsi words as well as topics on differences of lexical stress patterns within various dialects of Farsi language. In his analysis about intonation events he first examines briefly the stress patterns in Farsi poems, but then mostly focuses on the way stress can be used for emphasizing one word or a group of words within a sentence leading to different meanings, i.e. various readings of the same utterance convey different purposes. His discussion about intonation covers various Farsi sentence types including statements, and interrogatives, imperatives, etc. Kamyar tries to relate various forms of intonation to various intentions of the speaker thus emphasizing the speaker different attitudes and the meanings they communicate with the listener. The most important investigation in his study of intonation is identification of the location of the main sentence stress, i.e. the place of nucleus within sentence/intonation group (the syllable within the sentence/intonation group which receives the strongest stress). So his work as far as investigation of intonation is concerned can be considered as a shallow investigation of Farsi nuclear stress rules, describing changes of place of nucleus and the probable changes resulting from this in the meaning of the sentences (narrow focus). He also points out some changes in the stress patterns of individual words within a sentence such as the neutralization of word stress after nucleus within an intonation group. Throughout his dissertation Kamyar has pointed out (where relevant) some of the interactions of stress and grammar within Farsi Language as well.

Another group of works investigating Farsi prosody and intonation system consists of a few dissertations and projects mostly prepared in Iranian technical universities and research centers as part of a more general task within some speech synthesis and speech recognition pilot studies. Among them one can name some more relevant studies. One such study is "Farsi Language Prosodic structure, Research and implementation using a Speech Synthesizer" (H. Sheikhzadeh et al 1999). In this research, there are some investigations about prosodic features of Farsi and some attempt for quantification of major stress rules and some intonation rules for speech synthesis purposes. The research is mostly concentrated on pitch variations but as far as a serious research on intonation is concerned, only a few major empirical rules have

been suggested. According to Sheikhzadeh, they extracted a few major rules through listening and observations and then classified Farsi sentences to major groups of declaratives, simple questions, questions with question words, exclamations, and imperatives. Through experimentation, using the results of a grammatical analysis, they distinguished the major intonation carrying word (nucleus) in each type of sentence. Afterward they used a basic pitch curve to simulate natural pitch frequency variations, which fitted the proposed intonation pitch patterns[1].

Another work is "Implementation of a Text-To-Speech System for Farsi Language" (Abutalebi et al 2000), which is in fact an implementation of a Text-To-Speech system for Farsi language. It is a concatenative synthesizer, which concatenates Farsi syllables in a TD-PSOLA manner. For achieving a successful synthesis, the researchers have investigated about pitch variations in Farsi sentences and presentation of some rules for modeling these variations. Based on the location of stressed syllable, they obtained a primary pitch curve for each word. In a similar way, obtaining pitch curves for prosodic groups and sentence type effects, the final pitch contour can be determined[2]. The Most important aspect of this work however is that in addition to word stress and sentence type intonation pattern, there is an intermediate prosodic level called prosodic grouping. Such groupings are determined by some boundary markers (they determine begin/end of the groups.). These markers include some grammatical words (prepositions, conjunctions, etc.) or some punctuation marks (commas, semicolons, etc.).

The last work I am going to shortly discuss here is "Analysis of the structure of Farsi Utterances using prosodic Information of speech signal" (Farshad Almas Ganj 1998). In this work Almas G. has tried to use prosodic structure of words to establish a criteria for recognition of Farsi words. He has investigated acoustic correlations such as F0 and duration in the articulation of Farsi words quite carefully and has quantified the differences between productions of some phoneme in different environments (such as within a stressed syllable, word initially, etc.). Then relative to the rate of speech he has calculated a proportional value for each of them. This is a valuable work by itself, but as

---

[1] They selected the Hanning function to simulate natural pitch frequency variations, because it resembles the pitch contour in natural utterances. Due to slow variations of pitch frequency, they could consider a pitch contour to follow a continuous curve, even in unvoiced segments.

[2] So the pitch frequency contour of the input text is resulted by applying the effects of these three parameters: pitch contour of the word, prosodic grouping and the sentence type.

it rarely considers prosodic features beyond individual words, it does not contribute much to the study of Farsi prosody and intonation.

Although such works have contributed (to some extent) to the study of Farsi prosodic structure and intonation, the important thing which can be said about all of them in general is that they have mostly relied on a rough acoustic investigation of f0 over the whole utterance and proposed some general intonational patterns for Farsi sentences without examining in detail sophisticated aspects of the f0 events or even acoustic correlates other than F0 which may play an important role in Farsi prosodic and intonational events. Therefore it is also difficult to discuss these works as being relevant to current debates within the intonational phonology framework. Thus it can be stated that Farsi intonation still awaits a more seriously instrumental description, a task that in this current work the author tries to at least partially achieve.

Chapter 2

Theory of Intonational Representation

Prosody has an integrating function in the organization and production of speech by embedding semantic information, syntactic and morphological structure and the segmental chain in a consistent address frame (Dogil 2001). There are various models for representation and investigation of prosodic and intonation events within linguistic domain of knowledge. As I mentioned earlier, this work is within the framework of AM (Autosegmental Metrical phonology). This is an explicitly phonological approach to intonation in which tunes are characterized in terms of discrete elements, relatively high (H) or relatively low (L) tones, which map onto acoustic F0 targets. The tones are associated with either stressed syllables, lending prominence to a particular syllable or word, or with the edges of phrases, where they take on a junctural function. In this section I will first review some general aspects of prosody and intonation and then try to describe a model of intonation representation (the one which originally was proposed by Pierrehumbert) as the model I will be using in this survey.

2.1 Prosody and Intonation

Prosody and intonation are two important aspects of all languages of the world. They are phonologically obligatory in the sense that just as it is vital to have some choice of phonemes and syllables in the realization of an utterance, every single utterance should also have a prosodic structure representing some choice of intonation pattern. In general prosody refers to the grouping and relative prominence of the elements making up the speech signal. It is mostly reflected in the perceived rhythm of the speech, and affects various aspects of the speech signal. Intonation on the other hand refers to phrase-level characteristics of the melody of the voice. It is used by speakers to mark the pragmatic force of the information in an utterance. The alignment of the intonation contour with the

words is constrained by the prosody, with intonational events falling on the most prominent elements of the prosodic structure and at the edges. As a result, intonational events can often provide information to the listener about the prosodic structure, in addition to carrying a pragmatic message. The term intonation is often used, by extension, to refer to systematic characteristics of the voice melody at larger scales, such as the discourse segment or the paragraph (Pierrehumbert and Hirschberg 1990; Ladd 1996).

There are many important differences in the arrangement of prosodic structures and the organization of intonation events within various languages of the world. First of all various languages differ in the total inventory of intonational patterns and in the pragmatic meanings assigned to particular patterns. Languages with lexical tone (including tone languages such as Mandarin, and classic pitch accent languages such as Japanese) tend to have somewhat simpler intonational systems than intonational or stress languages like Farsi or English. That is presumably due to the fact that much of the F0 contour is taken up with providing phonetic expression of the tones in the words (Hayes 1995). Secondly other differences in the prosodic domain among various languages are reflected in the constraints these languages impose on the composition of the various units. For example at the phrasal level, they differ in how they set up the correspondence between intonational phrases and syntactic and semantic structures. Some languages for instance tend to locate prosodic breaks after a syntactic head, whereas others tend to locate breaks before, or some languages (such as English) permit the main prominence to be located anywhere in the phrase (for the purpose of highlighting or foregrounding particular words) while other languages make little or no use of variable placement of prominence within the phrase, instead moving new information to fixed prosodically prominent positions. Investigation of Farsi data however suggests that this language makes use of a variety of such possibilities. However within certain word order paradigms, Farsi seems to have a tendency toward presuming only certain prosodically prominent positions for highlight words.

2.2   Prosodic Phonology

During the last two decades, we have had significant developments in the phonological theory of prosody and intonation. Prosodic phonology investigates the way in which the flow of speech is organized into a finite set of phonological units. It also considers interactions between phonology and the components of the grammar. The interactions can be expressed in the form of mapping rules that build phonological structure on the basis of morphological, syntactic, and semantic notions, providing the set of phonological units necessary to characterize the domains of application of a large number of phonological rules. While the division of the speech chain into various phonological units makes reference to structures found in the other components of the grammar, a fundamental aspect of prosodic phonology is that the phonological constituents themselves are not necessarily isomorphic to any constituents found elsewhere in the grammar. So it appears that the central concepts at the heart of studies in prosodic theory are the prosodic units (e.g. the syllable, the foot, the intonation phrase, etc.) and the relations defined among these units. Such units are all temporally ordered with bigger units dominating the smaller ones. Within each unit, a relationship of strength is available that singles out one element as more prominent than the other elements of the same type in the group.

Also within some trends of prosodic phonology, intonation contours function as independent pragmatic morphemes. In this view contours indicate the relationship of each utterance to the mutual beliefs that are developed and modified in the course of a conversation. An H accent for instance may mark an intended addition to the mutual beliefs, whereas an L accent marks information that is marked as salient but not to be added. The tremendous variety of understood meanings of patterns in context arises from the interplay of these factors with the goals and assumptions of the interlocutors (Pierrehumbert and Hirschberg (1990).

2.3   Pitch Accents, Boundary Tones, and Prosodic Structure

This study uses an approach, which is based on a model of English intonation developed in Pierrehumbert (1980), later modified and applied to Japanese in Beckman

and Pierrehumbert (1986), and Pierrehumbert and Beckman (1988). In this model there are two level tones as primitives in phonological representation[1], i.e. high (H) and low (L). H and L are employed into two types of tonal events in an utterance: pitch accents and boundary tones.

In this view pitch accents are perceptually significant changes in f0 aligned with particular words in an utterance, giving them added prominence. For English, the inventory of possible pitch accents in Pierrehumbert and Hirschberg (1990) also includes combinations of these two tones as L+H*, L*+H, H*+L, H+L*. The asterisks, which follow individual tones mark the tone, which aligns with the stressed syllable of the word, stress here referring to the prominence assigned by lexical-phonological rules. The unasterisked tone represents a rapid change in f0 immediately preceding or following the stressed syllable in the word.

Boundary tones in this model are either H or L, which align with the "edge" of a phrase, at a phrase boundary. In English they align with two types of phrases, the intermediate (ip) and the intonational phrase (IP). ip boundary tones are phonetically realized as changes in f0 from the last pitch accent of the phrase to the end of the phrase. They are also called phrase accents, and in some texts are denoted by a dash following the tone like H- or L-.

In this model a complete English utterance consists not only of at least one pitch accent and one ip boundary tone, but also an intonational phrase (IP) boundary tone, denoted with a "%" after either the H or L tone. The acoustic cues demarcating intonational and intermediate phrases are the same: pause, pre-boundary syllable lengthening, and reset of the speaker's pitch range. The two phrase types differ from one another in terms of the relative degree of juncture - less for the intermediate phrase, more for the intonational phrase. This degree of juncture is judged from auditory impressions, and no absolute acoustic criteria have been specified as of yet. Unlike pitch accents, boundary

---

[1] The credit for this idea goes to Bruce thesis (1977) on the phonetic realization of the word accents in Swedish. Bruce showed that certain identifiable points in F0 contours are aligned with the segmental string in predictable and largely invariant ways. For example, the drop in pitch that marks a Swedish accented syllable begins in the preceding vowel for "Accent 1" words and in the accented syllable's onset consonant for "Accent 2" words. On the basis of his experimental evidence, Bruce proposed that the truly distinctive features of F0 contours are localized F0 targets, and that rises and falls are simply transitions from one target to another.

tones are not perceived by native speakers to be emphasizing the particular word they align with, but are rather a component of the complex contours that can appear at the ends of phrases, contours which can, for instance, function in negotiating relationships in meaning between whole phrases.

The intermediate and intonational phrases have been postulated for English, though they need not be used in describing other languages. The phrase types are two levels of prosodic structure out of multiple levels that have been claimed to be tonal domains in English and other languages. For instance, Selkirk (1986) proposes the following prosodic hierarchy: syllable, foot, prosodic word, phonological phrase, intonational phrase, and utterance. For Japanese, Pierrehumbert and Beckman (1988) have an accentual phrase as well as intermediate and utterance levels. For Korean, Jun (1993) used only the accentual phrase and the intonational phrase to account for her data. Ultimately, the prosodic structure of any language described will be determined by both the apparent domains of tone assignments as well as the domains of its phonological processes.

2.4   Pitch Register and Intonational Phonology

Any intonational description also has to account for shifts in the speaker's pitch register within an utterance, register defined as a selected f0 range out of the entire pitch range the speaker can possibly use. Pierrehumbert and colleagues claim that for English, a new register can be selected by a speaker for each intermediate phrase. Such shifts are not indicated by any additional diacritics other than brackets for each phrase. Upstep, or register expansion, is represented as either a succession of H* pitch accents, the second accent being upstepped, or if occurring at the end of an utterance, from a combination of an H- (phrase accent) and an H% (boundary tone). Downstep, the successive compression of the pitch range over an utterance, is said to occur after each bitonal pitch accent. A series of such bitonal pitch accents would be posited to describe a common intonational form in English, the "staircase" pattern of successively lower f0 peaks, most clearly observable in "list" intonation. Other shifts in register are presumed to the product of paralinguistic factors, such as speech styles or greater emotional

expression, and thus are absent from a phonological representation of f0. (Harnsberger 1994)

In essence, register shifts are the automatic result of successive tones in the tonal string. The alternative to this view, that register should be independently represented in English as well as other languages, appears in work by Ladd (1983, 1988, 1990, 1992, 1994), Inkelas and Leben (1990); van den Berg, Gussenhoven and Rietveld (1992); and Kubozono (1992) among others. The typical examples used to justify this position are cases of nested downstep (downstep across phrases as well as within them), pitch range reset triggered by following phrases, and "key-raising" in tone languages like Hausa, where adding phonological tones simply cannot account for the observed register shifts (Inkelas and Leben 1990). These examples have inspired various proposals to represent pitch range apart from the tonal string, such as the introduction of a separate register tier, or the employment of a metrical tree, or the addition of a third tone. These challenges to the standard theory have been met with several criticisms, many revolving around the "mix and match problem" (Hayes 1994), where the introduction of new elements into the representational system also entails a series of stipulations limiting their use in the system. Beckman and Pierrehumbert (1992) also point out the failure of many of the experiments that inspired phonological representation of pitch range to control for various pragmatic interpretations of test sentences. Much of the independent use of pitch range, they argue, may be accounted for with reference to pragmatic context, and thus should not appear in a phonological representation. (Harnsberger 1994)

2.5 Criteria for identifying phrase boundaries

To a certain extent most of the intonational phrase boundaries tend to fall on the end of major syntactic structures reflecting sort of syntax-phonology mapping. However such mapping is only partial and one can find many cases in which such mapping cannot fully be marked out i.e. they only correspond with constituents of sentences in a somewhat loose way. So identifying phrase boundaries between intonation groups is not always an easy task. Although in slowly carefully spoken utterances, the phonetic correlates of boundaries between intonation groups are quite straightforward, in actual connected

speech there are many cases where it remains difficult to decide whether a boundary is present or not. This problem with inexperienced investigator of data is even graver. That is due to the fact that spontaneous speech involves lots of hesitations, repetitions, false starts and incomplete sentences.

Judgment that a phrase boundary is present would in an ideal situation be based on external criteria, i.e. on phonetic cues present at the actual boundary. But in reality such phonetic cues may be either ambiguous or not present at all. Therefore internal criteria must also play a role here: that is our judgment that the application of the external criteria produces chunks of utterance all of which have pitch patterns which accord with acceptable 'whole' intonation pattern. The identification of phrase boundaries is therefore something of a circular business; we establish some intonation groups in cases where all the external criteria such as pause conspire to make the assignment of a boundary relatively certain; we note sorts of internal intonational structure occurring in such cases and this enables us to make decisions in those cases where the external criteria are less unambiguous, and in some difficult cases we take grammatical or semantic criteria into account. In short the external criteria are namely pause, lengthening of the final syllable of an intonation group, and the change of pitch on an unaccented syllable. (Cruttenden 1986)

Chapter 3

An Overview of Farsi Tonal Events, and Prosodic Constituent Structures

This chapter starts with a quick review of the corpora used in this study, the speakers participated in recording of the data and the relevant technical arrangements. I also investigate Farsi prosodic structures in terms of the number of prosodic levels, prosodic units and their various realizations, as well as pitch accent types in Farsi.

3.1 Corpus of Study

In this study two corpora have been prepared. The first corpus includes individual Farsi sentences consisting of different types i.e. various declarative sentences, interrogatives (including both yes/no questions and questions with Wh-words), and finally some imperative utterances. The second corpus includes three texts. The first one is a carefully designed interview based on a real interview with a contemporary Iranian writer published on a website, the next two are extracts from simple narratives (a fable story and a simple narration) chosen among some works in children's literature (Karimzadeh "Chehel Ghese"). Both the interview and narrations have been carefully modified to enable them better capture the target sentences. Investigation of the recorded data shows that texts of the second corpus mirror the natural articulation of Farsi sentences more realistically. This should be due to the pragmatic hints and clues readers pick up while reading the texts.

All sample sentences in the corpora have been recorded by four native speakers with Tehrani accent. It is the predominant accent in Tehran, but it can be considered as the

standard language register in Iran because it is widely used in media and in formal uses of language all through the country[1].

As speaker variability (such as gender, accent, age, speech rate, and phone realizations) is one of the main difficulties in investigating speech signals, an attempt was made to choose the speakers mostly within the same age group (3 of them are in their 20s). Also special attention was given not to use speakers with creaky and rough voices in order to maintain an f0 contour as smooth as possible. Also in selection of the speakers much attention was given to their educational background, social level, etc. This was for ensuring to give relatively homogeneous and stable recorded corpora. There are two male speakers aged 22, and 33, and two female speakers aged 23 and 27. In the first corpus, speakers read individual sentences with a short pause after each sentence, which provides enough spacing before and after each utterance. In the second corpus however there is a more natural flow of speech, in which sentences follow one another with shorter pauses and sometimes overlaps. However, due to carefully designed context, the target sentences have only been slightly affected.

Both for my recording and analysis I used Praat[2]. The data had been recorded in mono wave format with a sampling rate of 16000 KHz, 16 bit, and 44 Kb/second. Getting F0 measurement in Praat is actually a very easy task, but the accuracy of the contour is influenced by many factors. For example sharp jumps in pitch contour may be due to individual voice quality such as irregularities of the vibrations produced by the vocal folds. Also places with no pitch values show the presence of problems in recording (as well as normal voiceless consonants which are inevitable). Such characteristics however can also be due to malfunctioning of the software. Most of the times (in such cases) mistakes could be removed by adjustment of some analysis parameters such as adjusting the minimum and maximum F0 values so that they fit to the frequency range of

---

[1] It is vital to mention that there are numerous dialects within modern Farsi. Although most of them are mutually understandable, there are still considerable differences in the choice of phonemes and the general prosodic and intonation specifications. So to some extent the effects of such investigation on other varieties of Farsi may bring up different results.

[2] This is a comprehensive program developed by Paul Boersma and David Weenink in the Institute of Phonetic Sciences of University of Amsterdam. Praat can be used for speech analysis and synthesis. The manipulation package includes general as well as specialized tools built on a general-purpose GUI shell.

speaker's voice. But sometimes there was a need for rerecording of the data, which prolonged the recording phase beyond expectation.

Something worth of mentioning is that besides the normal differences between male and female speakers in their absolute pitch values[1], the female speakers had sudden jumps between the L and H tone steps in their pitch accents. Even among speakers of the same sex, these abrupt movements were different; one of them (the 27 year old woman) had even more rapid movements as well as more exaggerated pitch accents. The 23 year old female speaker was very relaxed in her readings thus she maintained a slower speech rate with more distinct pauses between intermediate phrases while the other female speaker had a higher rate of speech which led to more coarticulations and less distinguishable boundaries.

3.2 Farsi Prosodic units

As this is one of the first studies of Farsi prosody and intonation within AM model, the author only tries to include simple sentences (mostly utterances which consist of a single Intonational Phrase) in this work. Although in narrations used as contextual corpus there were some complex sentences (combinations of subordinating and main clauses), they were deliberately ignored as being the focus of investigation in this preliminary survey. So in identifying the prosodic structure of Farsi all Intonational phrases could easily be identified in the recorded data without any serious trouble (almost each one of them corresponds to a single IP.). However a shallow investigation of the longer instances of speech (compound and complex utterances) suggests existence of some intermediate prosodic units in Farsi.

The typical intonation pattern of simple Farsi declarative sentences (where the phrasal boundaries can be well distinguished) is the common terminal intonation pattern, in which the pitch decreases at the very end of the utterance (here each utterance is mostly identical to one intonational phrase), thus leading to a low boundary tone. But in longer instances of declarative sentences (such as compounds or complex sentences),

---

[1] Females have higher absolute pitch than males when speaking, due to shorter vocal tracts.

Farsi utterances take slightly different patterns. For example in complex Farsi utterances, whichever clause (subordinate or main clause) comes first, the last stressed syllable at the end of the clause rises considerably before the other clause starts, i.e. there is a high phase tone at the end of the phrase (sometimes when a high pitch is present, it will be preserved without any lowering) leading to a progressive intonation pattern, i.e. the pitch either increases slightly or does not show any lowering at the very end. Apparently this behavior of the pitch contour signals the message is not completed yet, thus making the hearer wait for the rest of the utterance[1]. In such complex or compound sentences the natural flow of speech treats the boundary tones within the sentence as weaker than final ones, thus suggesting some sort of hierarchical organization of intonation phrasing in Farsi in which some intonational units hold together more closely and firmly. The treatment of such phenomena (Ladd 1996) is still a place of dispute between various scholars. One can consider an additional middle prosodic unit namely an intermediate intonation phrase as Beckman and Pierrehumbert (1986) proposed for such apparently smaller units in AM theory. They suggest the phonological difference is that an intermediate phrase (ip) is followed by a 'phrase-accent / tone', whereas a full IP is followed by a 'phrase-accent / tone' and 'boundary tone'[2]. On the other hand, as Ladd suggests (to keep our prosodic units inventory as small as possible), we can give permission to compounds prosodic domains (CPD). Here CPD is interpreted as a prosodic domain of type IP (Intonational phrase), whose immediate constituents are again other IPs. The only difference between IPs here is the type of boundaries they take, thus releasing us from considering an additional separate intermediate prosodic unit.

For this study as I said most of the selected utterances consist of simple sentences (i.e. the IP dominates immediately an ip). So wherever there is a need of presenting the hierarchical structure of utterances, I present the structure of IP in terms of the accentual

---

[1] However further investigation of such phenomena reveals that in case of complex sentences when the main clause comes first, the whole sentence starts with a considerably higher pitch register compared to the time when the subordinate clause comes first. Also In the former case, the subordinate clause, which follows the main clause, will be considerably deaccented, while in the latter case (when the main clause follows the subordinate clause), pitch accents are relatively well preserved all through the clause.

[2] But preliminary investigation of Farsi data suggests that the most important difference here lies in differing degrees of end-of-phrase lengthening: This is greater at the end of IPs than ips, and in the absence of pause is the main indicator that the end of an IP has been reached.

phrases without indicating the structure of the ip. But I assume three prosodic units for Farsi marked by intonation, namely IP (intonational phrase), ip (intermediate phrase) and a smaller unit above word level called Accentual Phrase. IP and ip are marked by H or L tones while the default AP boundary tone is H.

3.2.1 Accentual phrases

Analysis of the hierarchical prosodic structure involves investigation of such issues as deciding over the number of prosodic levels, or resolving the tonal and intonational default patterns, etc. Such investigations have been carried on for some major languages of the world. However a consensus over such descriptions for any language is yet to be achieved. So it is obvious discussions on Farsi prosodic structure in such a preliminary work cannot be free from dispute. However one should admit that any attempt in this regard would pave the way for more focused investigations and better future surveys.

AP (Accentual phrase) is the smallest unit of the hierarchy in a prosodic model. The purpose of this part is the description of some of the properties of this unit. Investigation of Farsi utterances suggests that in this language the lowest prosodic hierarchy level is AP, which contains one or more content words. The default AP boundary tone is H, which often is right demarcated by the H* of the typical Farsi pitch accent pattern i.e. L+H* (figure 3-1). However it can be realized as L when the AP includes a focused word.

An AP in general includes only one word but in case of a focus or a relative clause construction it can include several words. Usually each content word is realized with a pitch accent thus representing an AP unless the word comes after a focused word, which makes it deaccented. In such cases the AP corresponds to chunks of utterance larger than individual words.

In order to make the structure of Farsi APs more concrete I will present here a simple declarative sentence (figure 3-2), which consists of one IP dominating 3 APs[1]. This is a

---

[1] The iP has been ignored here as there was only one iP dominated by IP.

Figure 3-1: Default Farsi pitch accent pattern (L+H*) in a short declarative sentence
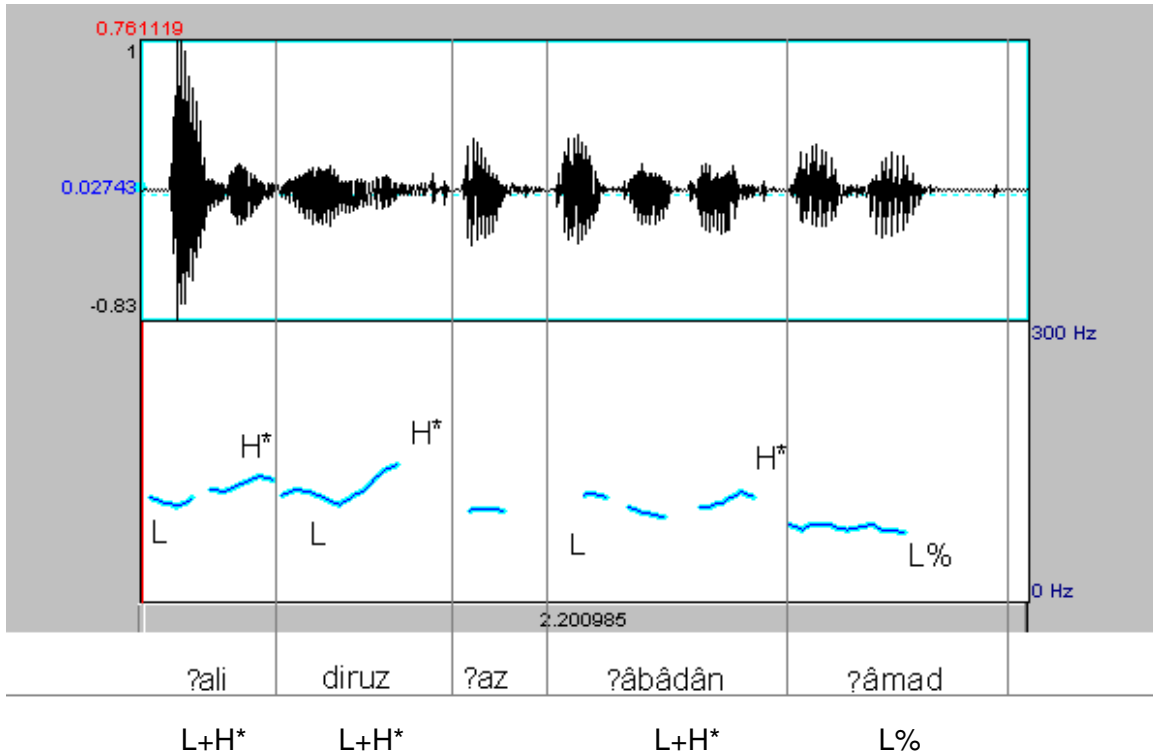
- ?ali   diruz   ?az ?âbâdân ?âmad. (Ali came from Abadan yesterday.)
  ali yesterday from abadan   came.

simple unmarked Farsi SOV sentence consisting of the subject "barâdar-aš" formed by joining "barâdar" (brother) and the enclitic pronoun "-aš" (his), the object "qazâ râ" (food Acc) and the verb "mipazad" (cooks). The pitch accents of this utterance are respectively L+H*   L+H*   H*   L%. Here "his brother" is realized as one word with the enclitic pronoun "-aš" (his) attached to "brother", the last syllable of "barâdar" carries the primary lexical stress while the enclitic "-aš" remains unstressed but rides on the high tone of the previous stressed syllable. The same is true with the object "qazâ râ" in which the specific object marker "râ" rides on the high tone of the previous stressed syllable (i.e. final syllable of "qazâ")[2]. The verb "mipazad" has its lexical stress on its first syllable. So

---

[2] Both in transcription and in segmentation of words, I follow the Farsi orthography i.e. writing the enclitic pronoun "-aš" as part of the word "barâdar" but the clitic marker "râ" as an independent word separated from "qazâ".

the first two APs get the default Farsi pitch accent pattern i.e. L+H*, while the last AP is realized as H* because the pitch accented syllable is the first syllable of the Accentual Phrase. Again examination of F0 in figure 2 confirms that pitch accents in Farsi can have a starred tone to indicate their association with the accented syllables.
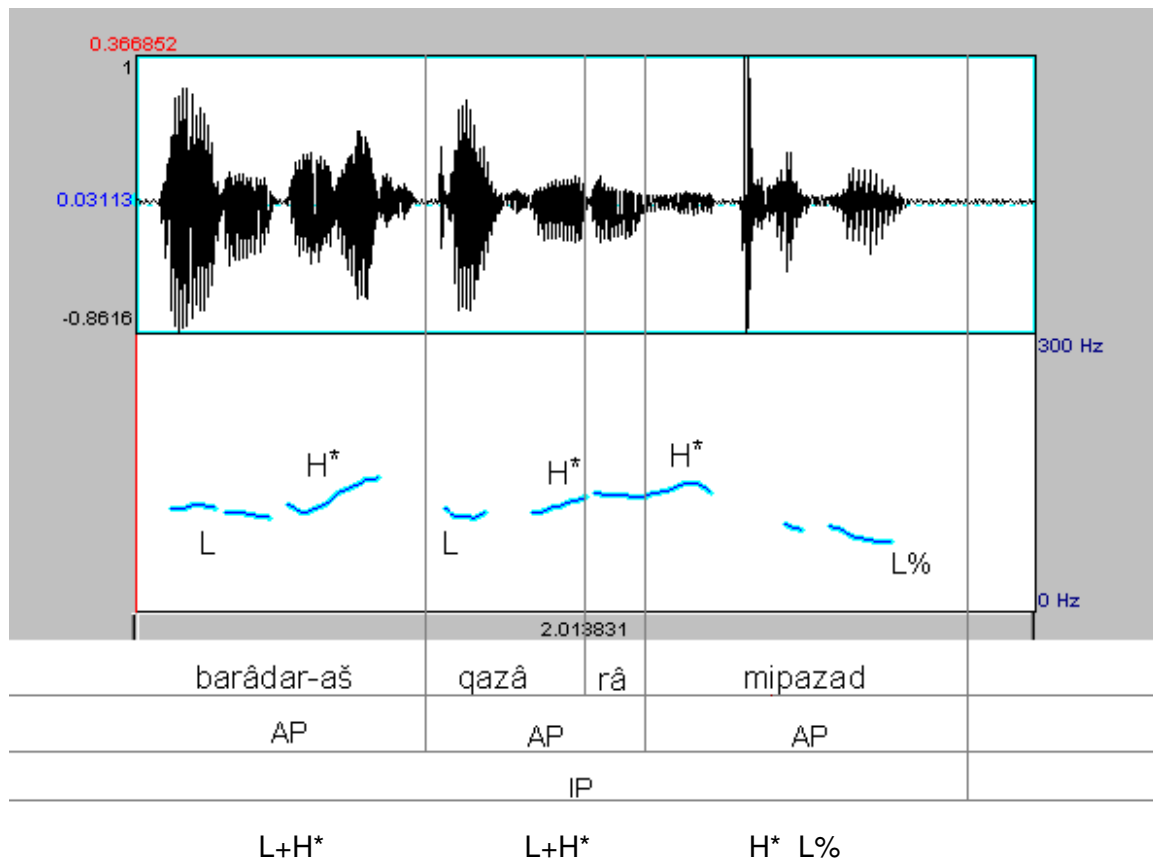

- barâdar-aš qazâ râ mipazad. (His brother makes the food.)
  brother-his food Acc makes.

Figure 3-2: simple declarative sentence hierarchical structure and pitch accent type

As it can be observed from the examples, the actual assignment of the L+H* tones to words (APs) is quite straightforward in Farsi. When there is a two-syllable word, the L tone assigns to the first syllable and H to the second. When L+H* assigns to words of three syllables or more, the H tone still appears on the last syllable (the stressed syllable of the word), however when the last syllable is unstressed like the example in figure 3-2,

the H extends over the rest (the unstressed syllable) of the word, thus making it ride on the high tone of the previous syllable. In monosyllabic words, both tones (L and H) assign to a single syllable. The phonetic realization for such an assignment is a rapid rise from low in the speaker's pitch range. (figure 3-3)
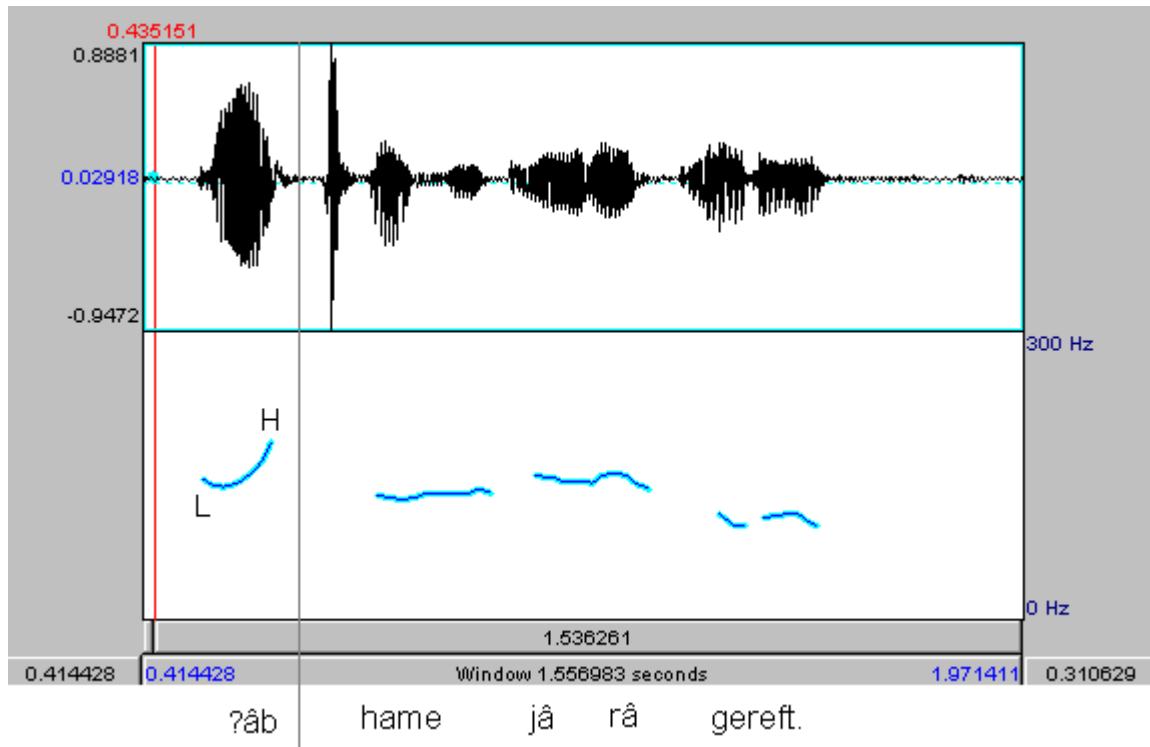

- ?âb   hame   jâ   râ   gereft.

 water   all   place Acc took-3SG

(Water filled all places.)

Figure 3-3: assignment of L+H* pitch accent to a monosyllabic word (AP)

3.2.2   Intermediate phrases and Intonational phrases

Based on the readings of statements by native speakers of Farsi (Tehrani dialect), a typical Farsi utterance consists of at least one intonational phrase (IP) which dominates

one or several intermediate phrases (ip) each one of them composed of one or more accentual phrases i.e. rising contours (LH) aligned with individual content words. The final constituent of the sentence, usually the verb (as Farsi is a verb final language), carries the boundary tone of the IP. The register in which H tones are realized gradually narrows across the utterance, a trend commonly observed in the intonation of other languages as well. Below is an example of a Farsi utterance with numerous instances of LH tones (This is called Declination) (figure 3-4).
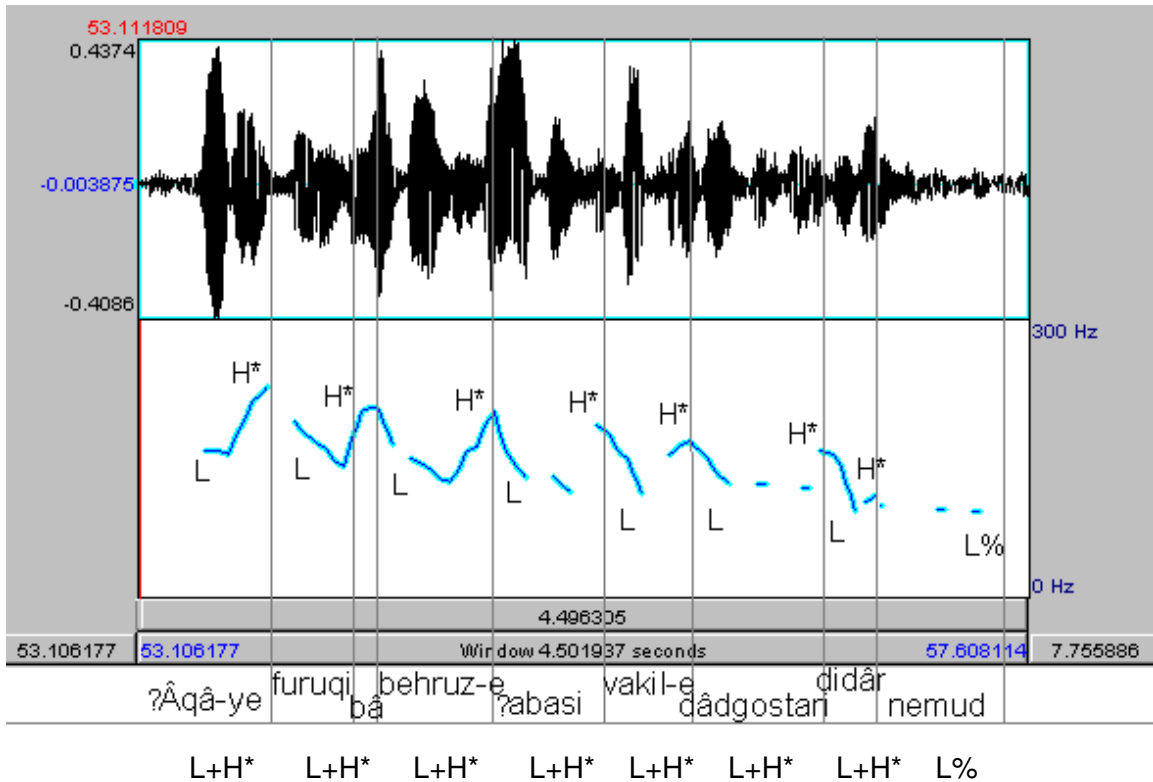
- ?Âqâ-ye   furuqi   bâ   behruz-e   ?abasi   vakil-e   dâdgostari   didâr   nemud.

  Mr.-Ezf[1]   forughi   with   behruz-Ezf   abasi   lawyer-Ezf   judiciary   meeting   did.

(Mr. Forughi met Behrooz Abassi, a lawyer of justice ministry.)

Figure 3-4: Example of a Farsi utterance with numerous instances of L+H* pitch accent showing downsetpping

---

[1] Ezafe, look at footnote page 6 for an explanation of ezafe

The Intonational Phrase (IP) in Farsi is signaled by L% boundary tones on the right edge of declarative utterances and Wh-questions, and H% for yes/no questions. However such boundary tones are not the only clues of signaling the IP. Another property that marks the right edge of an IP is lengthening of the final vowel. An Intermediate Phrase (iP), on the other hand, has no boundary tones, but can be identified by the blocking of downstep. The left edge of an iP corresponds to the left edge of an AP that is the initial AP in a syntactic phrase. An iP may contain one or more APs, and is thus the domain of downstep. The end of ip is marked either by H or L tones. Non-final iPs are specially marked by their higher pitch register. Figure 3-5 below shows a Farsi IP dominating two iPs immediately each one of them consists of some APs. First iP noticeably represent phrase without a boundary tone, i.e. the end of this phrase (the verb "bordam") is well marked with an H tone. Also it has a higher pitch register compared to an iP, which is followed immediately with a boundary tone (like the second iP).

3.3   The Tonal Inventory of Farsi

This section outlines the observed tonal inventory of Farsi based on the recorded corpora. Here the various boundary tones, and the inventory of pitch accents will be reviewed. It is however crucial to note that in case of pitch accent types further examination of more natural improvised Farsi conversations is needed before the inventory is considered as an exclusive and finalized list.

3.3.1   Pitch accents

The default pitch accent in Farsi is the bitonal pattern L+H* (figure 3-5), which corresponds to a sharp rise from a low point in the speaker's range to a high peak featured on the accented syllable. This default pattern can also be realized as H* when the pitch accented syllable is the first syllable of an Accentual Phrase. This mostly happens at end of the utterances i.e. the unmarked position of verbs (In Farsi it is usually the verbs which can get an accent on their first syllable.). (figure 3-2)
An idiosyncrasy of Farsi, which is worthy of mentioning, is the special tone used with relative clauses. The beginning and the end of the relative clause in this language are

marked by a high tone and the middle of the clause shows a low plateau (This shows a great degree of mapping between syntax and phonology.). (figure 3-6)

man   pul bordam tâ meqdâri qazâ begiram

iP	iP

IP
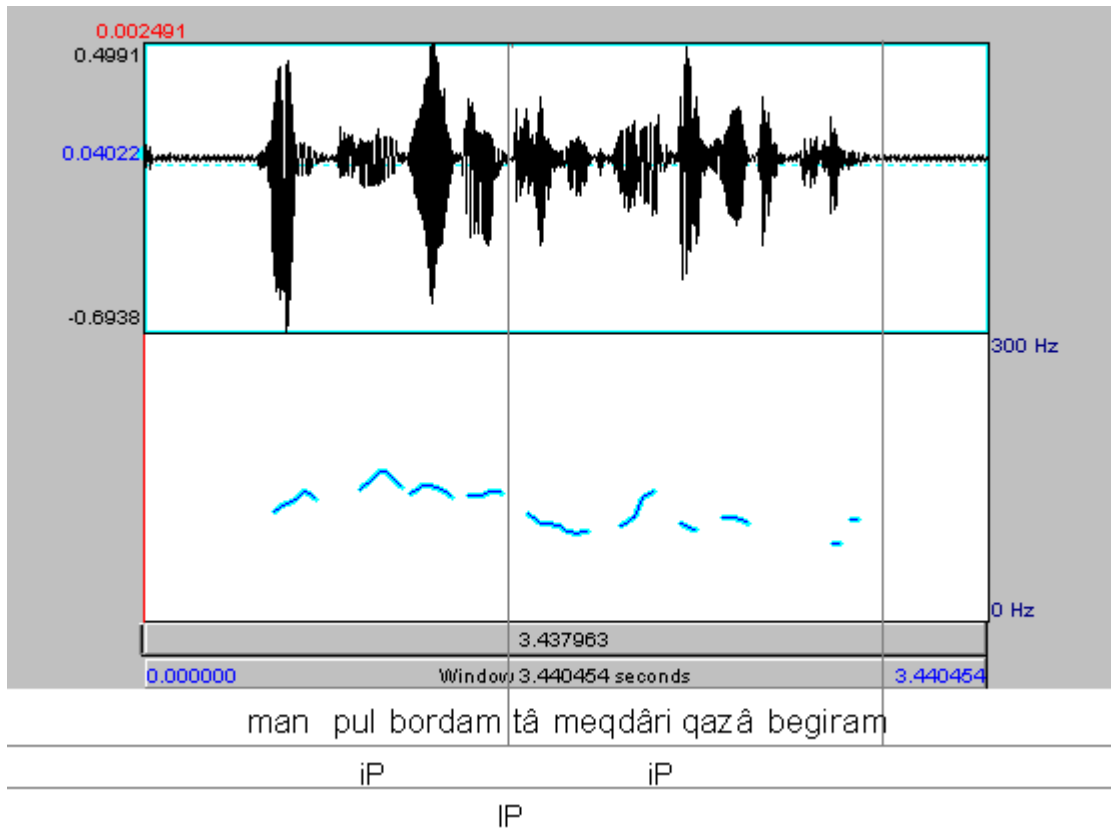
- man   pul   bordam   tâ   meqdâri qazâ   begiram.
   I   money   took-1SG in order that   some   food   get-1SG
 (I am taking some money to get some food.)

Figure 3-5: Example of an H tone with relatively higher pitch register iP

3.2.2 Edge Tones

Two tonally marked levels of phrasing may be suggested for Farsi; An intermediate phrase level and an intonational phrase level. The two phrase accents delimiting the edges of intermediate phrases in Farsi are:

L-: a low phrase tone, which controls the pitch between the last pitch accent and the edge of the intermediate phrase. If this stretch is not large, a fall of pitch to a low part in the speaker's range will be observed. If the stretch is long, L- creates a flat valley stretch between the nuclear and the edge of the intermediate phrase.

H-: a high phrase tone, which usually creates a slightly rising pitch for the stretch between a nuclear accent and the edge of the intermediate phrase (figure 3-5).
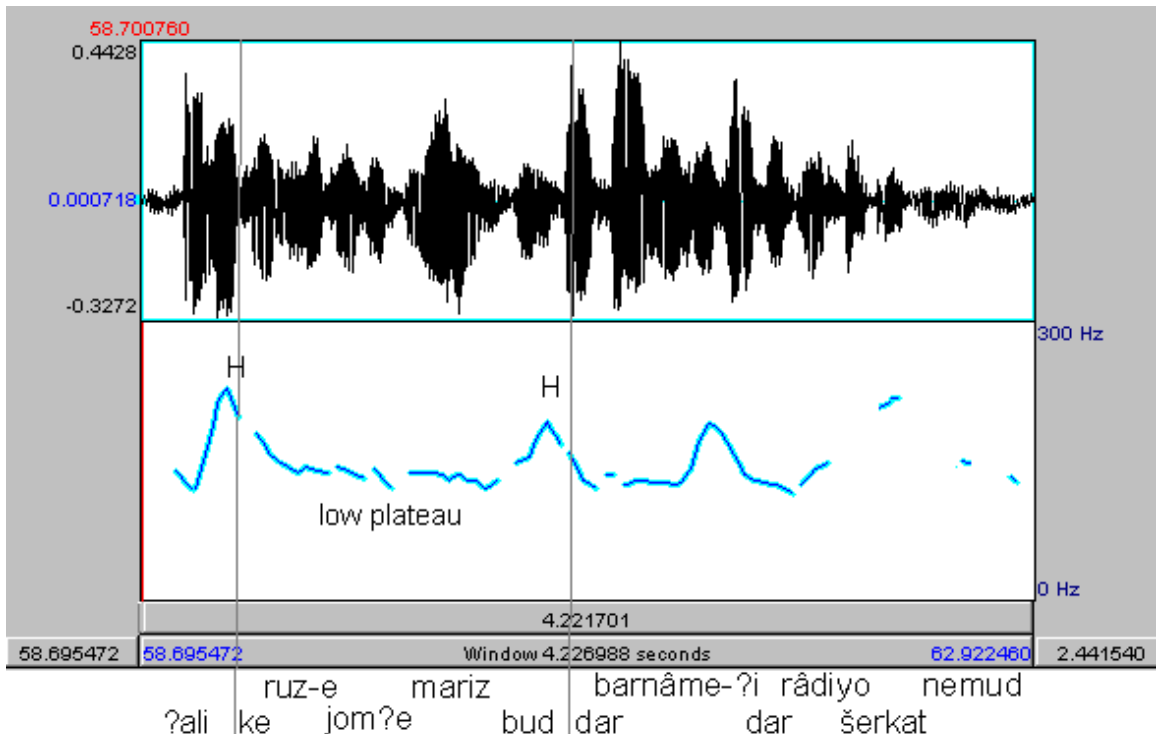
- ?ali ke   ruz-e jom?e mariz bud dar barnâme-?i   dar râdiyo   šerkat   nemud.
   ali who day-Ezf friday ill   was in program-Indef in   radio part-taking   did-3SG. (Ali who was ill on Friday participated in a program in radio.)

Figure 3-6: A relative clause in Farsi marked by two high tones at the beginning and end of the clause with a low plateau in the middle

3.3.2 Boundary Tones

The two boundary tones demarcating the edges of intonational phrases in Farsi are L% and H%. Intonational phrases are formed of at least one intermediate phrase. The last intermediate phrase occurring in an intonational phrase will have its phrase accent followed immediately by the boundary tone. In other words, the last intermediate phrase accent combines with the intonational boundary tones to yield one of the configurations, namely L-L%, L-H%, H-L%. L% can be observed on the final word in declaratives and H% typically ends yes/no questions. Further investigation of these tones is the aim of the next chapter.

Chapter 4

An Investigation of Farsi Intonation patterns

In this chapter I investigate the intonation structure of Farsi sentences and main clauses. Here I will analyze the main Farsi sentence types, namely declaratives, interrogatives, imperatives, and exclamatory sentences in terms of their prosodic structures, and the way they exploit the various prosodic features.

4.1   Farsi Intonation (A Quick Look)

Intonation, which is often called the melody of language, refers to the pattern of pitch changes used in speech. It involves the occurrences of recurring pitch patterns, each of which is used with a set of relatively consistent meanings, either on single words or on groups of words of varying length. Farsi is a stress language. This means that in this language pitch variations over words[1] doesn't change their meanings, but does change for example an utterance from a statement to a question, or emphasize different words for pragmatic functions[2]. Thus in Farsi like other stress languages, tone can be used to convey an attitude or change a statement into a question (but tone alone does not change the meaning of individual words).

In Farsi as well as other languages, there are various functions of intonation; among them one can name grammatical function, delimiting role, expressive function, etc. For

---

[1] It can be one word or several words grouped together.

[2] However an intonation contour in a tonal language like Chinese can change the meaning sometimes, but can also be used as it is in stress languages. Tonal languages differ from stress or non-tonal languages like English where pitch doesn't have those same functions.

example when pitch variations constitute the only difference between a question and an assertion, intonation has a grammatical function or when intonation signals continuation or termination, it has a delimiting function. In this section I will try to analyze first the way Farsi achieves arrangement of different intonation patterns (Which tonal events are employed and how for distinguishing different intonation patterns?) and second what meanings or implications are conveyed by these patterns.

4.2   Investigation of Intonational patterns in Farsi

As mentioned above, Farsi is a stress accent language, i.e. it uses pitch primarily for intonational purposes. Intonation in Farsi primarily affects sentence meaning e.g. question vs. assertion, etc. It is also used for indicating the information structure of an utterance i.e. signaling which information items are more important and new. Yet another use of intonation in Farsi is showing some extra-linguistic attitudes such as surprise, impatience, sarcasm, etc. Here we will try to have a close look at different sentence types in Farsi.

4.2.1   Declarative Sentences

The sample declaratives recorded for this study consists of some individual sentences with neutral focus, mostly covering the unmarked Farsi word order i.e. SOV. But there are a few sentences among them, which have marked word order like VSO, or SVO. Nearly all declaratives in this corpus are simple sentences in which the whole sentence maps into one intonational phrase. However in the narratives, there are several compound and complex declaratives as well as some simple statements with narrow focuses. In such sentences, the IP usually consists of more than one iP. Some of these sentences have been investigated throughout this chapter as well.

Nearly all simple declaratives investigated consist of accentual phrases with L+H* pitch accent pattern. Each AP contains one or more content words. However the H tone in the right hand side of the AP is sometimes realized as L when the AP includes a focused

word. Also if an AP is IP-final, the default H* tone will be replaced by L* followed by the boundary tone of the IP, either L% or H%.

By default in nearly all Farsi sentence types either declaratives, and interrogatives, or imperatives, the None-final APs preserve their default L+H* pitch accents all through the sentence. This pitch accent usually is in agreement with the lexical stress pattern of Farsi content words[1]. The final AP however occurs where usually the verb group stands by default. As it was mentioned in chapter 1, Farsi verbs have quite complex behavior in choosing the place of their lexical stress[2]. Considering this fact and the fact that the end of intonational phrase (here a simple sentence) is followed by a boundary tone makes the last AP get various patterns. However having a relatively unified pitch accent structure all through intonational phrases, various Farsi sentence types can mostly be differentiated according to type of their boundary tones.

Investigation of the declarative sentences in the sample corpora (each read by four different speakers including both sexes) suggests that in all Farsi statements f0 starts usually at the bottom of the speaker's pitch range and reaches the highest point in the speaker's pitch range for the given intermediate phrase within the first or second starting AP. In following the APs, again f0 drops in the first syllable of the AP to nearly the bottom of the speaker's pitch range, and then increases to the highest point in the speaker's pitch range at the boundary of the AP. However the peak of each subsequent H tone is lower than the preceding one. This leads to the establishment of a rather smooth declination in the f0 all through the statement. The longer the intonation phrase (here it corresponds to a declarative) the better this declination can be depicted. (Figure 3-4)

---

[1] As it was mentioned in chapter one most of the content words in Farsi have their primary stress in their final syllable, so the default AP pitch accent usually is in accordance with the lexical stress pattern.

[2] In fact different affixes added to the stem for indication of tense or aspect shift the default stress pattern of the verb.

4.2.1.1 Types of Declarative Sentences

As it was mentioned above, in Farsi pitch accent patterns are the same whether the sentence is declarative or interrogative. So the sentence types are distinguished mainly by the boundary tone. That is why here we will have a discussion about the declaratives in terms of their boundary tones.

In Farsi in a neutral declarative sentence typically the nuclear pitch accent is an H* on the last stressed syllable, followed by a L% boundary tone, and all pre-nuclear pitch accents are L+H*. This pattern leads into a terminal intonation in which the pitch decreases at the end, thus is seems as if it were signaling the message is completed. (figure 4-1)
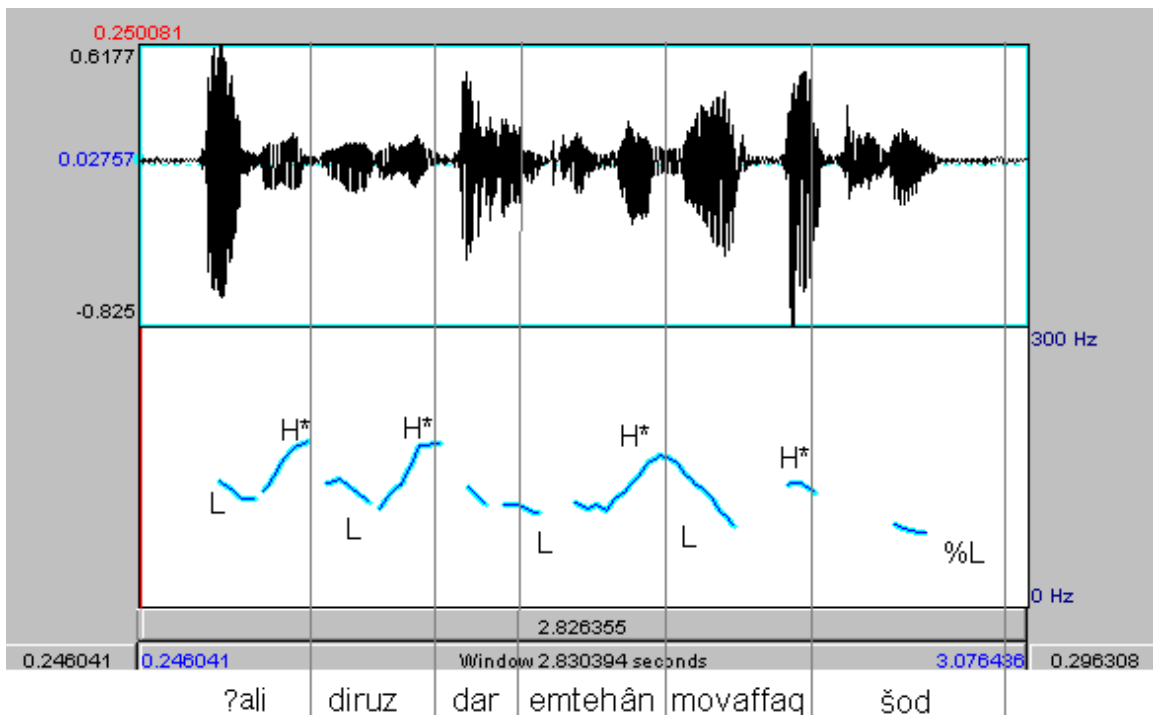

- ?ali   diruz   dar   emtehân movaffaq   šod.

  ali yesterday in   exam   success   became-3SG.

(Yesterday Ali succeeded in the exam.)

Figure 4-1: A typical neutral declarative with %L tone boundary

As majority of Farsi verbs have a compound construction e.g. consisting of a noun, adjective, etc. followed by a verb, the nucleus pitch accent usually occurs on the last syllable of the none verb element of the verb group.

Apart from this neutral declarative sentence pattern, which in fact occurs quite rarely in the natural flow of speech, there is a more common progressive intonation for declaratives in which throughout the declarative sentence the pitch either increases slightly or does not show any lowering at the end, thus leading to a high tone boundary %H. This resembles the high tone boundary in yes/no questions, however here differences in pitch range and duration of the last stretched syllable differentiates between them. This signals the message is not completed yet. It is typically used as the sentence pattern for declaratives in longer instances of speech (such as a narration, etc.). (figure 4-2) Compound or complex sentences also get the same rising at their junctions i.e. a raised pitch at the end of the preceding clause and a falling tone with a low boundary at the end of the last clause. At the same time some verbs[1] usually are used with this progressive intonation pattern like "porsidan" (ask), "goftan" (say), "pâsox dâdan" (answer), etc. (figure 4-3)

---

[1] They are in fact semantically determined and can be classified as members of one group.
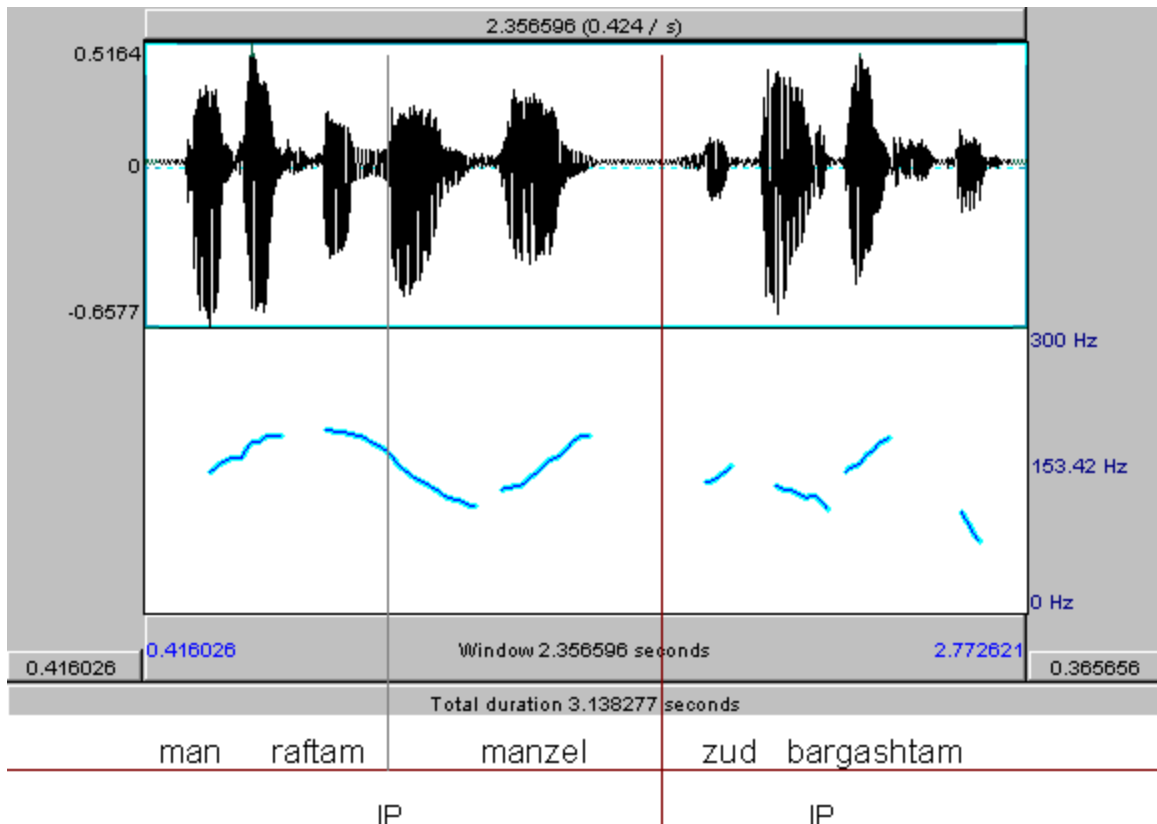
- man   raftam   manzel. zud bargashtam.

  I   went-1SG home.   soon returned-1SG.

(I went home. Soon I came back.)

Figure 4-2: A more common declarative with a raised pitch and a high tone boundary
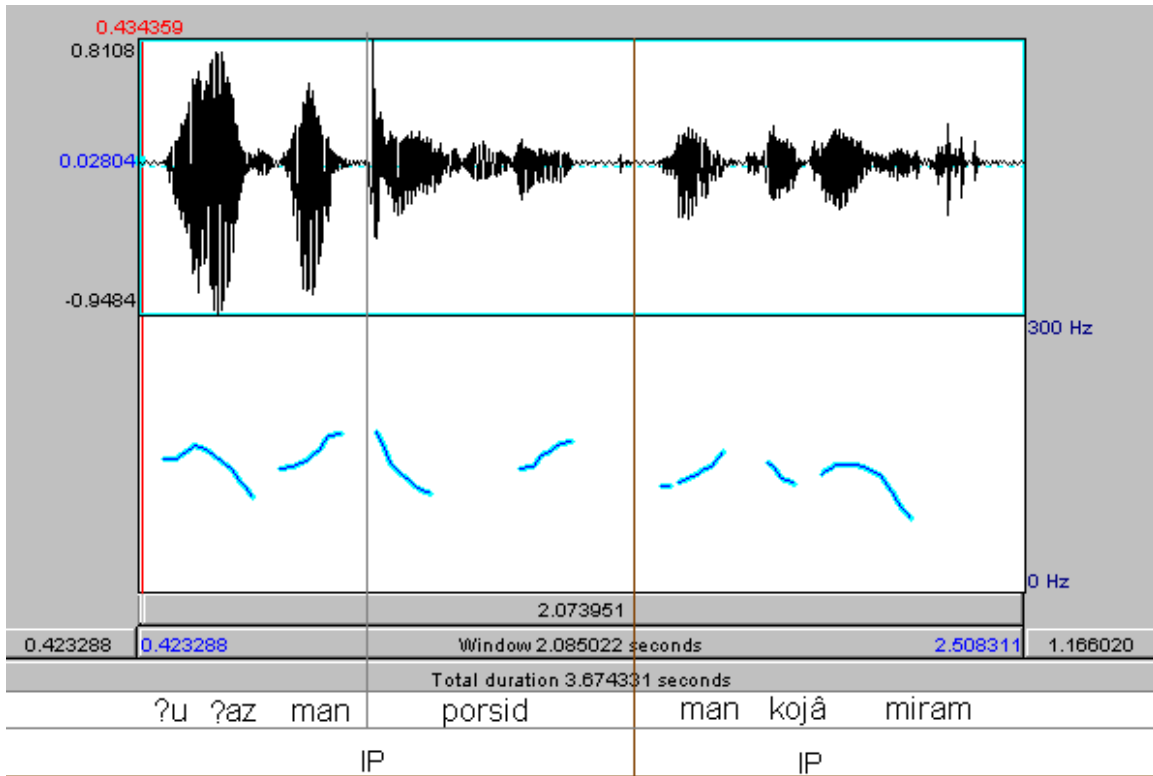
- ?u ?az   man   porsid   man kojâ   miram.
  he from   me   asked-3SG   I   where go-1SG.
(He asked me where I was going.)

Figure 4-3: A progressive tone of a declarative initiated by the verb "porsidan" (ask)

4.2.1.2   Declaratives with Focus

In Farsi, like many other languages, focus affects the prosodic structure of the whole sentence. It assigns the nuclear pitch accent to the focus word. As a result all post-focal pitch accents will be de-accented. Furthermore, on the phonetic level, the focused word lengthens in duration considerably, while words before and after it usually shorten in duration. (figure 4-4)
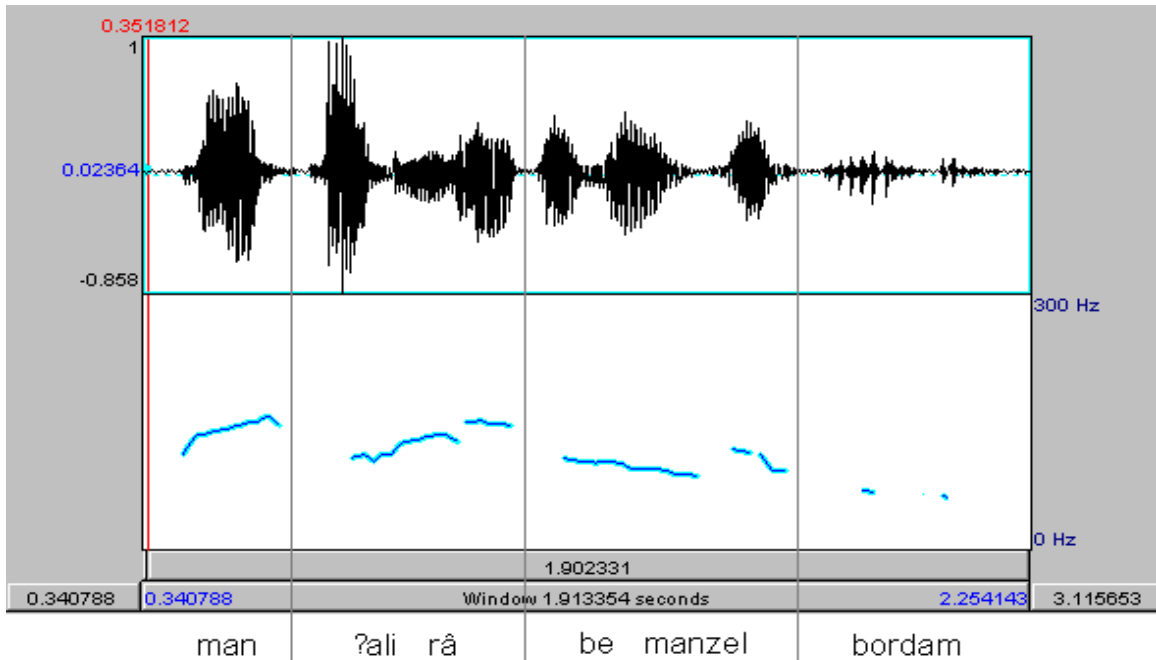
(a) man ?ali râ be manzel bordam.
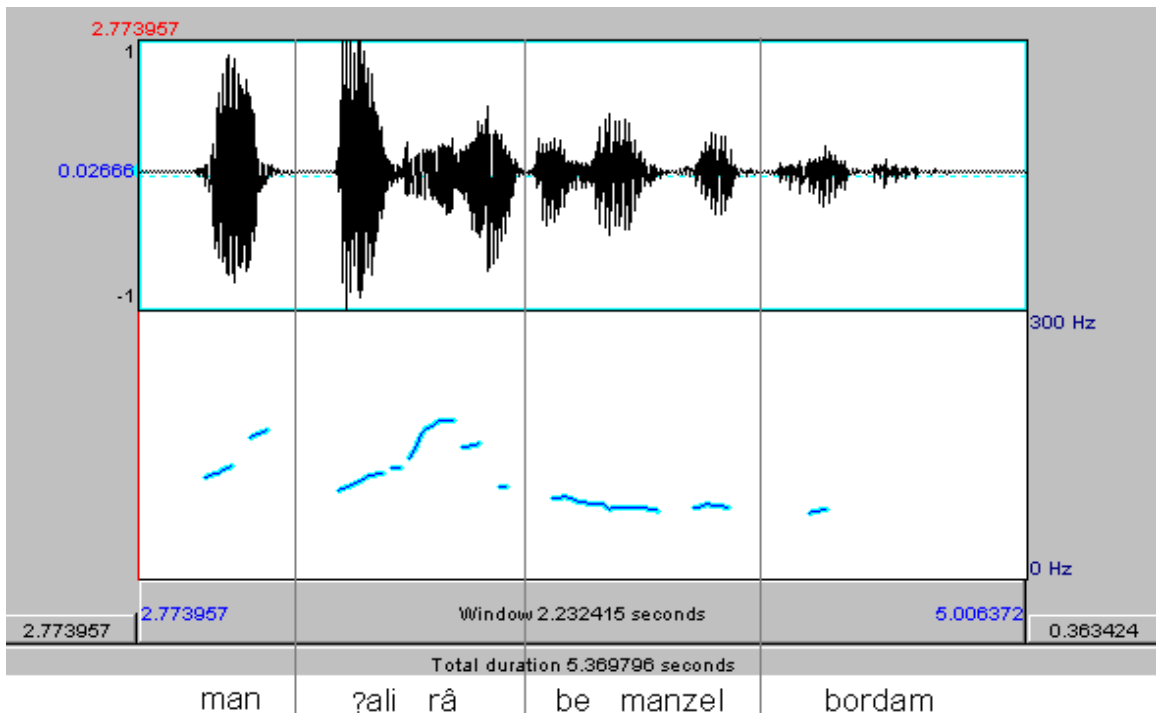

(b) man ?ALI râ be manzel bordam. (I took Ali home.)

     I ALI Acc to home took-1SG.

Figure 4-4: (a) a neutral declarative vs. (b) a declarative with a narrow focus on the object "?ali"

It is significant to mention that Farsi, being a free word order language, makes use of various devices (syntactic and prosodic) separately or in combination for focusing on the words and phrases of interest. This provides a relatively large number of potential word reorderings along with alternations in the prosodic gestures, which needs further investigations if one wants to depict a clearer picture of the behavior of Farsi prosodic system.

Usually Farsi uses topicalization for emphasis on some words and phrases. When the topicalized word or phrase appears at the beginning of the sentence, it forms slightly different variations on the original intonational pattern of the sentence, thus in purely descriptive terms it creates a special kind of intonation. However usually the reordering of the sentence elements does not change the overall structure of intonational phrasing, e.g. the overall declination of a declarative sentence, the low boundary tone at the end of the phrase, etc. However some of the marked word orders of Farsi makes crucial changes to the original prosodic phrasing of the sentence. For example when a verb immediately follows the subject while the object of the sentence (or other complements) shift their location toward landing sites at the end of the sentence, there is a crucial restructuring of the intonational phrase, i.e. when verb moves out of its unmarked position (the end of sentence) and lands immediately after the subject, the structure of intonation group changes from a simple IP immediately dominating an iP, to an IP which dominates two iPs. Proof for such changes can easily be traced through evidences such as picking up a higher pitch register, a sustained H tone at the end of iP where the verb is placed, etc. (figure 4-5).
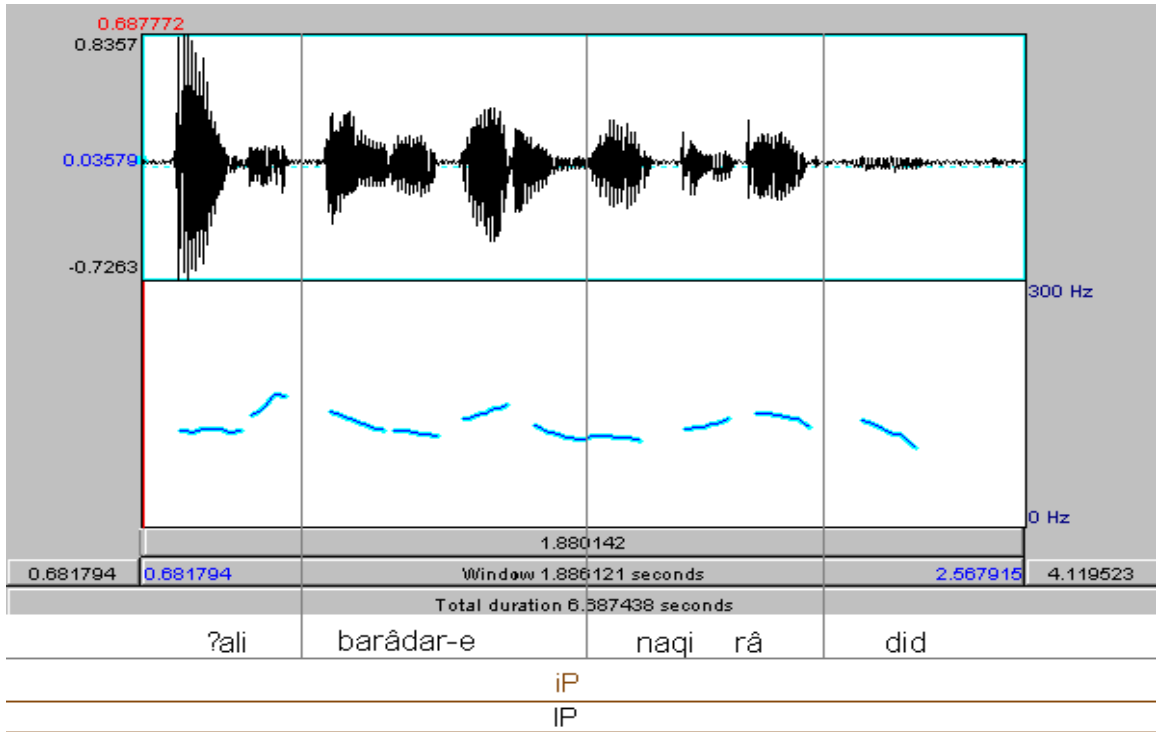
(a) ?ali barâdar-e   naqi   râ   did. (Ali saw the brother of Naghi)
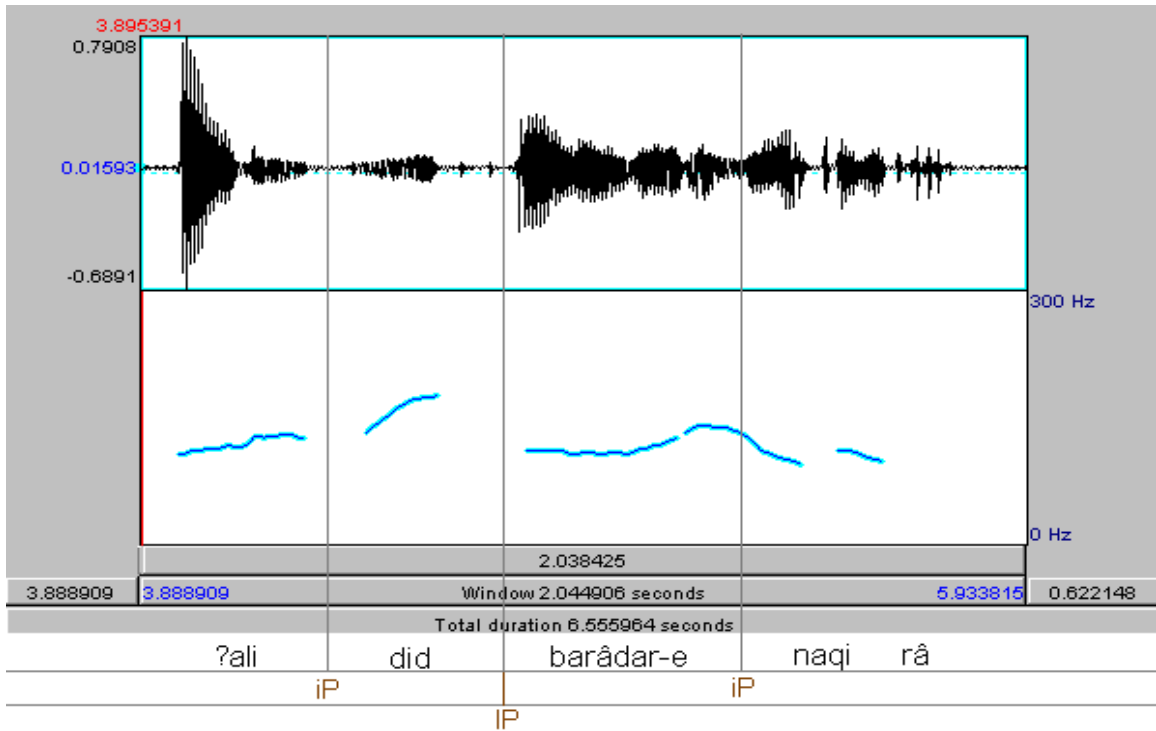
(b) ?ali   did   barâdar-e   naqi   râ. (Ali saw the brother of Naghi.)

    ali   saw-3SG   brother-Ezf   naghi   Acc

Figure 4-5: Change of prosodic phrasing due to a different word order

4.3.2   Interrogatives

As it was mentioned earlier by default in nearly all Farsi sentence types either declaratives, and interrogatives, or imperatives, the Non-final APs preserve their default L+H* pitch accents all through the sentence. So the most important clues in differentiating different sentence type is the boundary tone in the final accentual phrase at the end of the IP. In this part I will have a close look on Farsi interrogatives, including both yes/no questions and wh-questions, their boundary tone variations as well as their general prosodic behaviors. But before considering such events in detail, I will try to review quite briefly the grammatical structure of Farsi interrogative sentences and their variations.

4.3.2.1   Formation of Interrogatives in Farsi

In Farsi the yes/no questions are simply made by adding the word "?âyâ" to the beginning of sentences. After insertion of "?âyâ" the word order remains unaffected, however, in ordinary speech people only make such questions with using the proper rising intonation. The situation is a bit complicated for formation of wh-questions . Farsi unlike many other languages in which the wh-question is formed by either overtly moving the wh-word to the scopal position or leaving it in the situ, employs both strategies (Ganjavi 2000). That is to say wh-words can be in situ or they can be moved overtly to the beginning of the sentence.

- nâder   ki   râ   did? (Who did Nader see?)
   Nader who Acc saw

- ki   râ   nâder did? (Who did Nader see?)
   who Acc Nader saw

It is worth mentioning that such an option can only be seen in simple sentences. In Complex NPs, and nested clauses, overt extraction of the wh-element from an island results in total ungrammaticality.
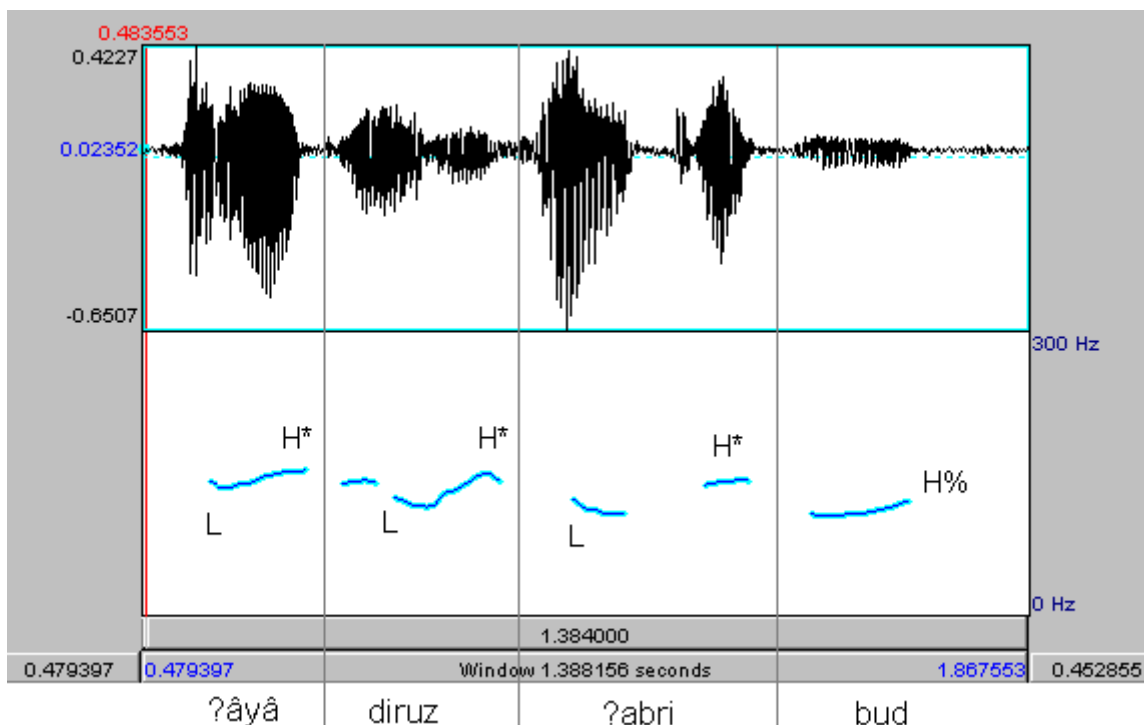
4.3.2.2 Yes/No Questions

In yes/no question generally the pitch slightly increases on the last syllable of the intonational phrase. This signals a wait for response from the speaker hence maintains a continuous flow of conversation between two interlocutors. In this way yes/no questions are different form wh-questions firstly due to the fact that they take a high boundary tone at the end of the phrase (sometimes the last syllable gets an LH boundary) while in wh-questions the boundary tone is almost always L%, secondly because in yes/no questions contrary to wh-questions there is no deaccenting of any words all through the utterance (figure 4-6).

Figure 4-6 shows a typical Farsi yes/no question with the default L+H* pitch accents and the final H% boundary tone spreads over the last syllable (here the verb "bud"). Just like declaratives there is a declination all through the utterance but with a much shorter downsteps. Investigation of many sample recorded utterances reveals that the lower bound of the pitch range in yes/no questions is higher than that of the lower bound of the pitch range of neutral declaratives (and obviously that of wh-questions' because due to deaccenting which happens after the wh-word, wh-questions have intrinsically much lower lower-bound pitch limit). However I couldn't observe (at least in the sample data I recorded for this survey) any significant difference between neutral declaratives and yes/no questions as far as their upper bound of pitch was concerned[1].

The question particle "?âyâ" which almost always only comes at the beginning of a yes/no question is not comparable to wh-words in wh-questions as far as its function and its effect on the utterance is concerned. Here "?âyâ" does not enforce any sort of deaccenting to rest of the intonational phrase. In fact All APs after "?âyâ" preserve their original pitch accents. However the presence of this particle at the beginning of the

---

[1] All investigations concerning degree in prosody are relative to individual native speakers of course.

- ?âyâ   diruz   ?abri   bud? (Was it cloudy yesterday?)
 Q-particle   yesterday   cloudy   was-3SG?

Figure 4-6: A typical Farsi yes/no question

utterance make the rest of the utterance take a slightly lower register compared to the pitch register of the corresponding question without question particle "?âyâ". Another difference is that the H% of such questions also takes a significant higher register compared to that of normal question[1] (figure 4-7).

---

[1] In daily conversations, native speakers of Farsi use such questions, however in more formal registers of Farsi and specially in written forms "?âyâ" is frequently used to initiate a yes/no question.
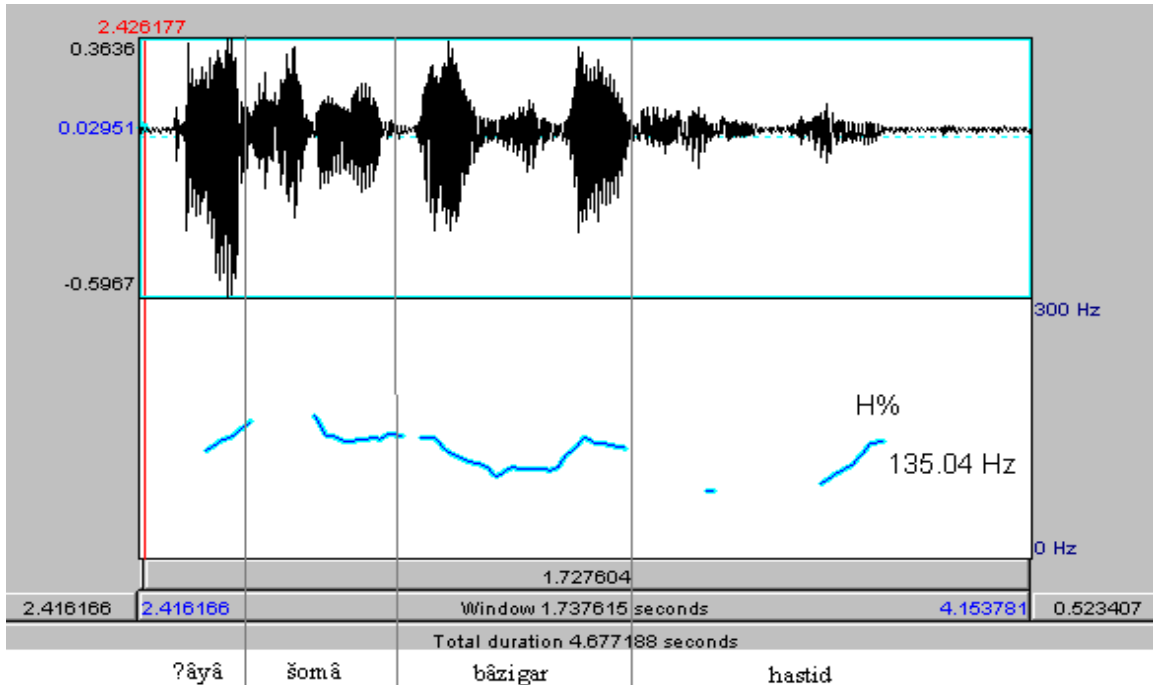
(a) "?âyâ" šomâ bâzigar hastid? (Are you an actor?)

Q-particle you actor are?
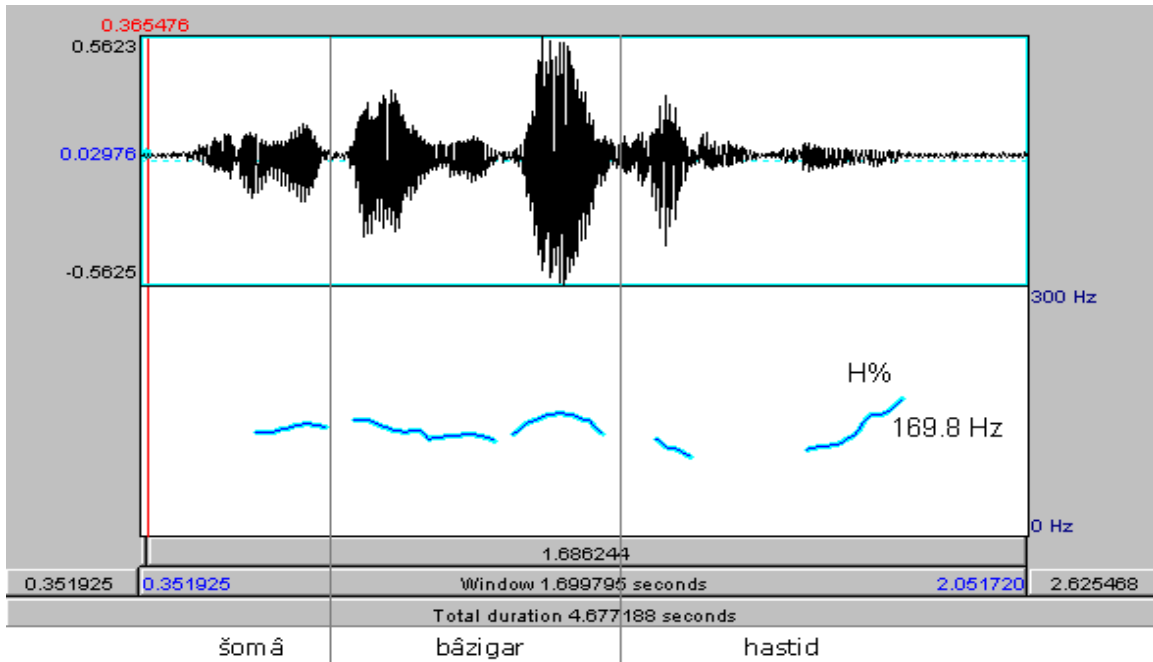
(b) šomâ bâzigar hastid? (Are you an actor?)

you actor are?

Figure 4-7: A normal yes/no question (a) vs. an echo question (b). The latter has slightly higher pitch register and a significant higher H%.

Now I would like here to add something about my observations on focus in such questions. Just like declaratives with focus here a narrow focus on any word affects the prosodic structure of the whole utterance. It assigns a prominent pitch accent to the focus word, and consequently all the post-focal pitch accents will be de-accented. However because of existence of a high boundary tone right at the end of the utterance, there is a small but noticeable upset in the low deaccented plateau area up until the very final syllable in which suddenly there is sharp rise of the pitch toward the H% boundary right at the end of the phrase. The interesting thing about this final rise is its unbelievably high pitch range at the boundary tone (in case of male speakers for example the H% tone could expand up to 220 Hz). (figure 4-8) The rise I am talking about is relative to the same speaker when he/she makes a yes/no question with or without a focus word. For example in the sample data presented here, the same speaker whose voice has the high pitch of 215 Hz at the end of the question with a focus word, quite rarely exceeds 170 Hz in normal questions without a focus word[1].
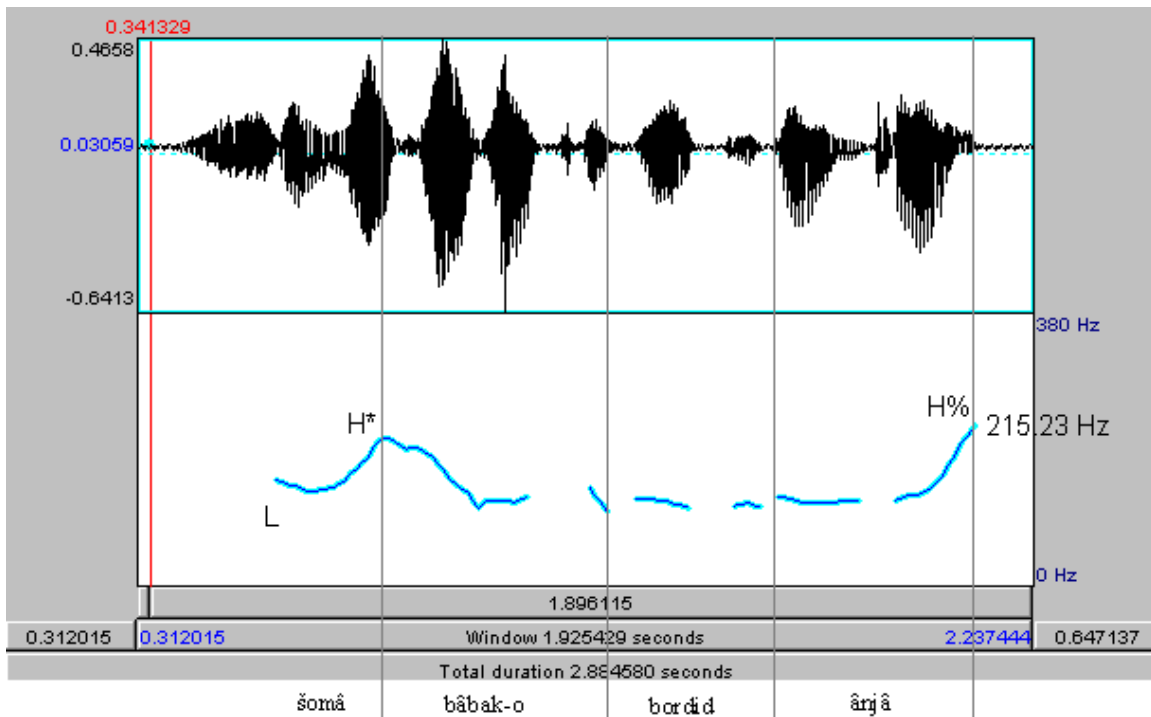
- ŠOMĂ   bâbak-o   bordid   ânjâ? (Did YOU take Babak there?)
 you   babak-Acc took-2PL there?

Figure 4-8: An echo yes/no question with focus

---

[1] To compare just look at figure 4-7 , (a) and (b). It is in fact the same speaker who has read the sentences in 4-8.
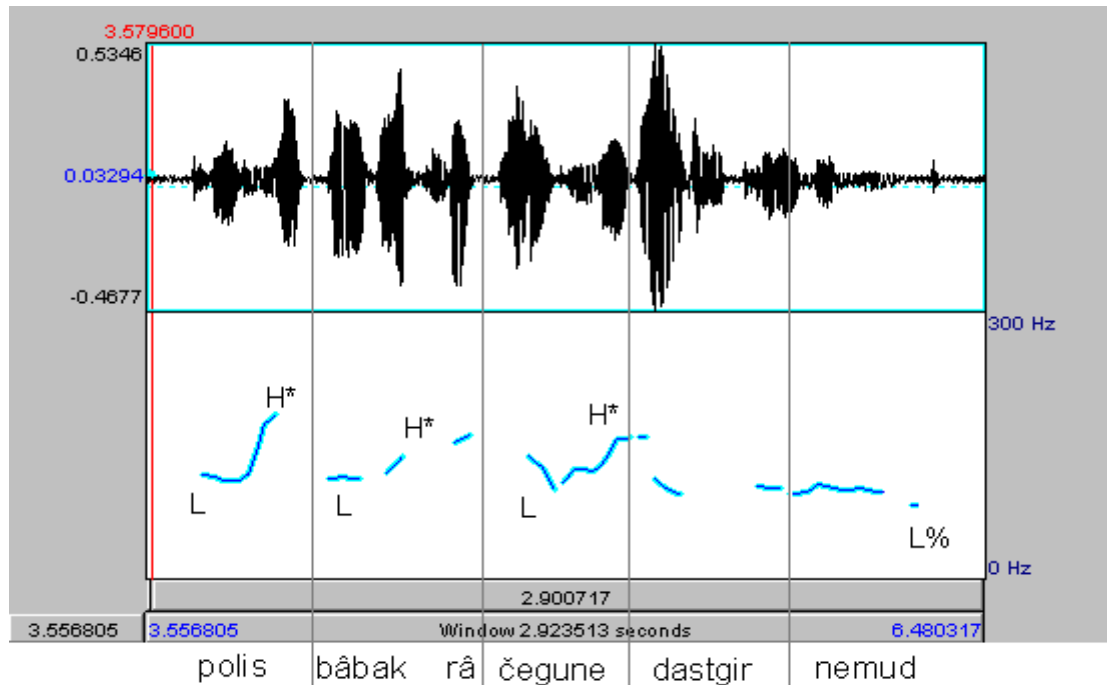
4.3.2.3 Wh-Questions

Wh-questions in general are tonally marked in a similar, though not quite identical, way as the narrow focus cases I explained in the previous section on declaratives. Where wh-word occurs in the utterance, the pitch increases on the stressed syllable of that wh-word. When the wh-word comes at the beginning of an utterance, the remaining of the utterance will become deaccented, but in case the wh-word occurs anywhere else, the phrases before the wh-word preserve their original pitch accents but the ones following the wh-word all will be deaccented (figure 4-9).

After the process of deaccenting, one can still traces the weak deaccented L+H patterns, however here the pitch movements are so lowered in their pitch range that L and H tones become almost identical situating at the same level. This as I said resembles the narrow focus case in declarative sentences. There I showed how focus affects the prosodic structure of the whole Farsi declarative, i.e. by assigning the nuclear pitch accent to the focus word, all post-focal pitch accents became de-accented. Yet one can differentiate between these two (wh-questions and declaratives with focus) on the bases of some acoustic evidence, i.e. the focus word uses a greater register expansion than the wh-word, consequently the after-focus-deaccenting in declaratives with focus is more severe than the deaccenting which happens after the wh-word, thus in the former case the f0 gets a noticeable lower frequency in its lower bound limit.
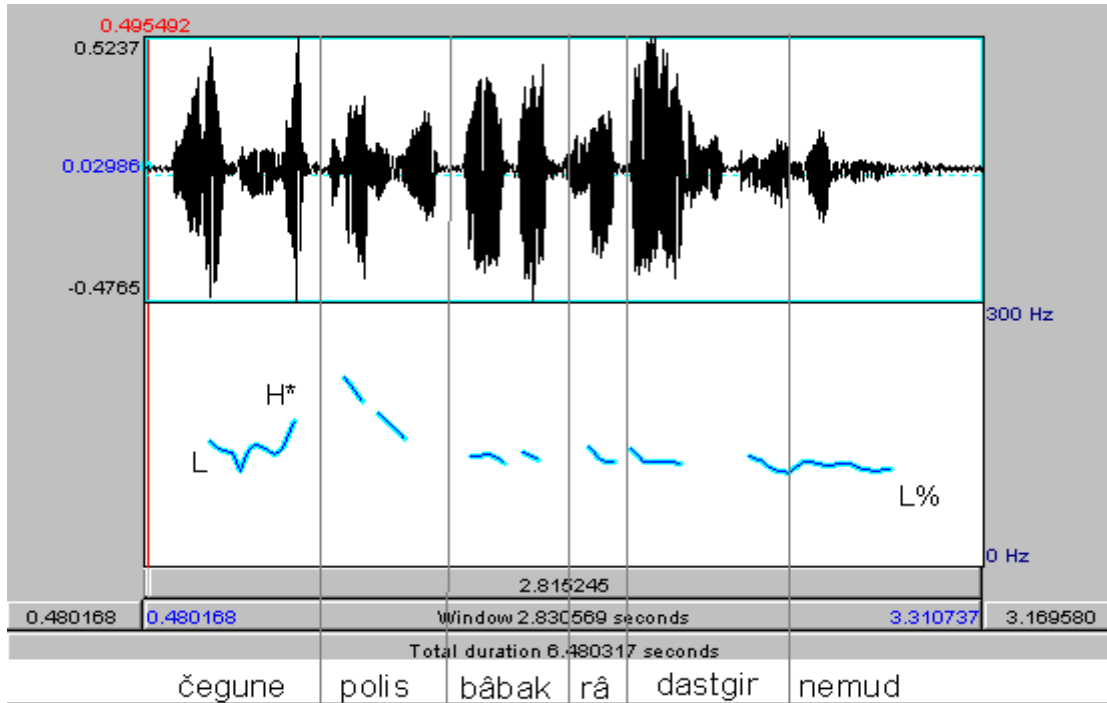
(a) čegune polis bâbak râ dastgir nemud? (How did the police arrest Babak?)

(b) čegune polis bâbak   râ dastgir nemud? (How did the police arrest Babak?)

   how   polis   babak Acc arrest did-3SG?

Figure 4-9: Loss of pitch accents after a wh-word in a Wh-question, (the words before wh-word all preserve their original pitch accents.)

4.3.2.3.1   Focus in Wh-Questions

Focus in wh-questions is a little bit complicated. In fact it is quite rich in terms of the presence of various prosodic features and intonational events. That is why I decided to have a closer look at this phenomenon in this work. There was not any data available in my recorded corpora covering this, because the idea of investigating focus in wh-questions occurred to my mind in the course of writing this chapter, i.e. well after the recording phase. So my investigations on this matter are on the bases of my own recorded speech.

What happens in a wh-question in case of the presence of focus is in fact bunching up of some various intonational events together, which lead to complicated interactions among them. As an example I try to analyze here figure 4-10. In this example sentence (a) is a normal wh-question with wh-word coming right at the beginning of the sentence.
As it can be observed the stressed syllable of the wh-word gets prominence and as a normal wh-question the remaining of the utterance become deaccented. Finally at the final syllable there is a very slight fall leading to the final L% boundary tone (the slightness is because of deaccenting process). In the second sentence (b), we expect "šomâ" (you) to be deaccented right after the wh-word, but because of focus on the object, i.e. "bâbak", there is a suspension over the deaccenting process of wh-word. What happens next is very interesting. The first phrase "kojâ šomâ" ends up forming like an independent iP accompanied by the expected high tone on its final syllable in such situations[1]. Now the rest of the utterance starts right with the focus word "bâbak". It gets the nucleus accent of the phrase and quite expectedly deaccents the final verb. This leads to a relatively steep fall merging rapidly with the final low boundary tone right at the last syllable of the utterance.

---

[1] I am not sure if there is an iP here, however because of the existence of a quite distinguishable pause at the end of this phrase, I would rather believe that there is an iP here. In fact omitting the pause here makes the whole utterance sounds like an abnormal and weird question.
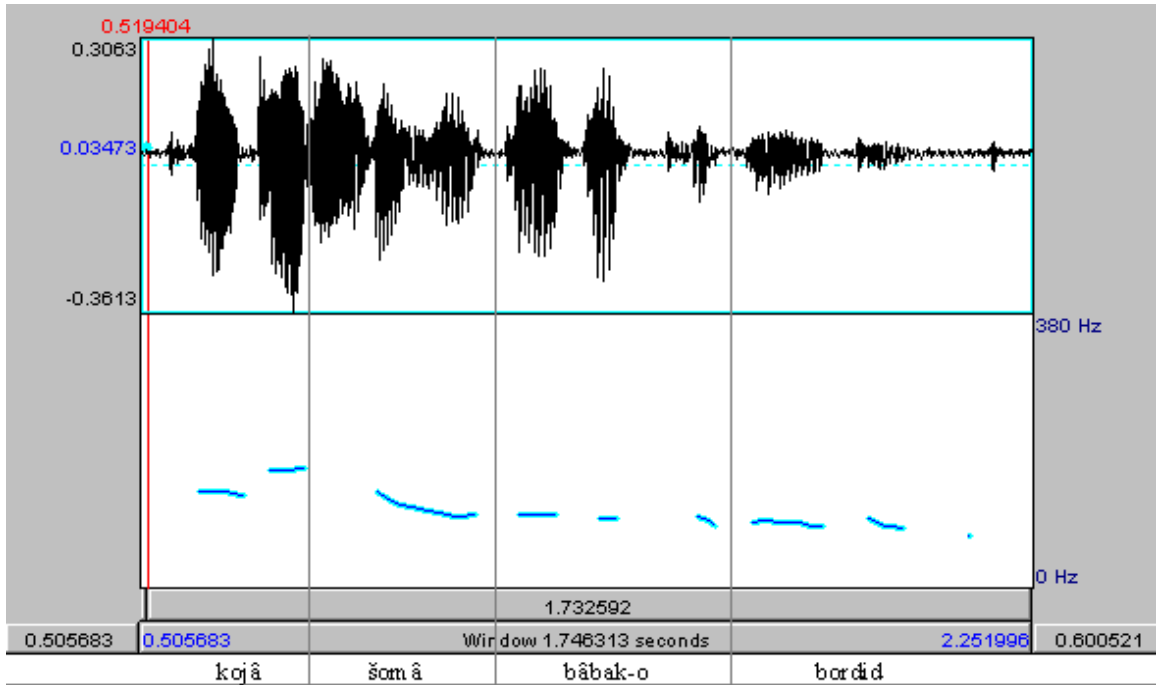
(a) kojâ   šomâ bâbak-o   bordid? (Where did you take Babak to?)

   where   you babak-Acc took-2PL?
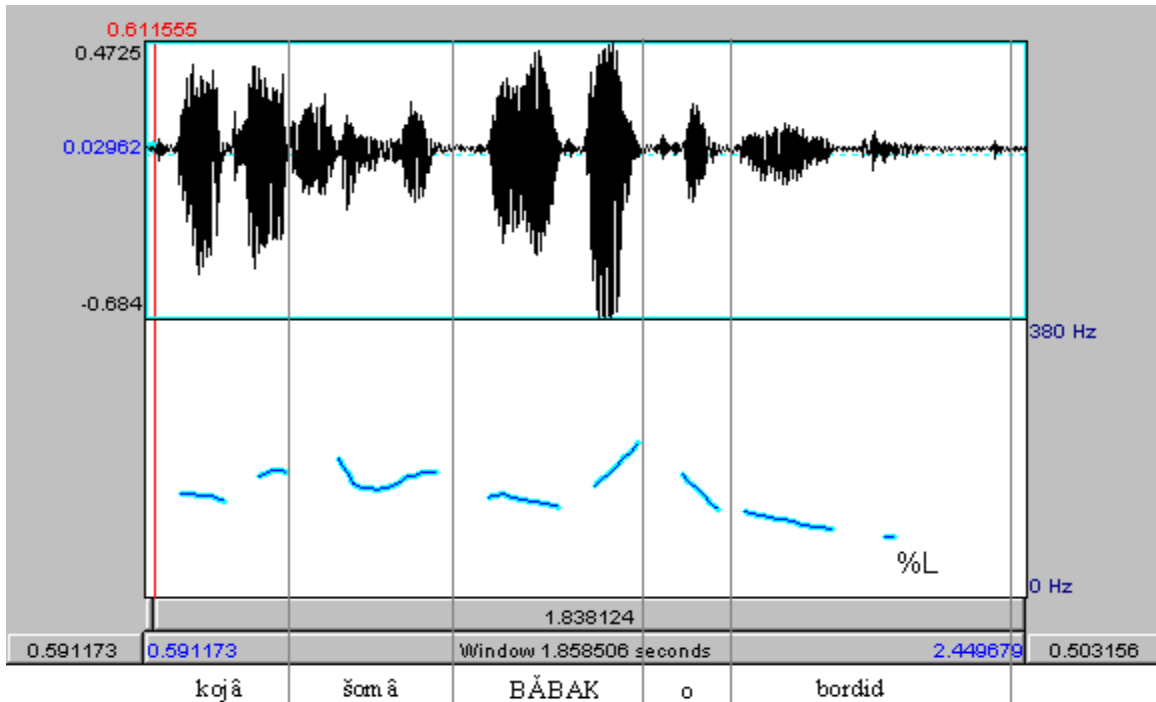
(b) kojâ   šomâ   BĂBAK-o   bordid? (Where did you take Babak to?)

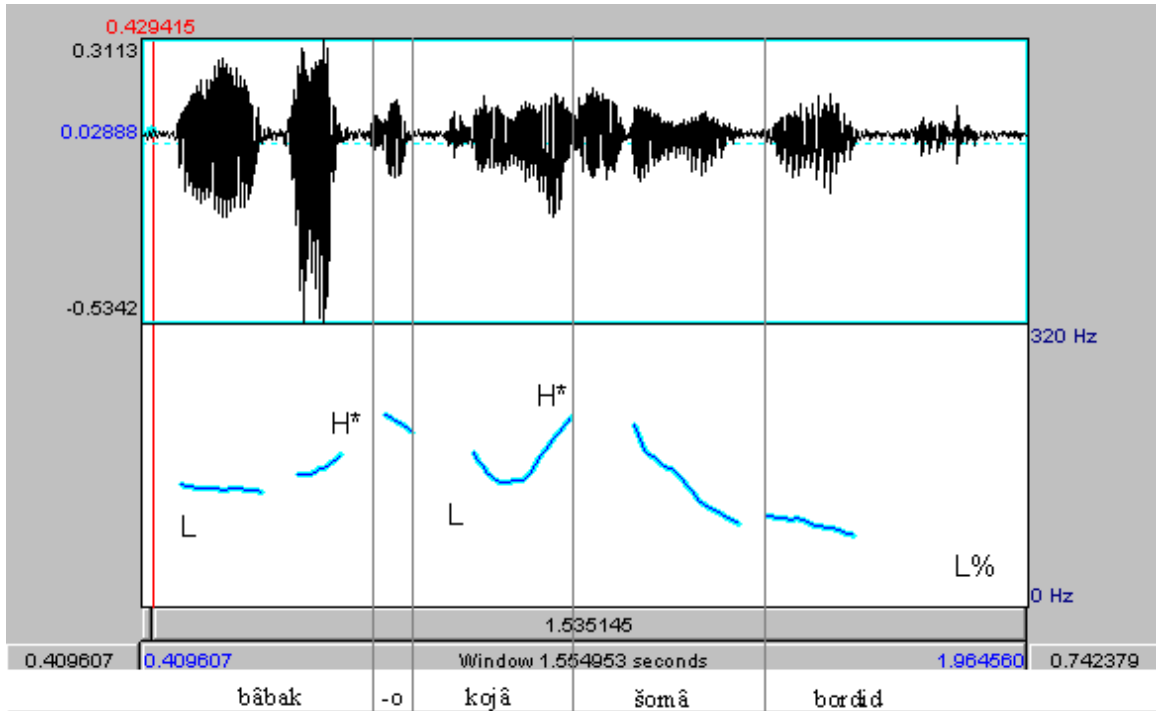   where   you   babak-Acc took-2PL?

Figure 4-10: mechanism of handling focus in a wh-question

What is interesting here is that either focused word or wh-word can cancel out the deaccenting process of the other one on the basis of their position, i.e. whichever occurs later, it will cancel out the deaccenting process of the other one which precedes it. Figure 4-11 confirms this claim of mine quite clearly. If we topicalized the focus (here "bâbak") we can observe clearly how this mechanism is handling the coexistence of these two intonational events.

In this example the focus word "bâbak" precedes the wh-word. What happens is the local cancellation of the deaccenting property of the focus (much more like the previous example). The focus word has its normal prominence thus preserving its L+H* pitch accent as expected. The deaccenting process of focus almost gets started at the Acc marker "râ" (here pronounced as "o"), but soon it gets cancelled by the presence of wh-word. Now here the wh-word becomes prominent and as we expect it deaccent the rest of the phrase. Finally we have the normal low boundary tone, which is typical of such utterances. The process is much the same as the previous example with one difference, i.e. there is no intuitional hint[1] to show the need for a compulsory pause before the second prominent word (here the wh-word). Thus no rephrasing of the utterance is expected[2].

---

[1] By this phrase I mean if a native speaker just uses his/her intuition and listen to this utterance with or without pause in either case it will sound natural, but in the previous example, the utterance sounds quite strange.

[2] That is still we have one IP dominating one iP, not two iPs.

- BĂBAK-o   kojâ   šomâ   bordid? (Where did you take Babak to?)
  babak-Acc where you   took-2PL?

Figure 4-11: The precedence of focus to the wh-word in a wh-question and the mechanism of handling this coexistence

The last example reveals yet another aspect of this complicated phenomenon, i.e. when the subject "šomâ" comes between the focus and the wh-word. What happens here is beyond expectation (at least to my expectation). The focus word has again its normal prominence thus preserving the default L+H* pitch accent as expected. The deaccenting process of focus again gets started at the Acc marker "râ" (here pronounced as "o"), but contrary to what we expect this time it gets cancelled out on the subject "šomâ", i.e. the subject stay out prominent and rejects the deaccenting. In fact it also preserves the default L+H* pitch accent. Now the interesting thing is that this time the deaccenting starts again right at the wh-word, i.e. contrary to our expectation the wh-word becomes deaccented. Finally comes the normal low boundary tone at the end of the phrase. As this is a descriptive work with little space even to handle its own purposes completely, I

tried just to describe it here. For a clear analysis of such phenomena further research, which is more focused, is needed.
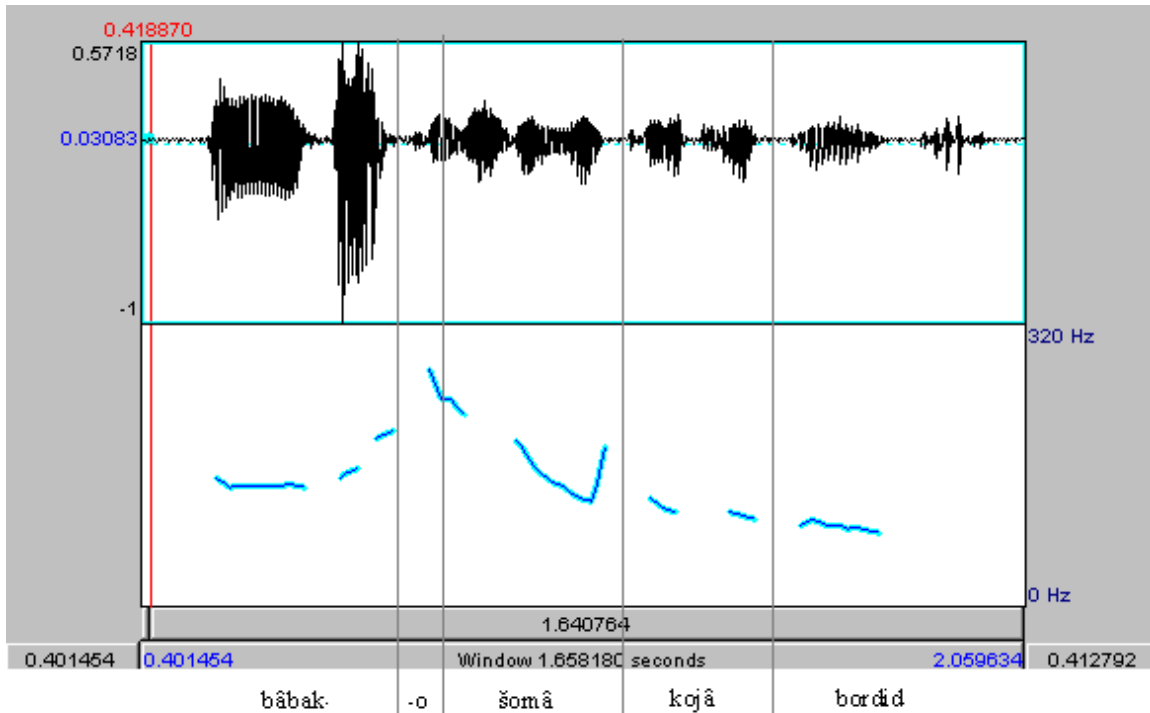
- BĂBAK-o šomâ kojâ   bordid? (Where did you take Babak to?)
 babak-Acc   you   where took-2PL?

Figure 4-12: A more complicated mechanism of treating focus and wh-word in a wh-question

4.3.3 Imperative Sentences

The intention behind using imperative sentences is mostly to order or request some one to do something or to prevent him/her form doing it (either for you, somebody else or for him/herself). You can also use them when you want to give someone directives or advice. It may be expressed as a polite request, sometimes it can be an impolite command, and other times it may take the form of a suggestion. Usually however in daily

social conversations very rarely people use an assertive direct command for communication due to the strong sense of impoliteness that it implies. So if somebody uses an imperative, it usually takes the form of a request, in that case he or she will pick up some polite adverbs like "lotfan" (please), etc. and makes the tone of his/her speech less assertive. But the point is that usually in Farsi a speaker has the option of using some other structures (such as an echo yes/no question) to fulfill this requirement and avoid using imperatives at all.

There were 10 sentences selected and recorded as imperatives in the corpora for this study. Some of them were assertive commands, and some were polite requests. But they are mostly in the form of assertive commands, which quite clearly represent imperative sentences and make them distinguishable from other types of sentences.

Imperatives take a short pitch range, in Farsi their average lower bound f0 is higher than that of neutral declaratives, and also their higher bound f0 is lower than that of declaratives. So all tones distribute within a quite narrow band frequency. Tones are quite flat, and there are not really many falling or rising tones within each AP, however the whole pitch contour gradually and slightly declines, thus giving a smooth and even f0 movement. At the very end of the utterance, however, there is An L tone shortly merging into a lower boundary tone (L%). Usually in imperative utterances speaker's voice starts from a higher pitch right at the beginning of the utterance[1] (figure 4-13)

---

[1] So it seems as if they shared this feature with wh-questions.
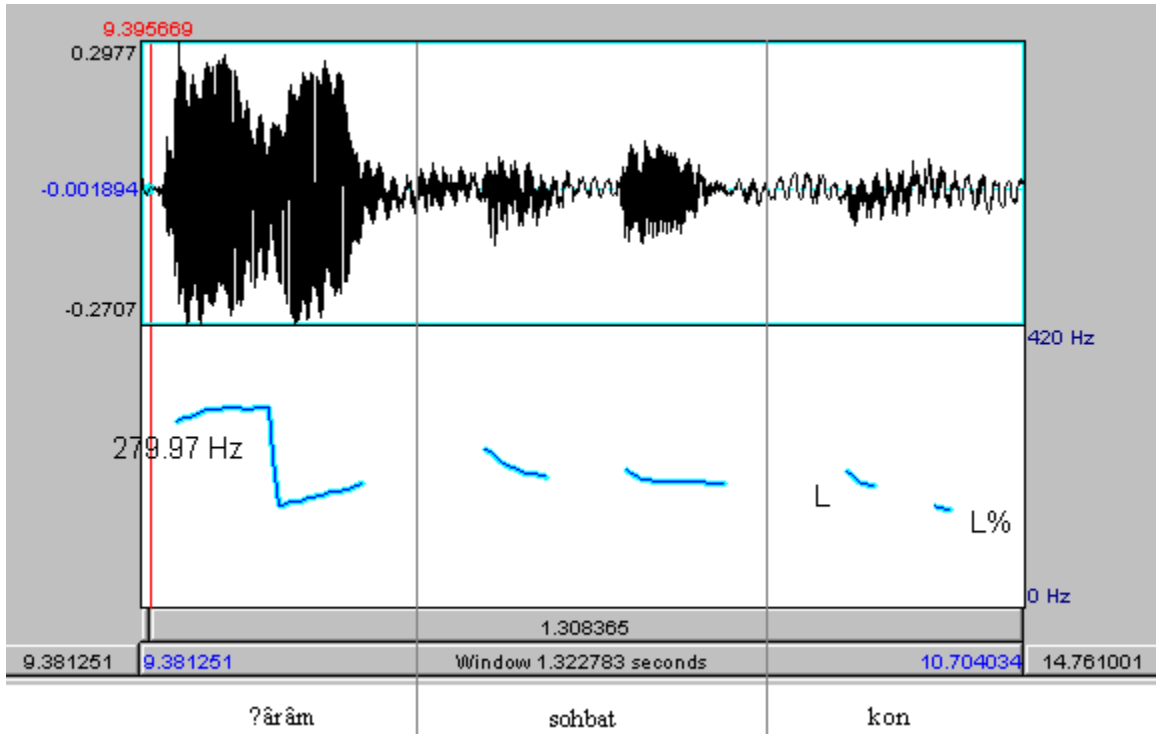
- ?ârâm   sohbat   kon. (Speak slowly.)
  slowly   speech   do.

Figure 4-13: An imperative sentence. It has the typical characteristics: a narrow band frequency, higher starting point f0, and relatively flat pitch contour.

However in polite requests, parts of a Farsi imperative sentence resemble exactly a declarative sentence preserving the default pitch accent pattern. But the verb group preserves the typical structure of imperative as I introduced above[1]. (figure 4-14)

---

[1] Refer to Farsi light verb construction in chapter 1
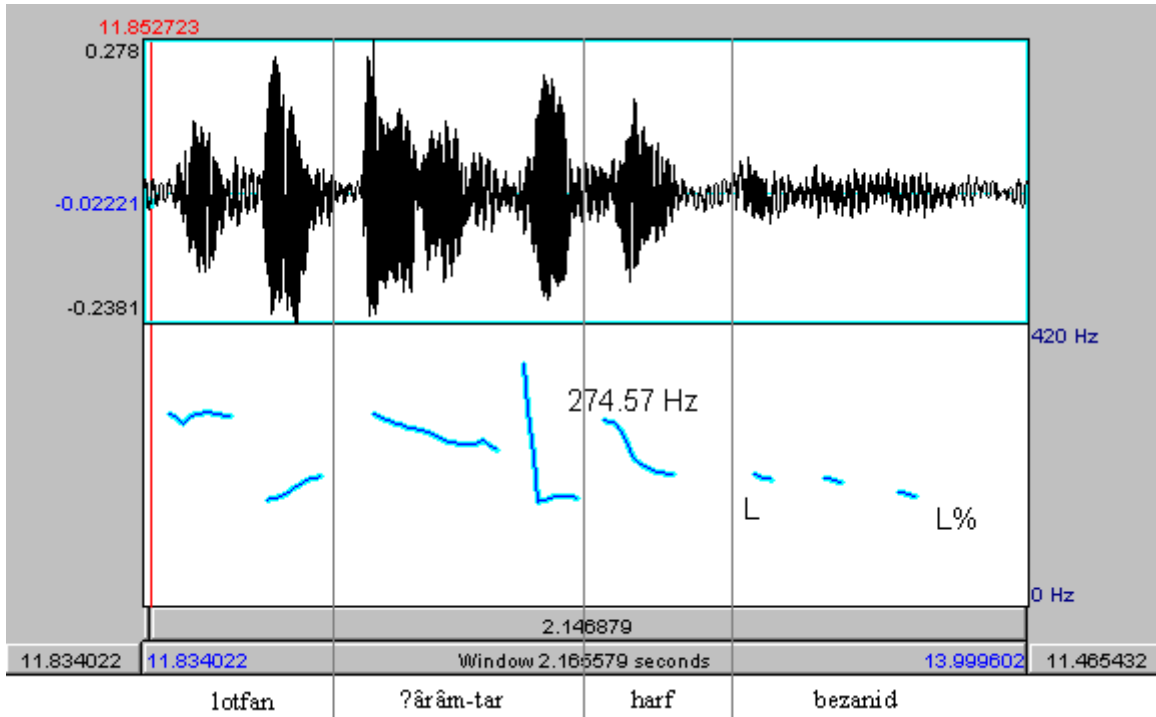
- lotfan   ?ârâm-tar   harf   bezanid. (Please speak more slowly.)
 please   slowly-more   speech   beat.

Figure 4-14: A polite request, only the final part of the utterance (verb group) truly represents the real imperative intonation.

4.3.4 Exclamatory Sentences

In Farsi exclamatory sentences express a wide range of emotions, passions and good or bad feelings like love or hate. That is why such sentences take a wide variety of different forms and structures. This makes them hard to study and any observations regarding their behavior should only cautiously be framed as some general and typical representation of such sentence types.

None of the sentences I prepared in my corpora contained an exclamatory sentence. Due to the factors I mention above, it is difficult to ask people make exclamations just by

seeing some sentences on the paper and putting and exclamation mark at its end.
Authentic data can only be collected in natural real situations when people truly are
impressed by some event and express themselves verbally at once. For analyzing a few
sentences here, I had no opportunity but relying on my own language intuition.

Because of short of space and time, I only try to briefly look at two very common types of
Farsi exclamation sentences namely; declarative exclamations, and wh-exclamations.

4.3.3.1 Declarative Exclamations

In Farsi it is possible to use a declarative sentence to express an exclamation. A general
pitch contour of such structures resembles the following example.(figure 4- 15)
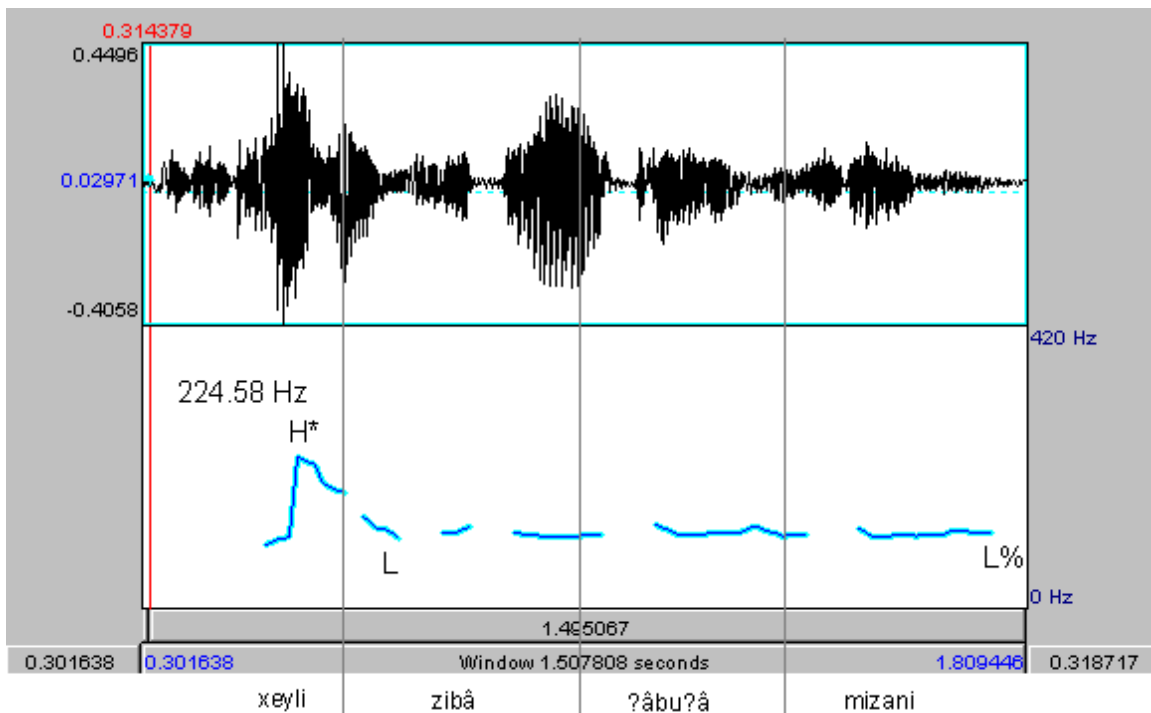

- xeyli   zibâ ?âbu?â   mizani! (You play oboe very nicely!)
  very   nice   oboe   beat-2SG!

Figure 4-15: A declarative exclamatory sentence

The pitch accent of the first AP, which is usually an adverb (this modifies an adjective, another adverb or a nominal which follows it immediately), is realized as an H*, that is the voice of speaker suddenly rises from the lower bound of his/her frequency to the limit of its upper bound (here about 225 Hz). Then there is a deaccenting all through the sentence until the last syllable combines with the L%. Such sort of deaccenting in Farsi declaratives is not a rare phenomenon (we have already seen the case of narrow focusing). But it appears that in other cases, this first AP is always realized as "L+H*" i.e. the default pitch accent in Farsi (in case the focus occurs in the first AP in non exclamatory cases). Duration and the relative value of the frequency of H* is also a good signal (This is relative to individual speakers. No absolute value in this regard can be suggested.). It is worth mentioning that this is only one of the ways an exclamation can be realized within a declarative. My intuition suggests it as the most frequent type. But there could be other variations even in this same example especially if one considers other geographical dialects within Farsi Language.

4.3.3.2 Wh-word Exclamations

Another structure, which Farsi speakers usually use to express their exclamations with, consists of a wh-word, which comes at the very beginning of an utterance. Usually an adverb that modifies the verb of the sentence (like the example below), or a noun phrase follows the wh_word. The general mechanism is like this. Almost never the wh-word gets the prominence, either the adverb or the noun, which follows the wh-word, gets the nucleus pitch accent and the rest of the utterance up to the very end will be deaccented. For example in the following sentence there are two possibilities; either "zibâ" (nice) will get the main pitch accent of the exclamation or the noun "?âlmâni" (German). Figure 4.15 depicts the latter case. (the focused word usually have an exaggerated duration)
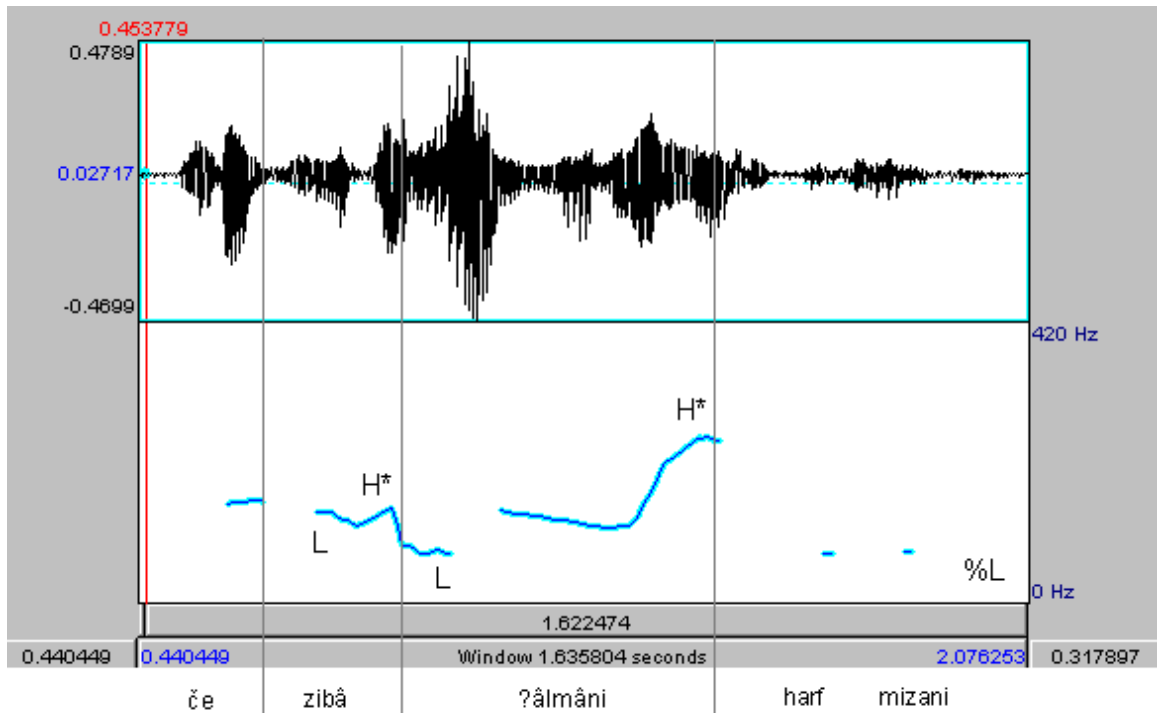
- če zibâ ?âlmâni   harf   mizani! (How nicely you speak German!)
what   nice German speech beat-2SG!

Figure 4-16: A typical wh-word exclamation sentence

Conclusion

This work was one of the first attempts in description of Farsi prosodic and intonation system within the framework of Autosegmental Metrical Phonology. Throughout the dissertation I tried to define the set of basic intonation units in Farsi, and further analyzed their distributions throughout Farsi prosodic phrases. Secondly I attempted to describe various intonation patterns of Farsi utterances in terms of the relevant prosodic features and properties.

The dissertation begins with a short introduction pointing out the aims of the survey, as well as the importance of the subject of study. In chapter one there is a quick review of the Farsi language providing the reader with some key information about the grammar, phoneme inventory, and lexical stress mechanisms. Of the three areas just mentioned the last one receives more attention because of it is highly relevant to the topic of this work. We find out that Farsi has a relatively simple syllable structure that of CV, CVC, and CVCC. Also as far as syllables' sensitivity toward weight is concerned, we can classify Farsi as a weight insensitive language. Later there is a review of various discussions about the superficial diversities in Farsi lexical stress system, a system in which the majority of content words (nouns, adjectives, adverbs and even some verb forms) receive their stress on their last syllable, while verbs violate this apparently well established tendency in a serious way, thus leading to many realizations of stress placement within verb groups. We see that grammar seems to be playing a major role in this phenomenon. Finally this chapter presents a relatively thorough review of the relevant literature written on Farsi prosodic and intonation system.

Chapter 2 is merely a review on the theory of intonational representation, covering firstly some general aspects of prosody and intonation and secondly the phonological theory of prosody. Here a model of intonation representation (the one which originally was proposed by Pierrehumbert) is introduced. It is on the basis of this model that in later

chapters various intonational patterns and prosodic features such as pitch accents, boundary tones, etc. are represented.

Chapter 3 begins with a review of the corpus of study covering issues like selection and modification of data, speaker's specification, recording format, etc. Afterwards throughout the chapter, the main attempt is providing an overview of the principal tonal events in Farsi and its common prosodic constituent structures. Farsi prosodic units namely intonational phrase, intermediate phrase and accentual phrase are mentioned, and their presence in Farsi hierarchical prosodic structure is investigated through some examples. Further in this chapter the tonal inventory of Farsi is examined. This includes a discussion about the default pattern of pitch accents in Farsi utterances (which is L+H*), and a review of Edge tones including the types of boundary tones.

Chapter 4 opens with the investigation of intonational patterns in Farsi. It includes all sentence types namely Declaratives, interrogatives, imperatives and exclamatory sentences. Among them declaratives and interrogatives are paid special attention, as they appear to have the most basic intonational patterns. Also some preliminary investigations are carried out on focus within declaratives and interrogatives.

Considering the short period of time, and the little space for an MSc dissertation, I should admit that this work has only partially achieved its ideal goals, i.e. a thorough study of Farsi prosodic features and Intonation. More carefully selected samples of spontaneous speech are needed. Also the number of speakers should increase so that more accurate and reliable generalizations can be reached at. Furthermore a focused study on the way Intonation conveys pragmatic information in Farsi is particularly valuable as it makes clear how Farsi treats given and new information using its prosodic mechanisms and intonation patterns.

Moreover the effects of word order, which only partially is investigated throughout this work, should receive more attention. As it was mentioned in chapter 1, Farsi is a verb-final language and to some extent the verb-final constraint is strictly observed in the formal registers of the language. In informal registers and daily conversations, however, many instances of utterances with non-canonical word order can be observed. Most of these non-canonical word orders turn up for native speaker as highly grammatical. The

presence of such utterances makes Farsi look like as a free word order language, thus allowing for many different rephrasing of sentences. However more investigation of the prosodic features present in such sentences (within natural spontaneous communication instances) reveals that although grammatically such possibilities are treated all as well formed, they are bound to a limited number of specific intonation realizations. It seems that such limitations are begin enforced by complex mechanisms interacting between syntax and pragmatics. All this and a lot more cries out for more focused surveys and serious studies on Farsi prosody. That is why the author's ambition is to continue on current investigation and probably depict a better picture of the whole issue.

Appendix A: Transliteration System

Farsi Consonants

p   voiceless bilabial stop

b   voiced bilabial stop

m   voiced bilabial nasal

f   voiceless labiodental fricative

v   voiced labiodental fricative

t   voiceless alveolar stop

d   voiced alveolar stop

s   voiceless alveolar fricative

z   voiced alveolar fricative

n   voiced alveolar nasal

r   voiced alveolar liquid

l   voiced alveolar liquid (lateral)

š   voiceless palatal fricative (equivalent to "ʃ")

ž   voiced palatal fricative (equivalent to "ʒ")

č   voiceless palatal affricate (equivalent to "ʧ")

j   voiced palatal affricate (equivalent to "ʤ")

y   voiced palatal glide/approximant (equivalent to "j")

k   voiceless velar stop

g   voiced velar stop

x   voiceless velar fricative

q   voiced uvular stop

?   voiceless glottal stop (equivalent to "ʔ")

h   voiceless glottal fricative

Farsi Vowels

i   high front unround vowel

e   mid front unround vowel

a   low front unround vowel (equivalent to "æ")

u   high back round vowel

o   mid back round vowel

â   low back round vowel (equivalent to "a")

Appendix B: IPA Table of Farsi consonants

	Bilabial	Labio-dental	Dental	Alveolar	Postalv.	Retro-flex	Palatal	Velar	Uvular	Pharyn-geal	Glottal
Plosive	p b			t d				k g	G		ʔ
Nasal	m			n							
Trill				r							
Tap or Flap											
Fricative		f v		s z	ʃ ʒ			x			h
Affricate					tʃ dʒ						
Lateral fricative											
Approximant							j				
Lateral approximant				l							

Bibliography

- Abutalebi Hamid Reza, Bijankhan Mahmood, 2000. "Implementation of a Text-To-Speech System for Farsi Language", Research Center of Intelligent Signal Processing (RCISP), Tehran, Iran.
- Aronoff and M. L. Kean (eds.) Juncture, 107-129. Saratoga, CA: Anma Libri.
- Beckman, M. and J. Pierrehumbert (1986). "Intonational Structure in Japanese and English", Phonology Yearbook 3, 255-309.
- Chodzko, A. 1852. Grammaire persane ou principes de l'iranien moderne, accompagnées de facsimilés pour servir de modele d'écriture et de style pour la correspondence diplomatique et familiere. Paris: Maisonneuve & Cie.
- Cruttenden Alan, 1986. "Intonation", Cambridge: Cambridge University Press.
- Dogil, Grzegorz and Bernd Möbius (2001): "Towards a model of target oriented production of prosody". Proceedings of the European Conference on Speech Communication and Technology (Aalborg, Denmark), vol. 1, 665-668.
- Elell_Sutton L.P. 1976, "The Persian Meters", Cambridge University press, Cambridge.
- Farshad Almas Ganj, 1998. "Analysis of the structure of Farsi Utterances Using through prosodic Information of speech signal", PhD Dissertation, Modares University.
- Ganjavi Shadi et al, "Against Optional Wh-Movement", Proceedings of WECOL 2000, Volume 12, 2000.
- H. Sheikhzadeh, A. Eshkevari, M. Khayatian, R. Sadigh, S. M. Ahadi, 1999. "Farsi Language Prosodic structure, Research and implementation using a Speech Synthesizer", EE Dept., AmirKabir Univ. of Tech., Tehran, Iran.
- Harnsberger James, 1994. "Towards an Intonational Phonology of Hindi", unpublished manuscript [http://www-personal.umich.edu/~jharns/hindi.html].
- Hayes, B. 1995. Metrical Stress Theory: Principles and Case Studies. Chicago: University of Chicago Press.
- Kahnemuyipour Arsalan, 2001. "Unifying Categories: Persian Stress Revisited", First Workshop on Modern Trends in Linguistics. Allameh Tabatabai University, Tehran, Iran. 2001

- Kamyar Taghi Vahidiyan, 1972. "Suprasegmental events", MA Dissertation, Tehran University.
- Karimzadeh Manuchehr, "chehel ghese (Forty stories)", find at the following URL http://www.geocities.com/persianstory/40Story/40Story.html
- Ladd, D. R. 1996. "Intonational Phonology", Cambridge: Cambridge University Press.
- Megerdoomian Karine, 1999. "Persian Syntax, Computing Research Laboratory " available at http://crl.nmsu.edu/Research/Projects/shiraz/ling/syntax.html
- Persian Profile - An academic linguistic profile from UCLA, available at http://www.lmp.ucla.edu/profiles/profp01.htm)
- Pierrehumbert, J., and J. Hirschberg. 1990. "The meaning of intonation contours in the interpretation of discourse" In P. Cohen, J. Morgan, and M. Pollack, Eds., Plans and Intentions in Communication. Cambridge, MA: MIT Press, pp. 271-312.
- Pierrehumbert, J. and M. Beckman. (1988). Japanese Tone Structure. (Linguistic Monograph Series 15).
- Samareh Yadollah, 1986. "Phonology of Farsi Language", Markaz e Nashr e Daneshgahi, Tehran.
- Selkirk E. 1980. "Prosodic Domains in Phonology: Sanskrit Revisited," in M.
- Windfuhr, Gernott. 1990. "Persian", In Comrie, Bernard (ed.), "The world's major Languages", New York: Oxford.
- World languages knowledge fair – Persian, available at http://www.lerc.educ. ubc.ca/ LERC/courses/489/worldlang/persian/description/phonology.htm