DE GRUYTER OPEN

DOI: 10.1515/qal-2017-0002

ISSN: 2299–8942

Intonation of Persian declaratives: Read vs. spontaneous speech

Nima Sadat-Tehrani

ABSTRACT This paper is an introductory investigation, comparing the intonation of Persian declaratives in read and spontaneous speech styles. The results indicate that 32% of the 254 spontaneous declaratives studied show one or more of the following intonational differences: a high or downstepped high tone at Intonational Phrase end, marking the incompleteness of the message; the existence of more pauses leading to a greater number of Intonational Phrases, pre-pause vowel lengthening, and pitch reset; a flatter contour and less pitch variation caused by a speaker's boredom or givenness of the information content; an initial high boundary tone resulted from a low degree of assertiveness.

Keywords: intonation, prosody, Persian, read speech, spontaneous speech, pitch track

1 Introduction

This research reports some preliminary observations regarding the differences between the intonational properties of read and spontaneous speech in Persian. In studies that make use of read speech (also known as lab speech), the speaker is provided with the written version of the test materials and is possibly given some instructions as to how to read them. This type of speech is usually elicited in a laboratory setting and done for a specific linguistic purpose. Spontaneous speech refers to utterances produced without any external stimulus, e.g., a natural conversation carried out between two interlocutors. The aim of this paper is to shed some light on the intonational differences between the two styles. The research on Persian intonation done in the autosegmental-metrical (AM) framework of intonation (Bruce, 1977; Pierrehumbert, 1980; Ladd, 2008) has so far implemented controlled speech (Mahjani, 2003; Jun, et al., 2003; Sadat-Tehrani, 2007; 2009; Taheri and Xu, 2012; Hosseini, 2014; and Rahmani, Rietveld, and Gussenhoven, 2016), and to the best of my knowledge, there has not been any documentation on the distinguishing characteristics of read and spontaneous speech. This paper is an initial investigation of this issue.

The present paper does not intend to argue for the use of spontaneous over read data; although the use of read speech has been criticized as unnatural and unrepresentative of real speech, it has its own merit. For instance, in a laboratory setting, the researcher has full control over many variables and determines their effect on intonation patterns, which is not the case when studying unscripted speech. Also, spontaneous data may not always contain all the structures that the researcher plans to investigate – for more arguments in favour of lab speech see Xu (2010). The view here, in line with that in Wagner, Trouvain, and Zimmerer (2015), is that making use of different types of data while developing a methodological awareness is the path to a more comprehensive and robust theory of intonation. This work is the first step in exploring the changes that take place in both speaking styles in Persian, and hence it suggests that these changes should be taken into consideration when interpreting the results of read speech analysis.

It must also be noted, as Face (2003) states, that the difference between read and spontaneous speech is not necessarily categorical; these are only the two ends of a continuum that contains other styles and techniques as well. Other types of data used in intonation research, with different degrees of spontaneity, include those based on narratives (Grabe, 1998; Bishop, 2002; Tuttle and Lovick, 2007; Ross, 2011), map tasks (Anderson, et al., 1991; Stirling, et al., 2001; Grice and Savino, 2003), guided questionnaires (Prieto and Roseano, 2010), and games (Krahmer and Swerts, 2001; Schafer, Speer, and Warren, 2005; Speer, Warren, and Schafer, 2011) – see Himmelmann and Ladd (2008) and Jun and Fletcher (2014) for a review of different elicitation techniques. The present study looks into the differences that exist between read and spontaneous data, making use of 254 Persian declaratives taken from a corpus of completely spontaneous telephone conversations (Canavan and Zipperlen, 1996) and comparing them to their read version pronounced by Persian native speakers.

The observed intonational differences between different speaking styles include the following characteristics. Generally, in more spontaneous styles, the contour has a lower mean F0 (Morris Haynes, White, and Mattys, 2015; Swerts, Stranger, and Heldner, 1996 for Swedish), is flatter (Tøndering, 2011 for Danish), has more phrase-final rises (Silverman, et al., 1992; Blaauw, 1994; Morris Haynes, White, and Mattys, 2015; Mixdorff and Pfitzinger, 2005 for German), and shows less declination and/or final lowering (Lieberman, et al., 1985; Laan, 1997; de Moraes, 1999 for Brazilian Portuguese; Face, 2003 for Spanish; Fuchs, et al., 2015 for German). Spontaneous style also exhibits more phrase-final lengthening (White, et al., 2010; Markó and Kohári, 2015 for Hungarian) and faster speech rate (Laan, 1997; Furui, et al., 2005 for Japanese; Yeung, et al., 2008 for Mandarin; Dellwo, Leeman, and Kolly, 2015 for Swiss German). There are fewer, shorter, and more regular pauses and fewer disfluencies, e.g., repetitions or deletions, in read speech (Howell and Kadi-Hanifi, 1991; Silverman, et al., 1992; Megyesi and Gustafson-Čapková, 2002 for Swedish; Wang, Li, and Yuan, 2008 for Mandarin).

The present paper does not seek to determine all the differences that exist between read and spontaneous style Persian declaratives; rather, it studies some of the more salient differences in this regard, namely those related to boundary tones, pitch variation, and pause and phrasing. I will leave other issues such as alignment of tones and nuclear accent placement for future studies. The theoretical framework of the paper is the AM theory of intonation, and its organization is as follows. Following this introduction, section 2 provides the reader with some key information on lexical stress and the prosody of declaratives in Persian. Section 3 discusses the observed differences between the speaking styles. Section 4 concludes the paper.

2 Overview of Persian stress and the prosody of read declaratives

I start this section with a brief description of some generalizations regarding Persian lexical stress, to enable the reader to follow the examples of the paper. There is a general tendency for word-final stress in Persian, accompanied by exceptions and deviations (Mahootian, 1997). Most Persian nouns, adjectives, and adverbs are finally stressed (*qæfæsɛ'* ‘cupboard’, *bolænd'* ‘loud’, *hærgɛz'* ‘never’), with some suffixes attracting the stress (*bolænd-tær'* ‘louder’). Verbs have the stress on the last syllable of the main constituent (*xær-id'-æm* buy-PST-1SG ‘I bought’), and verbal prefixes attract the stress (*mi'-xær-id-æm* DUR-buy-PST-1SG ‘I would buy’). Compound verbs – which comprise of a non-verbal element and a verbal element – are stressed on the non-verbal element (*zæng'+zæd-im* bell+hit.PST-1PL ‘We phoned’). Enclitics are stressless (*miz'-ɛt* table-your ‘your table’). For more on Persian stress, see, for example, Kahnemuyipour (2009) and Sadeghi (2017).

Based on the works done on Persian read data in the framework of the AM theory (e.g., Mahjani, 2003; Jun, et al., 2003; Sadat-Tehrani, 2007), the smallest unit of intonation in Persian declaratives is the Accentual Phrase (AP) with the pitch accent L+H* associated with the stressed syllable. The pitch accent is realized as H* for initially-stressed words and monosyllabic content words. An AP is normally composed of a content word plus its possible enclitics, but sometimes in long APs, the L can form a low plateau extending over several unstressed syllables, which can belong to more than one content word. An AP is marked by a high (Ha) or low (La) boundary tone at the right edge.[1] The low boundary tone is used for the nuclear final AP in most simple mono-clausal declaratives, and the pre-nuclear APs all have a high boundary tone. One or more APs make up an Intonational Phrase (IP), which usually corresponds to an utterance for mono-clausal declaratives. An IP is marked by a low or high right boundary tone (L% or H%). Within each IP, the materials following the final AP’s pitch accent undergo deaccentuation in most sentence types (Sadat-Tehrani, 2011; Rahmani, Rietveld, and Gussenhoven, 2016) or post focal compression (Taheri-Aardali and Xu, 2012; Taheri-Ardali, Rahmani, and Xu, 2014). The prosodic structure of Persian declaratives is illustrated in example (1) and Figure 1 for a typical SOV sentence. An accent mark follows the stressed syllable of an AP, and vertical dashed lines demarcate the AP limits. The speech analysis software used is Praat (Boersma and Weenink, 2016).[2]

(1) *mana'* *nun'=o* *bor-id-ɛ'-bud* .
Mana bread=SM cut-PST-PTCP-be.PST.3SG
‘Mana had cut the bread.’

[1] I have previously shown the AP boundary tones with h and l (e.g., Sadat-Tehrani, 2007; 2009), but here I adopt Jun’s (2014) notation as it is more transparent.

[2] Some authors have proposed additional prosodic levels besides AP and IP for Persian, for example, the Intermediate Phrase or the Clitic Group (e.g., Eslami and Bijankhan, 2002 and Abolhasanizadeh, Gussenhoven, and Bijankhan, 2010), but such additional phrasings have been left out of this overview, since the simpler two-level system (IP and AP) suffices for the analysis of this paper. Also, for more on the prosodic status of proclitics and function words see Hosseini (2014).

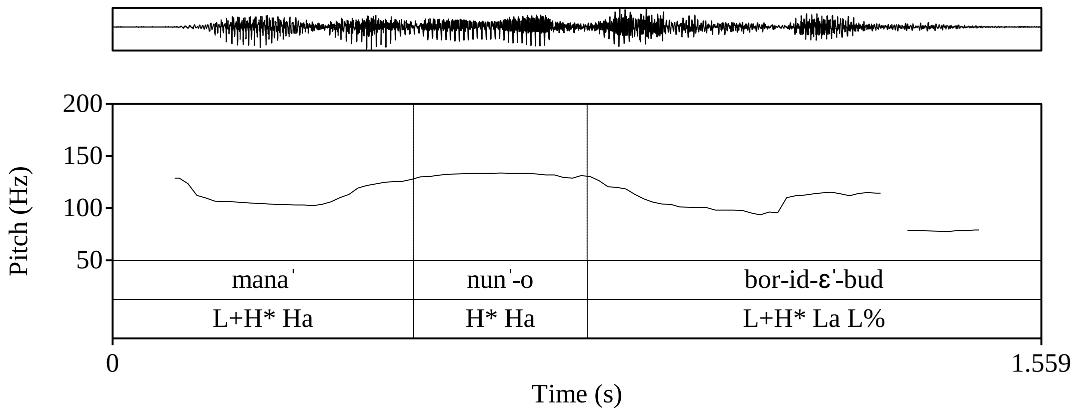

Figure 1: The read utterance *mana' nun'-o bor-id-ɛ'-bud* ‘Mana had cut the bread’

The utterance in (1) contains three APs. The first and the third, i.e., the subject *Mana* and the verb *bor-id-ɛ-bud* ‘had cut’, have a L+H* pitch accent. The mono-syllabic direct object plus its enclitic (*nun=o* ‘the bread’) is stressed on the first syllable and is marked by an H* pitch accent. The first two APs are high-boundary-toned and the last one is low-boundary-toned. The utterance contains one IP ending in a low IP boundary tone (L%).

3 The differences between read and spontaneous declaratives

The data of the present paper are taken from the Persian corpus of Canavan and Zipperlen’s (1996) thoroughly spontaneous telephone conversations. A total of 268 simple mono-clausal declaratives, uttered by 5 interlocutors, 3 females and 2 males, were collected, 14 of which were discarded due to factors such as low audio quality, disfluencies, voice overlaps, etc. The remaining 254 sentences were then read by 4 native speakers of Persian including the author. These speakers were 2 females and 2 males, all university-educated, with an age ranging from early 30s to early 50s, who had grown up as monolingual speakers of Persian. These 9 speakers are referred to in this paper as “spontaneous speakers” and “read speakers” and abbreviated as SF1 (spontaneous, female, 1), RM2 (read, male, 2), etc. The read speakers were requested to read the sentences in a natural way, without any directions given, except in a few cases where the spontaneous version had been pronounced with the nuclear accent placed on an element not nuclear accented by default, in which case that element was underlined and the reader was briefed on it.

The spontaneous and read declaratives were analyzed and labeled by the author, using Praat. One hundred and seventy-three of the spontaneous declaratives (68%) showed no systemic differences with the read version and conformed to the prosodic structure proposed for read data in previous accounts (see section 2 for sources). That is to say, they formed one low-boundary-toned (L%) IP which contained one or more APs with the pitch accent (L+)H* and the AP boundary tone Ha or La. The remaining 81 declaratives (32%) differed from the norm in one way or another. The main differences observed are given in sub-sections from 3.1 to 3.4.

3.1 High-ending declaratives

Thirty-six declaratives (14%) ended in a high tone. Such declaratives were realized in one of the following two ways. Half had the intonation of a pre-conjunction clause in a coordinated structure. In coordinated declarative sentences, the first clause, phrased as a separate IP, ends high (Sadat-Tehrani, 2007), illustrated in (2), where the pre-conjunction clause – i.e., ‘The letter arrived’ – has a high ending, shown in bold type. It is notable that Persian is a null subject language and allows subjects to be dropped, which has resulted in the second clause lacking the overt subject ‘I’.

(2)	L+H*Ha	L+H***HaH%** ...			
	namɛ-hɛ'	*rɛs-id'*	*væ*	*bɛ-hɛʃ'*	*zæng'+zæd-æm.*
	letter-the	arrive-PST.3SG	and	to-him	bell+hit.PST-1SG

‘The letter arrived and I called him.’

(Sadat-Tehrani, 2007, p. 102)

Now consider the example in (3) from the data, and Figures 2a and 2b, which show the spontaneous and read versions of it, respectively.

(3)	*diʃæb'*	*hærf'+zæd...*
	last night	talk+hit.PST.3SG

‘Last night she talked [with them]...’

(Canavan and Zipperlen, 1996, FA-6835, 7:01-7:02)

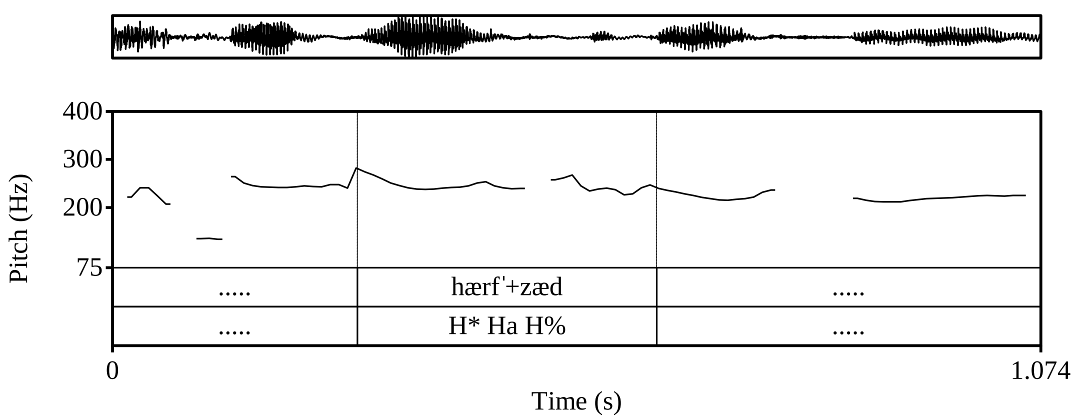

Figure 2a: The spontaneous declarative *diʃæb' hærf'+zæd* ‘Last night she talked’, ending in Ha H%; speaker SF2

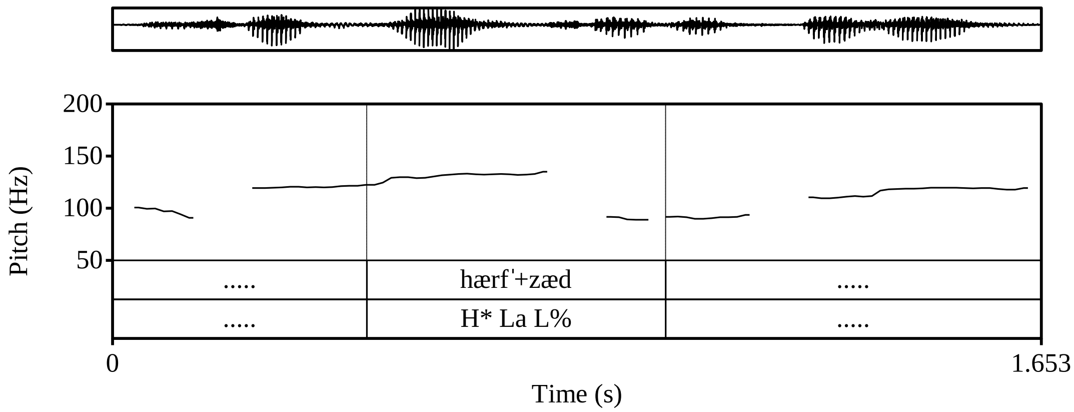

Figure 2b: The read version of *diʃæb' hærf'+zæd...* 'Last night she talked', ending in La L%; speaker RM1

The boundary tones of the spontaneous declarative of Figure 2a (Ha H%) are similar to a pre-conjunction clause, whereas its read pronunciation (Figure 2b) ends in low tones (La L%). The motivation behind the Ha H% sequence is the pragmatic "incompleteness" of such declaratives, which is a characteristic of coordination as well. In these declaratives in the data, the speaker intended to add a complementary idea. For example, (3) was followed by 'Akram is going to come for a visit', and the speaker meant to say that *last night she talked with them and she found out that Akram was going to come for a visit*; therefore, (3) cannot be considered as independent and demands additional information, i.e., the fact that Akram will visit, which has prevented the pitch from lowering.

In the other half of the high-ending declaratives, the nuclear AP boundary tone was realized lower than the previous H*, notated as !Ha, and the declarative ended at the same frequency level (!H%). This lowered AP boundary tone !Ha was not as low as the boundary tone seen in read declaratives (La), which was realized much lower, i.e., in the vicinity of the speaker's bottom range. These downstepped declaratives tended to have more IP final lengthening, followed by a longer pause. Example (4) and Figure 3 are illustrative.

(4) *æmæl'-ɛ* *sisɛkʃɛn'* *daʃt-æm*.
operation-EZ C-section have.PST-1SG
'I had a C-section.'

(Canavan and Zipperlen, 1996, FA-6768, 8:09-8:11)

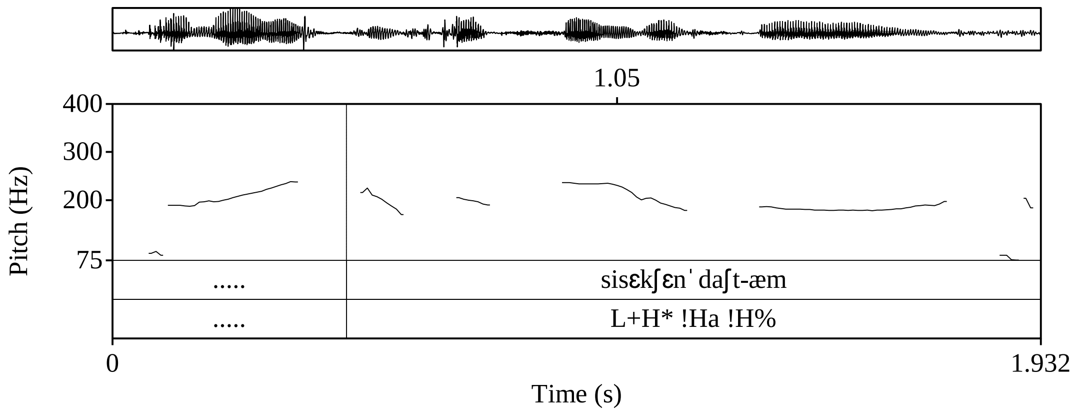

Figure 3: The spontaneous declarative *æmæl'-ɛ sisɛkʃɛn' daʃt-æm* 'I had a C-section', ending in !Ha !H%; speaker SF1

As can be seen, unlike Figure 2a, the pitch starts dropping after the pitch accent L+H* (marked at time 1.05 sec in Figure 3). To confirm that this lowered high tone (Ha) was different from the previous high (H*) tone, statistical analyses were done. Paired samples *t*-tests revealed a significant difference (at 5% level) between the frequency of the downstepped !H (measuring the average frequency over the syllable(s) involved) and its preceding H*. The results are given in (5). The test was carried out five times: for all female speakers combined, for all male speakers combined, and also for each of the speakers SF1, SF2, and SM2 separately (the number of samples for speakers SF3 and SM1 was not large enough for individual statistical analyses, i.e., one analysis for each speaker). Figure 4 shows the boxplot of the frequencies of H* and the downstepped !H for each of the above-mentioned cases. The calculations are done in terms of semitones, with a reference frequency of 100 Hz.

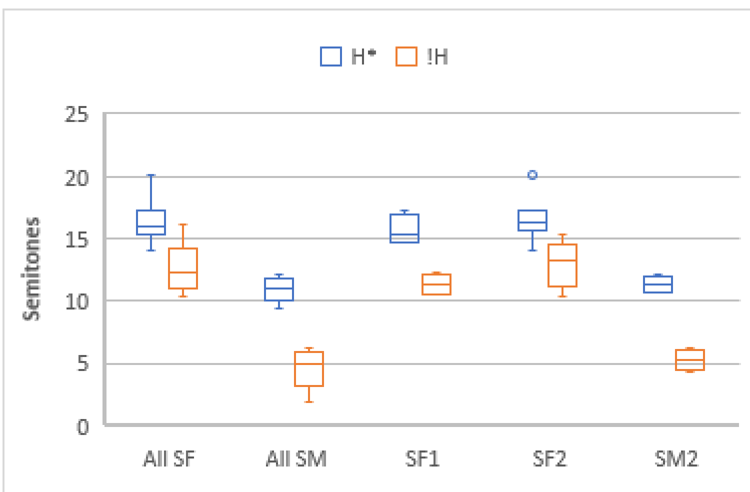

Figure 4: Boxplots of H* and !H frequencies in semitones (reference frequency: 100 Hz) for all spontaneous female speakers, for all spontaneous male speakers, and for each of the speakers SF1, SF2, and SM2. The boxplots display the minimum, the first quartile, the median, the third quartile, and the maximum

(5) All spontaneous female speakers: $t(12) = 11.8$; $p < 0.001$
All spontaneous male speakers: $t(4) = 14.9$; $p < 0.001$
Speaker SF1: $t(3) = 5.6$; $p = 0.01$
Speaker SF2: $t(7) = 9.3$; $p < 0.001$
Speaker SM2: $t(3) = 14.0$; $p < 0.001$

The results are to be interpreted that the downstepped tone is categorially different from its preceding high tone, meaning that speakers lower their pitch from H* to a distinct tone, !H. The pragmatic functions of this lowering and its occurrence in other structures, e.g., polar interrogatives, is currently being studied in another project by the author. For example, it is observed that when this tone occurs in a pre-nuclear AP in a declarative, it can function as a topic marker. The detailed results of this investigation will be shared upon the completion of the study.

A caveat concerning the use of the diacritic '!' for a downstepped tone is in order here. In the history of intonation studies, downstep has been regarded as both a phonetic and a phonological construct. Pierrehumbert (1980) considers a downstepped tone a redundant feature and does not include it in the phonological representation of a tune. Later works give the concept a more underlying status (Ladd, 1983; Beckman and Ayers, 1997). I have adopted the latter view here, which allows me to incorporate !H into the phonology of Persian intonation, and which is in line with the fact that the !H is not a predictable characteristic of the system, but rather a part of the grammar of some native speakers, who use it to convey certain meanings.

3.2 Pause and different phrasing

Twenty-one samples (8%) were uttered with at least one pause in the middle of the declarative (*M* = 204 ms, *SD* = 125 ms). Pauses are one of the acoustic cues of Intonational Phrases (e.g., Ladd, 2008), and a pause-containing declarative was realized as at least two IPs, as opposed to one IP which was the normal read pronunciation of the utterance. Compare the pitch tracks in Figure 5, which contain the spontaneous and read versions of the example in (6).

(6)	*xarɛʤ'*	*æz*	*ʃæhr'*	*bærq'*	*qæt'*	*ʃod-ɛ-bud.*
	out	of	town	electricity	disconnected	become.PST-PTCP-be.PST.3SG

'There was a power cut out of town.'

(Canavan and Zipperlen, 1996, FA-6768, 3:15-3:18)

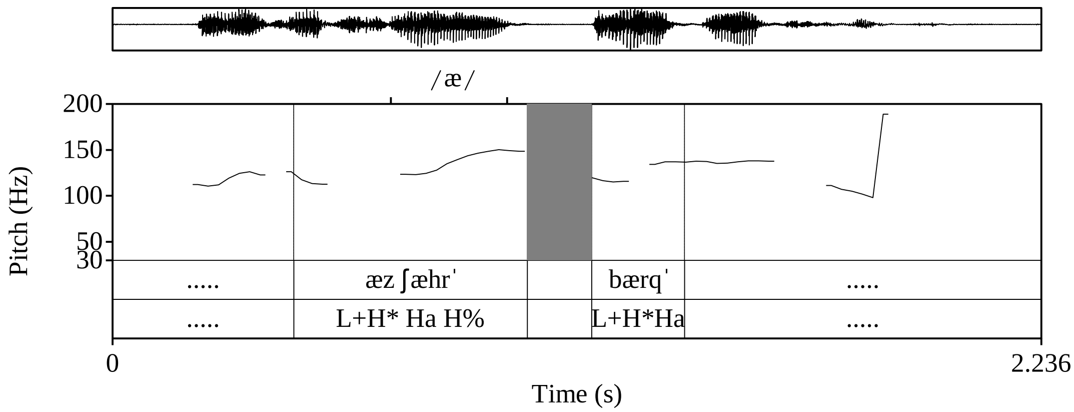

Figure 5a: The spontaneous declarative *xarɛʤ' æz ʃæhr' bærq' qæt' ʃod-ɛ-bud* ‘There was a power cut out of town’, containing a pause (shaded area); speaker SM1

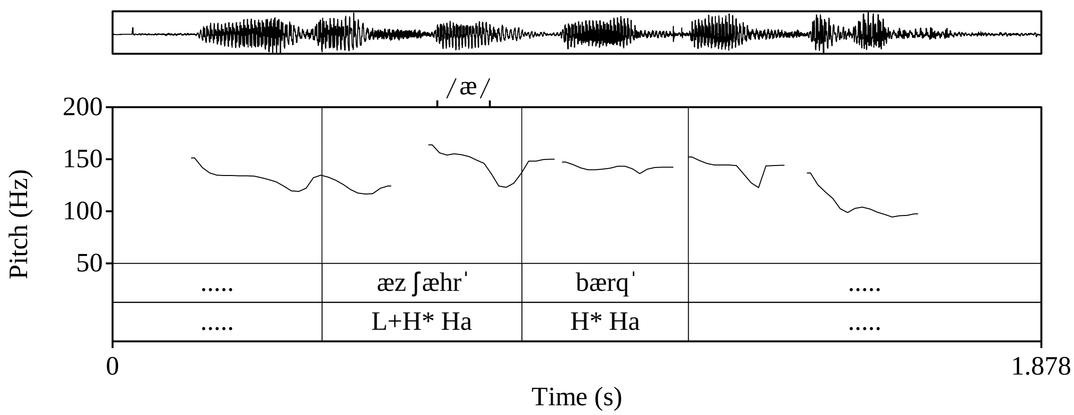

Figure 5b: The read declarative *xarɛʤ' æz ʃæhr' bærq' qæt' ʃod-ɛ-bud* ‘There was a power cut out of town’, without pause; speaker RM2

The spontaneous declarative clause corresponds to two IPs – separated by the pause shaded in the figure – whereas the read counterpart has only one IP.

The existence of pause produced two additional outcomes. The first was pitch resetting, which generated two changes in the tonal representation of the spontaneous utterance compared to the read version. One, an IP boundary tone (H%) was inserted at the end of the first IP (i.e., the end of *xarɛʤ æz ʃæhr* ‘out of town’). Two, the AP immediately following the pause always had the pitch accent L+H*, even when it contained a mono-syllabic content word (recall from section 2 that such words are usually realized as H* and not L+H*). So, in (6), the pitch accent on *bærq* ‘electricity’ is L+H*, and not H*, owing to the new IP-initial status of this word, which generates a rise.

An explanation with respect to the different representations of the two styles in Figures 5a and 5b is in order. Regarding the H% at the end of the first IP, it must be noted that this tone is not due to an underlying phonological distinction between the two utterances, rather, it is brought about by the pause. For example, the same H% could appear in scripted speech too if a sentence is read with a pause. In the same vein, the pitch accent L+H* in the spontaneous utterance is not phonologically distinct from the H* in the read utterance, since these two pitch accents are in allophonic variation in Persian and are realized as one or the other depending on factors such as location in the phrase and lexical stress pattern, without creating a meaning difference (Sadat-Tehrani, 2007). Thus, the argument here is that there are more pauses in spontaneous than read style, and the upshot of a pause is changes in phrasing, hence changes in notation.

The second outcome was that the pre-pause vowel was longer than the same vowel in read pauseless declaratives, which was due to being situated at IP end and having undergone phrase final lengthening. This vowel is shown for Figures 5a and 5b – the vowel /æ/ in the word *ʃæhr* ‘town’. The longer duration of this vowel in spontaneous samples was verified by one-way ANOVAs. The tests were run on the length of the pre-pause vowel as the dependent variable, and spontaneous and read conditions as independent variables. The durations were normalized in terms of the speaker’s utterance duration in order to control for speech rate. The ANOVA was done two times. Once, the analysis involved all the declaratives of the four spontaneous speakers combined compared with the production of each of the four read speakers of the same declaratives. The other comparison was made between the declaratives uttered by SM1 and the corresponding declaratives of each of the read speakers RF1, RF2, RM1, and RM2. The reason for including only one spontaneous speaker (SM1) was the small sample size for the others spontaneous speakers (two, one, three, and one for SF1, SF2, SF3, and SM2, respectively). The boxplots of the pre-pause vowel durations are given in Figure 6 and the summary results of the ANOVAs appear in (7).

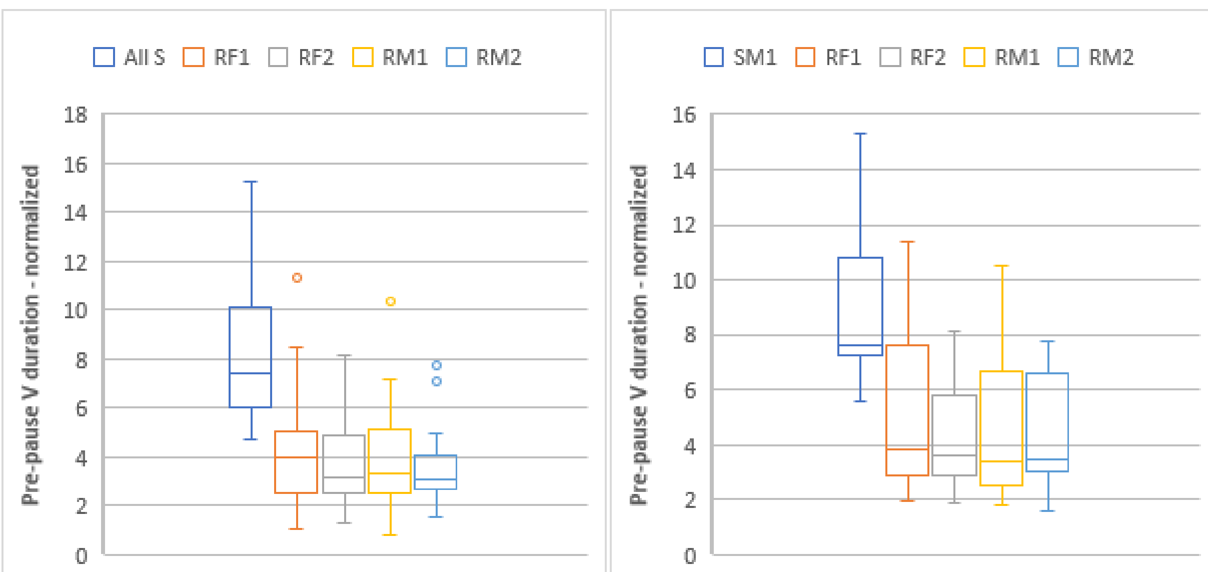

Figure 6: Boxplots of the normalized pre-pause vowel duration for all spontaneous speakers compared to each of the read speakers (left) and the declaratives of spontaneous speaker SM1 alone compared to the same declaratives by each of the read speakers (right)

(7) a. All spontaneous speakers
$F(4,100) = 14.7$; $p < 0.001$

b. Speaker SM1
$F(4,55) = 6.9$; $p < 0.001$

A significant difference in both cases was observed. Scheffé *post hoc* tests determined that the spontaneous samples were a different group and that the read speakers were not different from one another.

To further confirm the extra length of the pre-pause vowel, its duration (normalized) was compared with its preceding vowel (the penult vowel) in spontaneous speakers' productions by *t*-tests. The test was done three times, once for all five speakers combined, once for speaker SM1, and once for speakers SF1, SF2, SF3, and SM2 combined (as before, individual analysis for these latter speakers was not done due to their small sample size). The results are given in (8), accompanying the related boxplots in Figure 7.

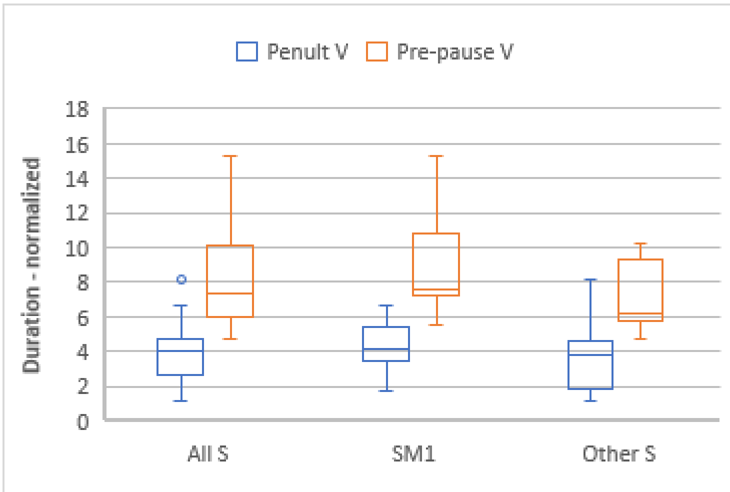

Figure 7: Boxplots of the normalized duration of the pre-pause vowel and its preceding vowel (penult V) for all spontaneous speakers, for SM1, and for the other four spontaneous speakers

(8) All spontaneous speakers: $t(20) = 6.3$; $p < 0.001$
Speaker SM1: $t(11) = 4.9$; $p < 0.001$
Speakers SF1, SF2, SF3, SM2: $t(8) = 4.1$; $p = 0.003$

In sum, the results suggest that the duration of the vowel before a pause in spontaneous speech undergoes significant lengthening, which is due to the realization of the pre-pause chunk as a separate IP.

3.3 Flatter contours

The contour of 6% of the data (15 utterances) were noticeably flatter and had lowered peaks. The example in (9) with its pitch track in Figure 8 is illustrative.

(9)	*halæt'-ɛ*	*bohran'*	*amæd-i'+birun'.*
	condition-EZ	crisis	come.PST-2SG+out

'You came out of the critical condition.'

(Canavan and Zipperlen, 1996, FA-6768, 25:56-25:58)

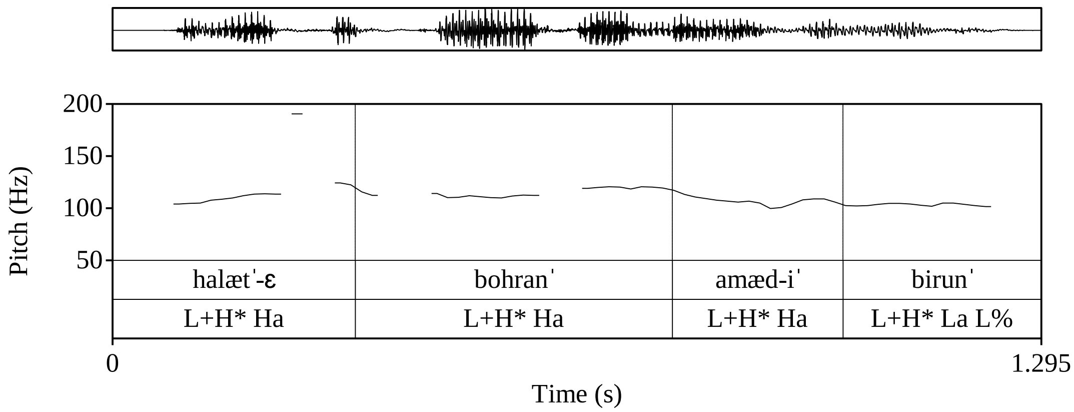

Figure 8: The spontaneous declarative *halæt'-ɛ bohrán amæd-i' birun'* 'You came out of the critical condition', flatter contour; speaker SM1

This degree of flatness was not seen in any of the read utterances, which normally exhibited more variation in pitch. Furthermore, a speaker's flatter contours were visibly quite different from his/her other contours. For example, the pitch range of the above utterance is about 21 Hz (123-102) while the normal pitch range of this specific speaker is around twice this amount.

What could be the motivation behind such contours? A potential cause for a flatter contour – or less declination – is an increase in utterance length (e.g., Ohala, Dunn, and Strouse, 2004; Yuan and Liberman, 2014). To check this, Mann-Whitney *U* tests were performed on the means of word count per utterance between flatter and normal contours. Since there was a large difference in the sample sizes (n_1 = 15, n_2 = 218), the test was done several times; once for all samples (n_1 = 15, n_2 = 218), and three times for equal size samples based on randomly choosing 15 utterances from the non-flat contours (n_1 = 15, n_2 = 15). The boxplots are given in Figure 9, and Mann-Whitney *U* test results appear in (10).[3]

[3] The reason for using the Mann-Whitney *U* test is the non-normal distribution of the second variable. The paused samples were excluded from this analysis.

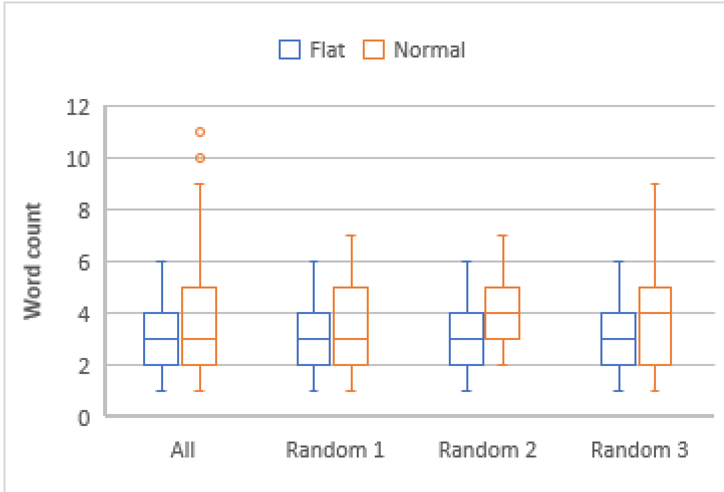

Figure 9: Boxplots of word counts in normal and flatter contours, for all samples and for three random sample sets of equal sizes

(10) All samples: $U = 1334$, $n_1 = 15$, $n_2 = 218$, $p = 0.19$
Random set 1: $U = 104$, $n_1 = 15$, $n_2 = 15$, $p = 0.75$
Random set 2: $U = 84$, $n_1 = 15$, $n_2 = 15$, $p = 0.24$
Random set 3: $U = 83$, $n_1 = 15$, $n_2 = 15$, $p = 0.23$

As can be seen, the word count per utterance in flatter and ordinary contours is not significantly different, so it is not the reason for the flatness of the contour. Instead, a few other factors seem to be relevant here. First, research has indicated that prepared speech generally involves more anticipatory speech planning (e.g., Fuchs, et al., 2015), and in this way, when speakers read from a script, there is more in-advance processing at work, which could lead to more declination. Therefore, this could be one of the reasons for the smaller amount of declination in these spontaneous declaratives. Second, a flatter contour in the data was linked to the attitude of boredom on the part of the speaker.[4] This observation reflects the impact of attitude on pitch, which is a well-known fact (e.g., Bänzinger and Scherer, 2005) and which is reported for many languages. The third reason concerns the link between pitch and information structure. Researchers have suggested that information packaging and degree of givenness interact with pitch (Brown, 1983; Steedman, 2000; Baumann and Grice, 2006). The utterance in (9) above is an example of this relation. After a conversation about her health problems, a speaker says that *now she's feeling better and the worst is behind her*. Then in response, her interlocutor remarks that *it's good that she is feeling better* and adds the utterance in (9) ('You came out of the critical condition'), which is basically a paraphrase of what is already mentioned in the discourse and so is not considered completely new information. As a result, the declarative is uttered with lowered peaks.

[4] This is a preliminary consideration as to the meaning assigned to this contour, and controlled perceptual tests are eventually required to provide more solid grounds for the semantic connection. This point is valid for the meaning of the initial high contour of section 3.4, too.

3.4 The initial high tone

Nine utterances (4%) started with a high tone. Following the original ToBI notation (Beckman and Ayers, 1997), I label this tone %H and treat it as an initial boundary tone, including it in the phonological representation. Alternatively, this initial high tone might be dealt with as an output of phonetic implementation rules, in which case it would normally not be reflected in the AM representation. I chose the former based on the fact that this tone was systematically employed by speakers for a specific purpose (to be discussed shortly). The boundary nature of this tone stems from its not being associated with an accented syllable, in which case it would have been notated with the pitch accent H*. Example (11) is an utterance containing the initial high tone. Figures 10a and 10b illustrate the spontaneous and read pronunciations of the utterance, respectively.

(11) *bædʃansi'-yɛ ma' bud.*
bad luck-EZ we be.PST.3SG
'It was our bad luck.'
(Canavan and Zipperlen, 1996, FA-6835, 18:19-18:21)

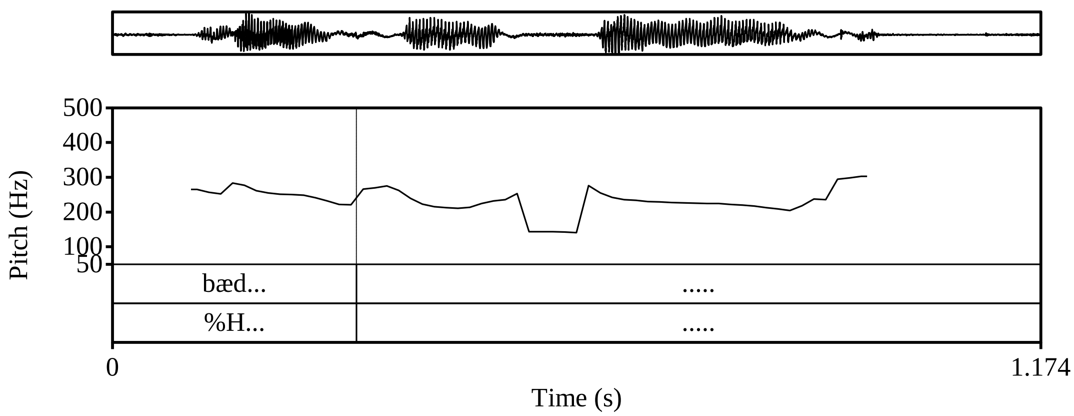

Figure 10a: The spontaneous declarative *bædʃansi'-yɛ ma' bud* 'It was our bad luck', initial %H; speaker SF3

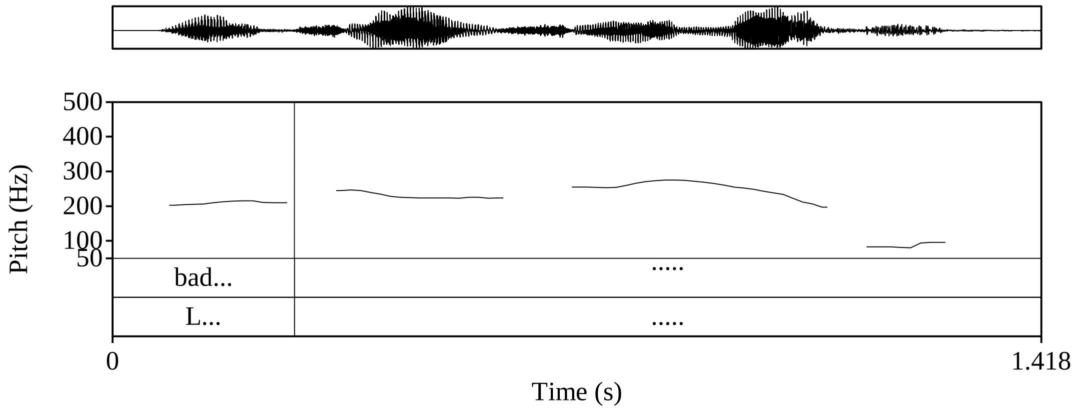

Figure 10b: The read declarative *bædʃansi'-yɛ ma' bud* ‘It was our bad luck’, starting with an L tone; speaker RF2

In Figure 10a, the utterance begins with a high tone associated with the unstressed syllable *bæd-*, while the same syllable in the read pronunciation (10b) carries the low tone of a normal L+H* pitch accent.

Let us reflect on what the driving force behind the production of this tone might be. One possibility to look at is its connection with cognitive preplanning of speech. As mentioned in the previous section, researchers believe that speakers plan speech in advance, which may be a factor connected with declination. Thus, an initial high tone could be a strategy that speakers use in order to facilitate declination in the course of an utterance; however, this is implausible in our case since the utterances in the data which had an initial high tone did not exhibit a substantial amount of declination. A related point is that researchers have suggested that the pitch level at the onset of an utterance could be positively correlated with the utterance length (e.g., Rialland, 2001). To investigate this possibility, Mann-Whitney *U* tests were carried out on the means of word count per utterance between high starting utterances and other utterances. As in section 3.3, due to the large difference between sample sizes, the test was done several times, once with all samples, and three times with equal size randomly selected samples. The results appear in (12) accompanied by comparative boxplots in Figure 11.

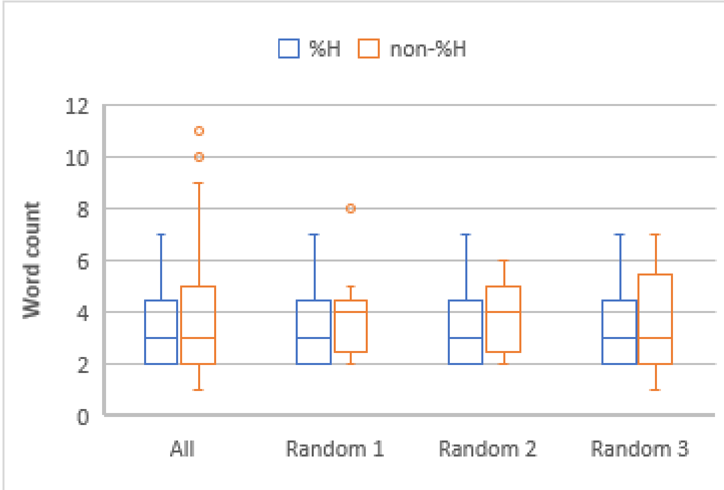

Figure 11: Boxplots of word counts in initial high and non-initial high utterances, for all samples and for three random sample sets of equal sizes

(12) All samples: $U = 959$, $n_1 = 9$, $n_2 = 224$, $p = 0.74$
Random set 1: $U = 36$, $n_1 = 9$, $n_2 = 9$, $p = 0.69$
Random set 2: $U = 35$, $n_1 = 9$, $n_2 = 9$, $p = 0.66$
Random set 3: $U = 40$, $n_1 = 9$, $n_2 = 9$, $p = 1.00$

As the results indicate, the utterances that start high are not longer than others, so utterance length is not inducing the %H tone.

Another potential cause for high pitches to be considered is segmental effects; for instance, high vowels or vowels following a voiceless consonant are usually higher pitched (Löfqvist 1989; Gussenhoven, 2004). This factor can also be ruled out as there is a variety of categorically different initial segments in these utterances, namely voiced plosives (/b/, /d/), fricatives (/s/, /h/), sonorant consonants (/r/, /m/), and a low vowel (/æ/). Notably, out of the nine samples with an initial high, only three started with a voiceless segment (and none with a high vowel); in addition, there were many other utterances beginning with the same voiceless segments that did not start as high as these.

The initial high tone in the data seems to be pragmatically motivated. Initial highs have been stated to serve different purposes: expression of surprise/exclamation (Sag and Liberman, 1975; Grice, et al., 2005 for Florentine Italian), topic refreshment (Gussenhoven, 2002), making requests (Frota, 2014, for European Portuguese), and asking certain yes/no questions (Prieto, 2014 for Catalan); however, in the Persian data, the initial %H tone mostly signifies some sort of "lack of importance" of the propositional content conveyed by the utterance and decreases its "assertiveness." For instance, in the dialogue leading to the above example, one speaker says that she is sorry that she has moved from a bigger apartment to a smaller one and, as a result, she cannot have the other speaker and her family over as guests. The other speaker replies that *it's OK and it's not your fault*, adding the utterance in (11) above ('It was our bad luck') meaning *let's just say it was me and my family's bad luck and forget about it so that you didn't feel guilty*. Another example of such an effect appears in (13).

(13) *maʃin'=æm=æm hala' rah'+mi'-r-ɛ.*
car=my=too now way+DUR-go.PRS-3SG
'My car is still running well.'
(Canavan and Zipperlen, 1996, FA-6835, 10:04-10:05)

The context is that the speaker's car is old, a fact known by both interlocutors, and the utterance in (13) is used as an answer to the question 'How's your car?' and means that *although my car is not in excellent condition, it's still running OK*, but in reality, the car is not in perfect condition, and hence the utterance is not a complete assertion of the fact that the car is running well.

4 Conclusion

A preliminary investigation was made into the intonational differences between read and spontaneous declaratives in Persian. Out of the 254 spontaneous samples studied, 32% were intonationally different from their read counterparts. Table 1 summarizes the findings of this paper.

Table 1: Summary of findings

		Characteristic	
Total number of spontaneous declaratives studied: 254	Similar to the read counterparts: 173 (68%)	(L+)H* La L%	
	Different from the read counterparts: 81 (32%)	Ending high: 36 (14%)	Ha 18 (7%)
			!Ha 18 (7%)
		Paused 21 (8%)	
		Flatter contour 15 (6%)	
		Initial %H 9 (4%)	

The spontaneous declaratives were characterized by one or more of the following differences. First, some spontaneous samples ended on a high pitch – either a high or a downstepped high boundary tone – as opposed to the normal low of a read utterance, which denoted the incompleteness of the message conveyed by the declarative and which indicated the intention of the speaker to continue. This comparatively high finish, although not involving a rise *per se*, parallels the findings of, e.g., Mixdorff and Pfitzinger (2005) for German, and Morris Haynes, White, and Mattys (2015). In these works, more final rises, as opposed to falls, were found in spontaneous speech map tasks, which served different purposes including establishing contact, topic continuation, and asking for feedback. The continuation function was observed in Persian as well, which acted as a non-finality signal. The higher number of high-pitch-ending phrases in spontaneous speech reflects a more communicative style, as opposed to read speech, which often involves more independent separate utterances.

Second, there were more pauses in spontaneous style, which is consistent with previous observations (e.g., Howell and Kadi-Hanifi, 1991; Silverman, et al., 1992). The unplanned nature of spontaneous speech increases the number of pauses in this speech mode, many of which occur in unexpected locations such as inside syntactic phrases, for instance, inside a noun phrase between a noun and its modifier (*tufan-ɛ [PAUSE] ʃædid* storm-EZ severe 'severe storm'). Those declaratives that were uttered with pause were realized as more than one IP. This resulted in the insertion of an IP boundary tone at the end of the first IP, and also caused the initial AP in the second IP to always start with a rise. In addition, the pause created the lengthening of the pre-pause vowel.

Third, some spontaneous contours were flatter and showed less declination, which provides converging evidence to previous findings, e.g., Laan (1997), de Moraes (1999) for Brazilian Portuguese, and Tøndering (2011) for Danish. This is explained by the fact that less planning is involved in spontaneous speech. The flatness of the contour was verified to be independent of utterance length and was motivated by the attitude of boredom on the part of the speaker or a lower degree of newness of the information content.

Fourth, a few of the spontaneous declaratives were marked by an initial high tone, not associated with any stressed syllable. This tone was not correlated with utterance length, nor was it impacted by segmental influence, e.g., by voiceless consonants. Rather, the production of the tone seemed to be driven by pragmatic factors such as indication of less assertiveness in the declarative and attachment of less importance to the message.

Abbreviations

AP = accentual phrase, DUR = durative, EZ = the ezafe particle, IP = intonational phrase, PRS = present, PST = past, PTCP = participle, SG = singular, SM = specificity marker; the equal sign (=) in the examples marks an enclitic boundary; the plus sign (+) separates the two parts of a compound verb

References

Abolhasanizadeh, V., Gussenhoven, C., and Bijankhan, M. (2010). The position of clitics in Persian intonational structure. *Speech Prosody 2010, 100108,* 1-4. Retrieved from http://gep.ruhosting.nl/carlos/speech_prosody_2010_zadeh_gussenhoven.pdf

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., ..., and Sotillo, C. (1991). The HCRC map task corpus. *Language and Speech, 34*(4), 351-366. doi: 10.3115/1075671.1075677

Bänzinger, T. and Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication, 46,* 252-267. doi: 10.1016/j.specom.2005.02.016

Baumann, S. and Grice, M. (2006). The intonation of accessibility. *Journal of Pragmatics, 38*(10), 1636-1657. doi: 10.1016/j.pragma.2005.03.017

Beckman, M. E. and Ayers, G. M. (1997). *Guidelines for ToBI labelling, version 3.* Online manuscript available at: http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/

Bishop, J. B. (2002). *Aspects of intonation and prosody in Bininj gun-wok: Autosegmental-metrical analysis.* (Doctoral dissertation). University of Melbourne, Melbourne.

Blaauw, E. (1994). The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech. *Speech Communication, 14*(4), 359-375. doi: 10.1016/0167-6393(94)90028-0

Boersma, P. and Weenink, D. (2016). Praat: Doing phonetics by computer [Computer program]. Version 6.0.19. Retrieved 14 June 2016 from http://www.praat.org/

Brown, G. (1983). Prosodic structure and the given/new distinction. In A. Cutler and D. R. Ladd (Eds.), *Prosody: Models and measurements* (pp. 67-77). Berlin: Springer. doi: 10.1007/978-3-642-69103-4_6

Bruce, G. (1977). *Swedish word accents in sentence perspective.* Lund: Gleerup.

Canavan, A. and Zipperlen, G. (1996). CALLFRIEND Farsi. Linguistic Data Consortium, Philadelphia.

Dellwo, V., Leemann, A., and Kolly, M.-J. (2015). The recognition of read and spontaneous speech in local vernacular: The case of Zurich German. *Journal of Phonetics*, *48*, 13-28. doi: 10.1016/j.wocn.2014.10.011

Eslami, M. and Bijankhan, M. (2002). NEzam-E ahæng-E zæban-E farsi [Persian intonation system]. *Iranian Journal of Linguistics, 34,* 36-61.

Face, T. (2003). Intonation in Spanish declaratives: Differences between lab speech and spontaneous speech. *Catalan Journal of Linguistics, 2*, 115-131.

Fletcher, J., Grabe, E., and Warren, P. (2005). Intonational variation in four dialects of English: the high rising tune. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 390-409). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199249633.003.0014

Frota, S. (2014). The intonational phonology of European Portuguese. In S.-A. Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing* (pp. 6-42). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199567300.003.0002

Fuchs, S., Petrone, C., Rochet-Capellan, A., Reichel, U. D., and Koenig, L. L. (2015). Assessing respiratory contributions to f0 declination in German across varying speech tasks and respiratory demands. *Journal of Phonetics 52*, 35-45. doi: 10.1016/j.wocn.2015.04.002

Furui, S., Nakamura, M., Ichiba, T., and Iwano, K. (2005). Analysis and recognition of spontaneous speech using Corpus of Spontaneous Japanese. *Speech Communication, 47*(1), 208-219. doi: 10.1016/j.specom.2005.02.010

Grabe, E. (1998). *Comparative intonational phonology: English and German*. Nijmegen: Max Planck Institute.

Grice, M., D'Imperio, M., Savino, M., and Avesani, C. (2005). Strategies for intonation labeling across varieties of Italian. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 362-389). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199249633.003.0013

Grice, M. and Savino, M. (2003). Map tasks in Italian: Asking questions about given, accessible and new information. *Catalan Journal of Linguistics, 2*, 153-180.

Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and phonology. *Speech Prosody 2002, Aix-en-Provence, France* (pp. 47-57). Retrieved from http://gep.ruhosting.nl/carlos/aix2002guss.pdf

Gussenhoven, C. (2004). *The phonology of tone and intonation.* Cambridge: Cambridge University Press. doi: 10.1017/cbo9780511616983

Himmelmann, N. P. and Ladd, D. R. (2008). Prosodic description: An introduction for field workers. *Language Documentation & Conservation*, *2*(2), 244-274.

Hosseini, A. (2014). *The phonology and phonetics of prosodic prominence in Persian.* (Doctoral dissertation). Tokyo University of Foreign Studies, Tokyo.

Howell, P. and Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication*, *10*, 163-169. doi: 10.1016/0167-6393(91)90039-v

Jun, S.-A. (Ed.). (2014). *Prosodic typology II: The phonology of intonation and phrasing.* Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199567300.001.0001

Jun, S.-A. and Fletcher, J. (2014). Methodology of studying intonation: From data collection to data analysis. In S.-A. Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing* (pp. 493-519). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199567300.003.0016

Jun, S.-A., Scarborough, R., Arbisi-Kelm, T., Esposito, C. M., and Barjam, P. (2003). *Intonational phonology of Farsi.* (Unpublished manuscript). UCLA, Los Angeles.

Kahnemuyipour, A. (2009). The syntax of sentential stress. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199219230.001.0001

Krahmer, E. and Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication, 34,* 391-405. doi: 10.1016/s0167-6393(00)00058-3

Laan, G. P. (1997). The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication*, *22*(1), 43-65. doi: 10.1016/S0167-6393(97)00012-5

Ladd, D. R. (1983). Phonological features of intonational peaks. *Language*, *59*, 721-759. doi: 10.2307/413371

Ladd, D. R. (2008). *Intonational phonology.* (Second Edition). Cambridge: Cambridge University Press. doi: 10.1017/cbo9780511808814

Lieberman, P., Katz, W., Jongman, A., Zimmerman, R., and Miller, M. (1985). Measures of the sentence intonation of read and spontaneous speech in American English. *The Journal of the Acoustical Society of America*, *77*(2), 649-657. doi: 10.1121/1.391883

Löfqvist, A., Baer, T., McGarr, N. S., and Story, R. S. (1989). The cricothyroid muscle in voicing control. *The Journal of the Acoustical Society of America*, *85*(3), 1314-1321.

Mahjani, B. (2003). *An instrumental study of prosodic features and intonation in Modern Farsi (Persian).* (Unpublished Master's thesis). University of Edinburgh, Edinburgh.

Mahootian, S. (1997). *Persian: Descriptive grammars.* London: Routledge. doi: 10.4324/9780203192887

Markó, A. and Kohári, A. (2015). Glottalization and timing at utterance final position in Hungarian: Reading aloud vs. spontaneous speech. In the Scottish Consortium for ICPhS

2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, paper 722. Glasgow, UK: The University of Glasgow. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0722.pdf

Megyesi, B. and Gustafson-Čapková, S. (2002). Production and perception of pauses and their linguistic context in read and spontaneous speech in Swedish. *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP), Denver, USA* (pp. 2153-2156). Retrieved from https://pdfs.semanticscholar.org/7d79/389b1a6882ee86240f5a6481461fa1d4e243.pdf

Mixdorff, H. and Pfitzinger, H. R. (2005). Analysing fundamental frequency contours and local speech rate in map task dialogs. *Speech Communication*, *46*, 310-325. doi: 10.1016/j.specom.2005.02.019

Moraes, J. A. de (1999). F0 declination in Brazilian Portuguese in read and spontaneous speech. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)* (pp. 2323-2326). San Francisco, USA: The Regents of the University of California. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/papers/p14_2323.pdf

Morris Haynes, R., White, L., and Mattys, S. L. (2015). What do we expect spontaneous speech to sound like? In the Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, paper 1011. Glasgow, UK: The University of Glasgow. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS1011.pdf

Ohala, J. J., Dunn, A., and Sprouse, R. (2004). Prosody and phonology. *Speech Prosody 2004, Nara, Japan* (pp. 161-163), retrieved from https://pdfs.semanticscholar.org/af63/4e68a1534a750cdba4fa5a34bcf55b629159.pdf

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation.* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA. Retrieved from http://faculty.wcas.northwestern.edu/~jbp/publications.html

Prieto, P. (2014). The intonational phonology of Catalan. In S.-A. Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing* (pp. 43-80). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199567300.003.0003

Prieto, P. and Roseano, P. (2010). Transcription of intonation of the Spanish language. *Lincom Studies in Phonetics.* Munich, Lincom Europa.

Rahmani, H., Rietveld, T., and Gussenhoven, C. (2016). Persian word accent is deletable. *Speech Prosody 2016, Boston, USA* (pp. 440-444). doi: 10.21437/speechprosody.2016-90

Rialland, A. (2001). Anticipatory raising in downstep realization: Evidence for preplanning in tone production. In S. Kaji (Ed.), *Proceedings of the Symposium on Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese accentology, and other topics* (pp. 301-321). Tokyo: ILCAA.

Ross, B. B. (2011). *Prosody and grammar in Dalabon and Kayardild.* (Doctoral dissertation). University of Melbourne, Melbourne.

Sadat-Tehrani, N. (2007). *The Intonational grammar of Persian.* (Doctoral dissertation). University of Manitoba, Winnipeg, Manitoba. Retrieved from http://mspace.lib.umanitoba.ca/handle/1993/2839

Sadat-Tehrani, N. (2009). The alignment of L+ H* pitch accents in Persian intonation. *Journal of the International Phonetic Association*, *39*(2), 205-230. doi: 10.1017/s0025100309003892

Sadat-Tehrani, N. (2011). The intonation patterns of interrogatives in Persian. *Linguistic Discovery, 9*(1), 105-136. doi: 10.1349/ps1.1537-0852.a.389

Sadeghi, V. (2017). Word-level prominence in Persian: An experimental study. *Language and Speech*, *Online first*, *First published Jan. 23, 2017*, 0023830916684862. doi: 10.1177/0023830916684862

Sag, I. and Liberman, M. (1975). The intonational disambiguation of indirect speech acts. In R. E. Grossman, L. J. San, and T. J. Vance (Eds.), *Papers from the 11th Regional Meeting of the Chicago Linguistic Society* (pp. 487-498). Chicago: Chicago Linguistics Society.

Schafer, A. J., Speer, S. R., and Warren, P. (2005). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In M. Tanenhaus and J. Trueswell (Eds.), *Approaches to studying world situated language use: Psycholinguistic, linguistic and computational perspectives on bridging the product and action tradition* (pp. 209-226). Cambridge, Massachusetts: The MIT Press.

Silverman, K., Blaauw, E., Spitz, J., and Pitrelli, J. F. (1992). A prosodic comparison of spontaneous speech and read speech. *Proceedings of the 2nd International Conference on Spoken Language Processing (ICSLP), Banff, Canada* (pp. 1299-1302). Edmonton: University of Alberta.

Speer, S. R., Warren, P., and Schafer, A. J. (2011). Situationally independent prosodic phrasing. *Laboratory Phonology*, *2*(1), 35-98. doi: 10.1515/labphon.2011.002

Steedman, M. (2000). *The syntactic process.* Cambridge, MA: The MIT Press. doi: 10.1017/s0022226702281918

Stirling, L., Fletcher, J., Mushin, I., and Wales, R. (2001). Representational issues in annotation: Using the Australian map task corpus to relate prosody and discourse structure. *Speech Communication*, *33*(1), 113-134. doi: 10.1016/s0167-6393(00)00072-8

Swerts, M., Strangert, E., and Heldner, M. (1996). F0 declination in read-aloud and spontaneous speech. *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP), Philadelphia, USA* (pp. 1501-1504). doi: 10.1109/icslp.1996.607901

Taheri-Ardali, M. and Xu, Y. (2012). Phonetic realization of prosodic focus in Persian. *Speech Prosody 2012, Shanghai, China* (pp. 326-329). Retrieved from https://pdfs.semanticscholar.org/d172/72af61aa39645b91164121940e0ae4a6f007.pdf

Taheri-Ardali, M., Rahmani, H., and Xu, Y. (2014). The perception of prosodic focus in Persian. *Speech Prosody 2014, Dublin, Ireland* (pp. 515-519). Retrieved from http://www.homepages.ucl.ac.uk/~uclyyix/yispapers/Taheri-Ardali_etAl_SP2014.pdf

Tøndering, J. (2011). Preplanning of intonation in spontaneous versus read aloud speech: Evidence from Danish. In the Scottish Consortium for ICPhS 2011 (Ed.), *Proceedings of* *the 17th International Congress of Phonetic Sciences* (pp. 2010-2013). Glasgow, UK: The University of Glasgow. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Tondering/Tondering.pdf

Tuttle, S. G. and Lovick, O. (2007). Intonational marking of discourse units in two Dena'ina narratives. *Nouveaux cahiers de linguistique française, 28*, 305-316.

Wagner, P., Trouvain, J., and Zimmerer, F. (2015). In defense of stylistic diversity in speech research. *Journal of Phonetics, 48*, 1-12.

Wang, X., Li, A., and Yuan, C. (2008). A preliminary study on silent pauses in Mandarin *Speech Prosody 2008, Campinas, Brazil* (pp. 673-676). Retrieved from https://pdfs.semanticscholar.org/82d4/2f59ad48452931a58d3b416e559c0e0458bd.pdf

White, L., Wiget, L., Rauch, O., and Mattys, S. L. (2010). Segmentation cues in spontaneous and read speech *Speech Prosody 2010, 100218*, 1-4. Retrieved from http://speechprosody2010.illinois.edu/papers/100218.pdf

Xu, Y. (2010). In defense of lab speech. *Journal of Phonetics*, *38*(3), 329-336. doi: 10.1016/j.wocn.2010.04.003

Yeung, Y. T., Qian, Y., Lee, T., and Soong, F. K. (2008). Prosody for Mandarin speech recognition: A comparative study of read and spontaneous speech. *Proceedings of the 9th Annual Conference of the International Speech Communication Association (INTERSPEECH 2008)* (pp. 1133-1136). Brisbane, Australia: ISCA. Retrieved from https://pdfs.semanticscholar.org/994a/91610b2cc5a0975f27a24dd18acacf48cd63.pdf

Yuan, J. and Liberman, M. (2014). F0 declination in English and Mandarin broadcast news speech. *Speech Communication*, *65*, 67-74. doi: 10.1016/j.specom.2014.06.001

Nima Sadat-Tehrani
Centennial College
Department of English
941 Progress Avenue, Toronto ON, M1G 3T8
e-mail: nsadatte@my.centennialcollege.ca