Prosodic Cues for Broad, Narrow, and Corrective Focus in Persian

Mortaza Taheri-Ardali [iD]
Department of English, Shahrekord University, Iran; IfL–Phonetics, University of Cologne, Germany; School of Linguistics and Language Studies, Carleton University, Canada

Simon Roessig [iD]
Department of Language and Linguistic Science, University of York, UK

Lena Pagel [iD] **and Doris Mücke** [iD]
IfL–Phonetics, University of Cologne, Germany

Abstract
Previous studies have demonstrated that focus significantly alters sentential prosody in Persian. However, research on the phonetic realization of non-corrective narrow focus is scarce compared to that on broad and corrective focus. This paper presents a systematic production study investigating whether Persian speakers distinguish between three focus structures on target words that bear a pitch accent, that is, broad, narrow, and corrective focus. In a multidimensional phonetic analysis, we investigated the parameters of intensity, duration, and F0. Taking a local perspective, results show that the duration of the target word is a robust cue for focus marking in both syllables of the word, exhibiting a three-step pattern (corrective > narrow > broad). In the first syllable, intensity is a reliable cue to distinguish broad focus from the other two focus types, with higher intensities in broad focus. In the accented syllable, a different two-step pattern is observed, with narrow and corrective focus showing larger F0 spans than broad focus. Taking a global perspective that considers the parts of the utterance before and after the target word, we find a lowering of F0 and decreased intensity for narrow and corrective focus in the pre-target region. In the post-target region, we find strong evidence for differences in mean F0 and intensity with lower F0 in corrective focus than in broad and narrow focus, while the intensity is lower in narrow and corrective focus than in broad focus. Our analysis deepens our understanding of Persian prosody.

1 Introduction

Information structure refers to the structuring of sentences (information packaging in Chafe's terms) in different kinds of information blocks (Chafe, 1976; Féry & Ishihara, 2015). *Focus*, as one of the basic notions of information structure, is defined as the presence of alternatives relevant

Corresponding author:
Mortaza Taheri-Ardali, Department of English, Shahrekord University, Shahrekord, Chahar Mahal va Bakhtiari Province, Iran.
Emails: taheri@sku.ac.ir; mtaheria@uni-koeln.de; mortazataheriardali@cunet.carleton.ca

to the interpretation of linguistic expressions (Krifka, 2008). This concept has been extensively explored in different disciplines of linguistics—in semantics, pragmatics, syntax, morphology, and prosody. From a prosodic perspective, a distinction can be made between broad and narrow focus, differing in the size of the focus constituent (Ladd, 2008). All information is new in the broad focus sentence (1), and the focus constituent, as indicated by square brackets, covers the entire sentence.

1. Q: What happened?

   A: [Peter stole the cookie.]$_F$

In a narrow focus (Krifka, 2008, p. 250), some of the information is already given by the previous context, for example, a preceding question, and only some information is new. For example, in the mini-dialogue in (2), the word *Peter* is in focus, and the remaining part of the sentence is considered as the background.

2. Q: Who stole the cookie?

   A: [Peter]$_F$ stole the cookie.

Another pragmatic use of focus is to correct information. When the focus marks a constituent that directly rejects an alternative, either previously spoken by the speaker or by the interlocutor, the focus is called corrective (Gussenhoven, 2007, p. 91; see also Repp, 2016).

3. A: The capital of Finland is Oslo.

   B: The capital of Finland is [Helsinki]$_F$.

Languages use different prosodic cues, in addition to syntactic and/or morphological means, to express focus and draw a listener's attention to a particular part of an utterance (Chen et al., 2014; Kügler & Calhoun, 2020; Ladd, 2008; Roessig & Mücke, 2019; Zimmermann & Fery, 2010; among many others). During the past years, a plethora of acoustic studies from several language families revealed that fundamental frequency is an important marker of focus structure, followed by cues related to syllable duration and intensity (Baumann et al., 2007; Cooper et al., 1985; Frota, 2000; Kügler & Calhoun, 2020; Ladd & Arvaniti, 2023; Mücke & Grice, 2014; Roessig et al., 2022; Xu, 2011; among others). All studies on Persian so far have concentrated on the manifestation of one particular focus type, namely corrective focus, compared to broad focus as a baseline condition. Accordingly, it is unclear how different types of focus, including non-corrective narrow focus, affect the phonetic properties of prosodic prominence in Persian in a more fine-grained way.

The present research will experimentally investigate phonetic parameters of Persian utterances in three different focus conditions (broad, narrow, and corrective focus), which are elicited as answers to pre-recorded questions. We aim to characterize intonation contours (F0 mean, F0 span, F0 minimum, and F0 maximum), in addition to duration and intensity. The study will look at both the more local syllable level (two syllables of the target word) and at a more global level, investigating changes in the region before the target word (pre-target region) and after the target word (post-target region).

This article is structured into five key sections. Section 2 elaborates on the prosodic focus in Persian. Section 3 outlines the methodology, including aspects such as participants, speech material, procedures, recording techniques, annotation and measurement, and statistical analysis. The

results, presented in Section 4, covers visualizations of average F0 contours and detailed analyses of the syllables in the target word, as well as the surrounding pre-target and post-target regions. The final section offers a discussion of the findings and their broader implications.

2 Prosodic focus in Persian

The study of Persian prosody has a rich history within the literature, particularly evident in classical Persian grammar books. Yet, Vahidian-Kamyar's (1972) PhD dissertation stands as the pioneering comprehensive investigation within the framework of British tradition laid out by Crystal (1969). Since then, several significant contributions have been published, including the work by Towhidi (1974), Eslami (2000), Mahjani (2003), Kahnemuyipour (2003), Scarborough (2007), Sadat-Tehrani (2007, 2009), Abolhasanizadeh et al. (2012), Taheri-Ardali & Xu (2012), Taheri-Ardali et al. (2014), Taheri-Ardali & Xu (2015), Hosseini (2013, 2014), Rahmani (2019), Rahmani et al. (2015, 2018), Sadeghi (2017a, 2017b, 2022), and Shiamizadeh et al. (2017a, 2017b, 2018, 2019). Among these publications, some have addressed the issue of prosodic focus from different perspectives, identifying various phonetic and phonological correlates as indicators of corrective focus in Persian (Abolhasanizadeh et al., 2012; Hosseini, 2013; Rahmani et al., 2018; Sadat-Tehrani, 2007; Scarborough, 2007; Taheri-Ardali et al., 2014; Taheri-Ardali & Xu, 2012). It is worth bearing in mind that we attempt to include the original terminologies used in the literature, but for the sake of clarity, the current study's equivalent terms have been partly added in [because you changed the square brackets to curly in the following paragraph so I removed "square"] brackets. Particularly, broad focus, which was traditionally regarded as the baseline in prior studies, remains a focus type on its own within our current study. Note that some authors do not regard broad focus as a focus—constituents in broad focus are then non-focused or neutral to focus.

According to Sadat-Tehrani (2007), a focused accentual phrase has a larger pitch excursion (F0 span) and longer duration than its non-focused counterpart (broad focus). In addition, the non-focused part of a focused sentence is usually pronounced at a faster rate than the part directly under the focus. The focused accentual phrase has a low boundary tone and causes de-accenting up to the end of the utterance. Following this, Abolhasanizadeh et al. (2012) argue that extremely small durational, intensity, and spectral differences exist between focused and non-focused syllables. However, F0 differences are substantial. In addition, their production and perception experiments show post-focus compression (PFC), since pitch excursions in the post-focal speech are extensively reduced compared to other positions in the sentence. In a production experiment on corrective focus in Persian declarative sentences, Taheri-Ardali & Xu (2012) report that corrective focus considerably changes the pre-, on-, and post-focus regions of the utterance. This change is so intense that the domain of a single focus is considered to encompass three temporal zones, with distinct pitch range, duration, and intensity adjustments for pre-, on-, and post-focus components. In particular, for the on-focus components, they observe that the focused words (corrective focus) have a higher pitch and longer duration than those in the neutral focus (broad focus) condition. There is no significant change in intensity in this region. Pre-focus elements, however, are realized with weaker intensity than non-focused counterparts. The effect of focus structure on the mean F0 and duration of the pre-focus part is insignificant, but the maximum F0 is higher in this region. In the post-focus region, maximum F0, mean F0, and mean intensity of post-focus words are significantly decreased compared to the corresponding neutral focus ones, whereas no difference in duration is found. Hosseini (2013) argues in a cross-linguistic comparison of corrective focus marking in Japanese and Persian that although lengthening occurs in some focused accentual phrases, it is rather an idiolectal phenomenon and cannot be regarded as a consistent mechanism in Persian. In other words, the

prolongation of correctively focused words is used only by some speakers, but not all. He further states that a possible reason why previous studies consider an increase in duration in focused words is that they look at the mean values of duration, neglecting the significant individual differences between speakers. His results suggest that pitch expansion and post-focal reduction are the main mechanisms concerning focus marking in Persian while intensity and vowel quality are irrelevant.

Despite the differences in approach, data design, and study objectives in existing phonetic studies, a common finding has emerged that is particularly noteworthy. Persian is characterized as a language with PFC, where post-focal constituents exhibit a significantly reduced F0 range—a phenomenon widely documented in the literature (e.g., Mahjani, 2003; Rahmani et al., 2018; Sadat-Tehrani, 2007; Taheri-Ardali & Xu, 2012; among others). Rahmani et al. (2018), with the most recent study on Persian focus, report that word prominence is effectively removed after focus, to the extent that minimal pair members become homophonous. The authors further assert that while post-focus de-accentuation is governed by the grammar of focus, the typical F0-raising reported in some studies (e.g., Sadat-Tehrani, 2007; Scarborough, 2007; Taheri-Ardali & Xu, 2012; among others) is not inherently part of the focus grammar, and they attribute this to the systematic phonetic variation in the realization of focus. Rahmani et al. (2018) conclude that there is no phonological distinction between the accent for focus and the neutral accent, suggesting that F0-raising for focus is attributable to paralinguistic strategies employed by speakers. To support this, they refer to their data, which do not consistently show that the accent for focus is higher than the neutral accent, indicating that F0-raising is not a consistent mechanism for marking focus in Persian. In addition, drawing on perceptual data from the study by Taheri-Ardali et al. (2014), it has been suggested that PFC plays a crucial role in marking the position of the focused item. Moreover, Rahmani et al. (2018) mention Karimi's (2005: p. 132) syntactic approach to Persian focus, which claims that contrastive (corrective) focus and narrow focus exhibit different prosodic realizations, although the exact differences are not clearly defined. The authors note the absence of any experimental study systematically comparing the intonational effects of narrow focus and contrastive (corrective) focus in Persian. However, impressionistically, they claim that the existing data do not indicate any obvious differences between these focus types in terms of intonation.

Overall, the present paper aims to expand the phonetic analysis of Persian focus marking by systematically comparing the prosodic realizations of broad, narrow, and corrective focus on a more fine-grained scale in Persian. Furthermore, in light of the discussion outlined earlier, if we attribute the F0 changes in focused constituents to paralinguistic factors, it remains crucial to examine the role of phonetic parameters from further dimensions of speech production, that is, duration and intensity. In a broader context, and considering the ongoing debate in the literature regarding the role of duration in Persian prosody, we will address these issues in the discussion of this article.

The present study has an exploratory character and provides experimental evidence to answer the following questions:

1. Do Persian speakers differentiate between broad focus, non-corrective narrow focus, and corrective focus?
2. Which parameters are affected? Parameters that relate to F0, duration, and/or intensity?
3. Which parts of the sentence are affected, and to what extent? Are effects of focus structure found in pre- and post-focal regions?
4. Which parts of the focused target word are affected? Only the syllable that is expected to carry the pitch accent or also the preceding syllable?

Table 1. The Carrier Sentences of the Experiment.

[Y]	tupp = o	be	[target word]	mi-de
[Y]	ball = om	to	[target word]	ipfv-give.npst.3sg

'Y gives the ball to [target word].'

[Y] توپوبه [target word] میده .

[Y]	tupp = o	be	[target word]	pas_mi-de
[Y]	ball = om	to	[target word]	pass_ipfv-give.npst.3sg

'Y passes the ball to [target word].'

[Y] پاس میده [target word] توپوبه .

3 Method

This section outlines the methodological framework employed in this study. It details participant selection, the design of the speech material, experimental procedures, the recording setup, annotations, and data analysis techniques.

3.1 Participants

For the experiment, 12 native Persian speakers (7 female and 5 male) were recruited from the Iranian community living in Cologne, Germany. At the time of recording, the speakers were aged between 25 and 43 years (*M*: 34.58; *SD*: 4.85). They acquired Persian as their mother tongue in Iran and learned English and/or other languages later in life. At the time of the recording, the speakers had mostly moved to Germany in recent months or years—eight of them had been in Germany for a maximum of 3 years, while four had been there for more than 3 years (*M*: 3.8 years). They used Persian at home in Germany as their primary language of communication and in their communications with their families back at home in Iran. Prior to their participation in the experiment, informed consent was received from all participants, and they were compensated for their involvement in the study. In addition, their linguistic profile was collected using a questionnaire primarily derived from the introductory sections of the *Atlas of the Languages of Iran* (ALI) questionnaire (Anonby, Taheri-Ardali et al., 2015–2025). None of the participants reported any speech or hearing impairments.

3.2 Speech material

We composed a corpus of Persian question-answer pairs to provide a natural context to trigger the intended focus types. In this corpus, we constructed two sentences differing in the final verb, which were produced by the participants as answers to preceding questions, that is, as target utterances of the study (cf. Table 1).

Eight Persian-sounding pseudo-names were created as target words to be incorporated in the aforementioned carrier sentences. They were دیما (/dimɑ/), زیما (/zimɑ/), ریبا (/ribɑ/), ژیبا (/ʒibɑ/) and زامی (/zami/), ژامی (/ʒami/), دابی (/dɑbi/), and زابی (/zabi/). We expect the penultimate syllable to be the stressed/accented syllable. These eight target words and the first syllable of the verbs were designed to be /i-ɑ-i/ and /ɑ-i-ɑ/ vowel sequences. This was achieved by always having the target words /dimɑ, zimɑ, ribɑ, ʒibɑ/ be followed by the verb میده (/mide/) and /zami, ʒami, dɑbi, zabi/ by the verb پاس میده (/pɑs mide/) (cf. Table 1). Since the study was also intended to measure articulatory movements using 3D electromagnetic articulography, these pseudo-names with the

Table 2. An Example From the Question-Answer Pairs to Elicit the Focus Structures.

Focus structure	Example trigger and target sentence	Persian orthography
Background	Q: Maryam tuppo be Dimā mide? 'Does Maryam give the ball to Dimā?' A: [Shirin]$_F$ tuppo be Dimā mide. '[Shirin]$_F$ gives the ball to Dimā.'	سوال: مریم توپو به دیما میده؟ پاسخ: [شیرین]کانون توپو به دیما میده.
Broad focus	Q: Dāre chi mishe? 'What is happening?' A: [Shirin tuppo be Dimā mide]$_F$. '[Shirin gives the ball to Dimā]$_F$.'	سوال: داره چی میشه؟ پاسخ: [شیرین توپو به دیما میده]کانون.
Narrow focus	Q: Shirin tuppo be ki mide? 'Who does Shirin give the ball to?' A: Shirin tuppo be [Dimā]$_F$ mide. 'Shirin gives the ball to [Dimā]$_F$.'	سوال: شیرین توپو به کی میده؟ پاسخ: شیرین توپو به [دیما]کانون میده.
Corrective focus	Q: Shirin tuppo be Maryam mide? 'Does Shirin give the ball to Maryam?' A: Shirin tuppo be [Dimā]$_F$ mide. 'Shirin gives the ball to [Dimā]$_F$.'	سوال: شیرین توپو به مریم میده؟ پاسخ: شیرین توپو به [دیما]کانون میده.

The target word in these examples is "Dimā."

specific constant structure were chosen as target words. We filled the Y space in the carrier sentence with the common Persian female names مریم (Maryam), شبنم (Shabnam), مهگل (Mahgol), and گندم (Gandom).

The target sentence was produced by the participant as an answer to a question. Four types of questions elicited four types of focus structures of the target utterance. Table 2 shows an example of this question-answer paradigm. For the question eliciting the target word to be in corrective focus, competing names were used, which were نگین (Negin) and شیرین (Shirin) for the /i-ɑ-i/ sequence and بهار (Bahar) and پوران (Puran) for the /ɑ-i-ɑ/ sequence.

We recorded 1,152 sound files: 12 speakers × 8 target sequences × 4 focus conditions × 3 repetitions. Thus, each speaker produced 24 sentences in each focus condition. We had to exclude two utterances from the corpus as a result of minor disfluencies. The total number of productions analyzed in this study is thus 1,150.

3.3 Procedures

The question-answer task was designed as a game where participants interacted with a virtual avatar. The speaker adhered to the instructions on a screen that was positioned at a distance of around 3 meters to them. In the game scenario, they were told that they were a soccer fan going to the stadium to watch a match of their favorite team with their friend Marie, who was the virtual avatar. Marie unfortunately forgot her glasses and could not follow the game. The speaker had to respond to her questions and explain the events on the screen.

In the beginning, the target words were introduced to the participants as the names of the female soccer players on the pitch. After a training phase, the experiment began. All speakers reported that they quickly became familiar with the procedure and found the experiment to be straightforward. After seeing an animation of the soccer scene, participants read and heard the question, which had been pre-recorded by a female Persian speaker. After that, participants produced the target utterance as written on the screen as an answer to the question. Figure 1 shows a screenshot of one example trial.
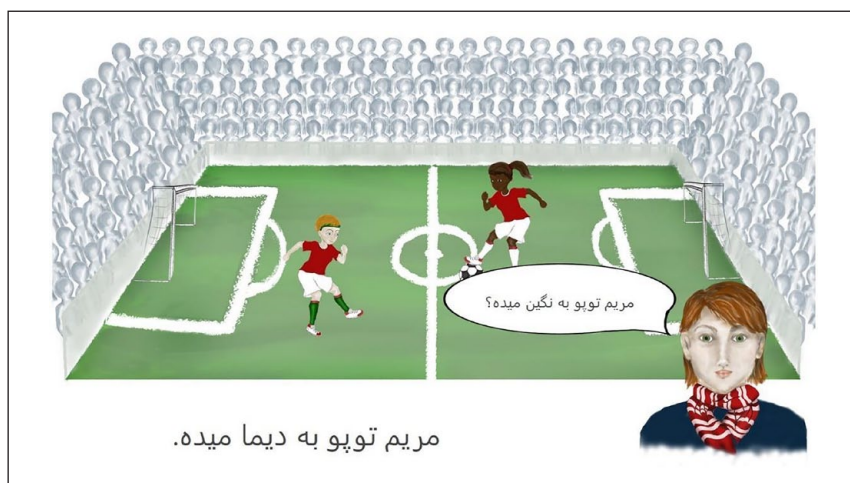
Figure 1. A screen from the soccer game scenario during a trial with a corrective focus condition (Question in bubble: "Does Maryam give the ball to Negin?" Target sentence below soccer pitch: "Maryam gives the ball to Dimā").
Source. Image: CC-BY 4.0.

Using HTML/CSS and JavaScript, the experiment was constructed as an interactive website with jQuery to animate the movements in the soccer game. The experimenter used a keyboard to control the flow of the experiment. If the participant produced speech disfluencies, the trial was repeated. To obtain natural, unbiased data, the experimenters provided no further explanations except a brief introduction into how the experiment should run. Each participant was assigned a unique randomized list of trials. The randomization was designed so that the same target word could not appear in subsequent trials. In addition, more than two subsequent trials with the same focus condition were not possible. And, 12 filler sentences were added to each recording session to prevent predictability.

3.4 Recording

The recordings were carried out at the IfL–Phonetics Department of the University of Cologne. Acoustic recordings were conducted with an AKG C520 headset microphone connected to a computer via a PreSonus AudioBox 22 VSL interface at a sampling rate of 44.1 kHz (mono channel, 16-bit). The headset microphone was placed at a constant distance of 7 cm from the speaker's mouth. Since the study additionally aimed to measure supralaryngeal movements, we tracked the movements of the articulators with 3D electromagnetic articulography (Carstens AG501), synchronized with the acoustic recordings. The present paper does not include results on supralaryngeal articulation and focuses on acoustic modulations related to syllable duration, F0, and intensity.

3.5 Annotations and measurements

The following four time intervals were manually annotated in Praat (Boersma & Weenink, 2001) by a native speaker of Persian:

1. First syllable of the target word
2. Second syllable of the target word

3. Pre-target region: from the start of the utterance to the start of the target word
4. Post-target region: from the end of the target word to the end of the utterance

F0 was extracted on a time-normalized scale, with 50 points per interval. The F0 extraction was done with Praat (Boersma & Weenink, 2001) through the Python interface Parselmouth (Jadoul et al., 2018). The Pitch object was created with default parameters except for the floor (50 Hz) and ceiling (500 Hz), and F0 was smoothed with a Gaussian filter (Praat's "Smooth" function with a bandwidth of 10 Hz). F0 was then normalized for each speaker: Each F0 measurement point was expressed in semitones relative to the average F0 of the speaker that produced the data point, that is, $12 * log_2\left(\dfrac{F0}{Average_{Speaker}}\right)$. Intensity was calculated using Praat Intensity objects created with default parameters except for the minimum pitch (50 Hz). The mean values were then extracted by calling Praat's "Get mean" function on each Intensity object. In the four intervals (first and second syllables of the target word, pre- and post-target region), the following six phonetic parameters were extracted for the analysis:

-   F0 mean over the duration of the interval (semitones with reference to speaker average)
-   F0 minimum in the interval (semitones with reference to speaker average)
-   F0 maximum in the interval (semitones with reference to speaker average)
-   F0 span as the difference between the maximum and the minimum (semitones)
-   Duration (milliseconds)
-   Intensity (decibels)[1]

For the visualization of F0 average contours, the time-normalized F0 contours were averaged at each time point. The average was calculated only for those time points where at least 20 F0 measures were available (otherwise, the point is NA in the average contour vector). A rolling mean with a window of 5 points was applied to the average contours.

3.6 Statistical analysis

The data were modeled statistically using Bayesian linear mixed models for each phonetic parameter with brms (Bürkner, 2018), an interface to Bayesian inference in Stan (Carpenter et al., 2017). A model was fit for each of the four intervals and each of the six phonetic parameters. All models contained focus condition as a fixed effect. Random intercepts were included for speakers and target words. In addition, the models included by-speaker random slopes for the effect of focus condition. All models ran with four chains with 4,000 iterations. Weakly informative priors for regression coefficients with a mean of zero and a standard deviation of 10 were used. All other priors were the default priors of brms. Convergence of the models was checked by ensuring no model yielded R-hat values larger than 1. The R-hat statistic considers the variance of the samples of each sampling chain to that from all samples across all chains. Its value indicates whether the different chains reached a similar outcome (Franke & Roettger, 2019).

3.7 Ethics statement

This study was approved by the ethics committee of the Deutsche Gesellschaft für Sprachwissenschaft (2020-04-200327). Each participant gave written informed consent before participating. The research was conducted in accordance with the Declaration of Helsinki.
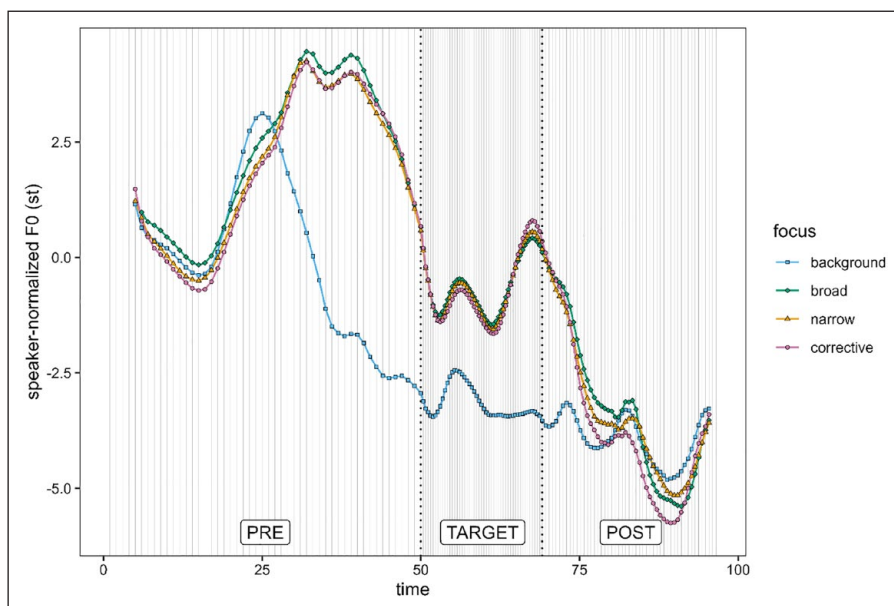
Figure 2. Average F0 contours over the complete sentence for all four focus conditions.

4 Results

We begin the presentation of the results by looking at the average F0 contours for all four focus conditions over the entire sentence. This step is mainly a visualization of the data, which provides a first overview of the intonational characteristics of the pre-target, target, and post-target regions of the utterance. To get a deeper understanding of the modulations across the three focus types—broad, narrow, and corrective focus—we then present results for the statistical modeling of the six parameters: F0 mean, F0 span, F0 minimum, F0 maximum, duration, and mean intensity. In doing so, we take both a local and a global perspective: Initially, we consider the two syllables of the target word (the local perspective). Subsequently, we expand our analysis to include an examination of the pre-target and post-target regions (the global perspective).

4.1 Visualization of average F0 contours

Figure 2 shows the average F0 contours for all four focus conditions during the entire utterance. The sentence is divided into three parts: (1) the region from the start of the sentence to the start of the target word; (2) the target word; (3) the region from the end of the target word to the end of the sentence. The first part is referred to as the "pre-target region" (abbreviated as "PRE" in the figure), and the third part is termed the "post-target region" (abbreviated as "POST" in the figure). The two syllables of the target word are not distinguished here. Note that the terms pre-target, target, and post-target shall not be confused with the terms pre-focal, focal, and post-focal. This distinction will become important later on.

The average contours are extracted on a normalized timescale with 50 time steps in each region. In the figure, this timescale is warped for each region to reflect the proportions of the three regions. The gray vertical lines visualize the time steps. The black dashed vertical lines show the boundaries of the three regions.

The pre-target region is pre-focal in narrow and corrective focus, while it is part of the focus in the broad focus condition, where the entire utterance is in focus. In the background condition, the
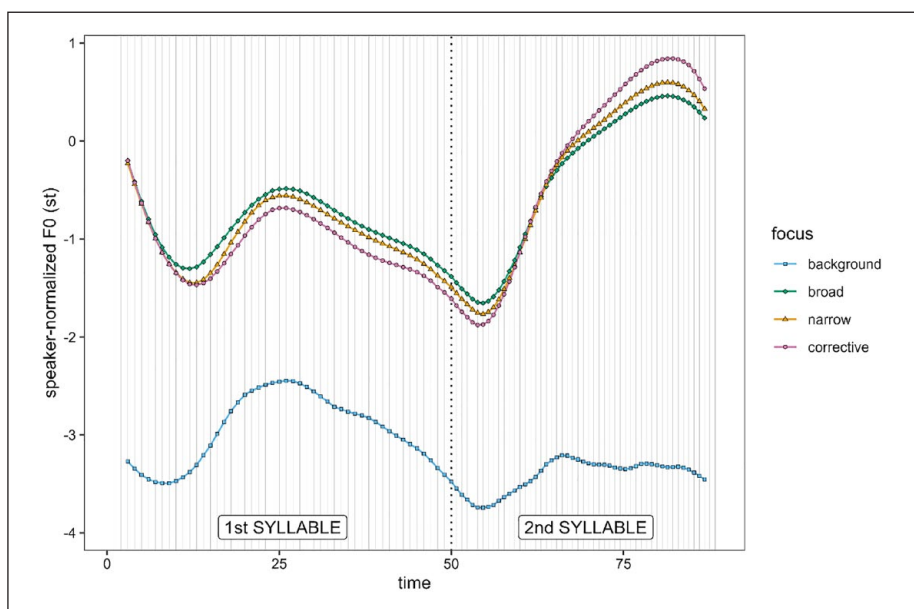
Figure 3. Average F0 contours over the target word.
Note. The x-axis represents normalized time.

pre-target region contains the element in focus (which is a female name that is different from the target word) and a part of the post-focal region. The post-target region is post-focal in all focus conditions, except for broad focus. The target region is focal in broad, narrow, and corrective focus. In the background condition, it is post-focal. See Table 2 again for clarification.

It is apparent in Figure 2 that the background condition has a very different average F0 contour from the other three conditions. This result is not surprising since the location of the focus is early in the sentence (on the first word). Consequently, we get an early F0 peak followed by a relatively flat F0. The other three conditions, where the target region is focal, are more similar to each other. Nevertheless, this first look at the data already reveals some differences: Broad focus seems to have a higher F0 peak in the pre-target region than narrow and corrective focus. In the target word, its course is first above the other two and then below them. We now zoom in on the target word to get a clearer picture of the differences in this region.

Figure 3 shows the average F0 contours during the target word. The word is divided into the first and second syllables. Again, the extraction of the F0 measurement points works on a time-normalized scale, and the time steps are warped in this display to reflect the proportions of the syllables. Here, the differences between broad, narrow, and corrective focus, albeit being quite subtle, become more apparent. In the first syllable of the word, F0 appears to be overall higher in broad focus than in narrow focus, and higher in narrow focus than in corrective focus. In the second syllable, the picture is reversed: F0 reaches the highest peak in corrective focus, followed by narrow focus, followed by broad focus. Interestingly, at the beginning of the second syllable, we additionally find the lower minimum for corrective focus, followed by narrow focus, and then by broad focus. This means that not only is the peak higher but also the preceding dip is lower, leading to the whole F0 excursion in the second syllable to be enlarged.

It should be noted that there is considerable tonal movement in the first syllable of the target word in the case of broad, narrow, and corrective focus. Since we do not expect an H component

Table 3. Descriptive Means and Standard Deviations for All Phonetic Parameters in the First Syllable of the Target Word.

Focus	F0 mean (st)		F0 span (st)		F0 minimum (st)		F0 maximum (st)		Duration (ms)		Intensity (dB)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Broad	−0.84	1.10	2.67	2.02	−2.06	1.78	0.61	1.32	198.89	31.17	70.94	4.68
Narrow	−0.91	0.85	2.71	1.46	−2.11	1.24	0.60	1.27	211.30	29.91	70.58	4.68
Corrective	−1.00	0.84	2.83	1.54	−2.21	1.30	0.62	1.36	224.62	31.75	70.29	4.67

of the pitch accent on this syllable, it is plausible to assume the F0 movement here may be influenced substantially by microprosodic effects induced by the onset consonants. However, since the segmental makeup of the target words remains constant across focus conditions (and hence also the influence of microprosody on the F0 contour), we can observe that there is a level difference between the focus types in the first syllable, with broad focus having the highest and corrective focus having the lowest F0 level, and narrow focus lying in between.

The background condition shows a very different F0 course both in the global picture (whole sentence, cf. Figure 2) as well as in the local picture (target word, cf. Figure 3). As mentioned earlier, this is because the focus structure of the background condition is very different: Here, there is an early focus on the first word of the utterance, and the target word occurs in the post-focal region. The comparison between the other three focus conditions is theoretically more interesting, since these are the conditions in which the target word is focused. Hence, in the following statistical analyses of F0, duration, and intensity parameters, we concentrate on broad, narrow, and corrective focus.

4.2 First syllable of target word

Table 3 presents the descriptive statistics (means and standard deviations) for all six phonetic parameters in the first syllable of the target word. The results of the Bayesian regression analysis are presented in Figure 4 and Table 4: The figure shows the expected means for the three focus conditions with 90% credible intervals (CIs) given the data and the regression model. The table lists the estimated differences (or contrasts) between focus conditions. If, for instance, the difference "narrow − broad" is positive, we can conclude that narrow focus has a greater mean value for a given phonetic measurement than broad focus. If the difference is negative, we can conclude that narrow has a smaller value than broad focus. The columns Q5 and Q95 indicate the limits of the 90% CI. This interval indicates how strong the evidence for the estimated difference is. We consider evidence for an effect compelling if the CI does not include zero.

We can observe that this is the case for all differences regarding duration. Narrow focus is characterized by a longer duration of the first syllable of the target word compared to broad focus. Corrective focus shows an even longer duration (broad < narrow < corrective). In the case of intensity, there is strong evidence for differences between broad and narrow focus, as well as between broad and corrective focus (broad > narrow, and broad > corrective). Interestingly, the pattern is reversed for intensity compared to duration: Broad focus has higher intensities than narrow and corrective focus. There does not appear to be compelling evidence for a difference between narrow and corrective focus in intensity.
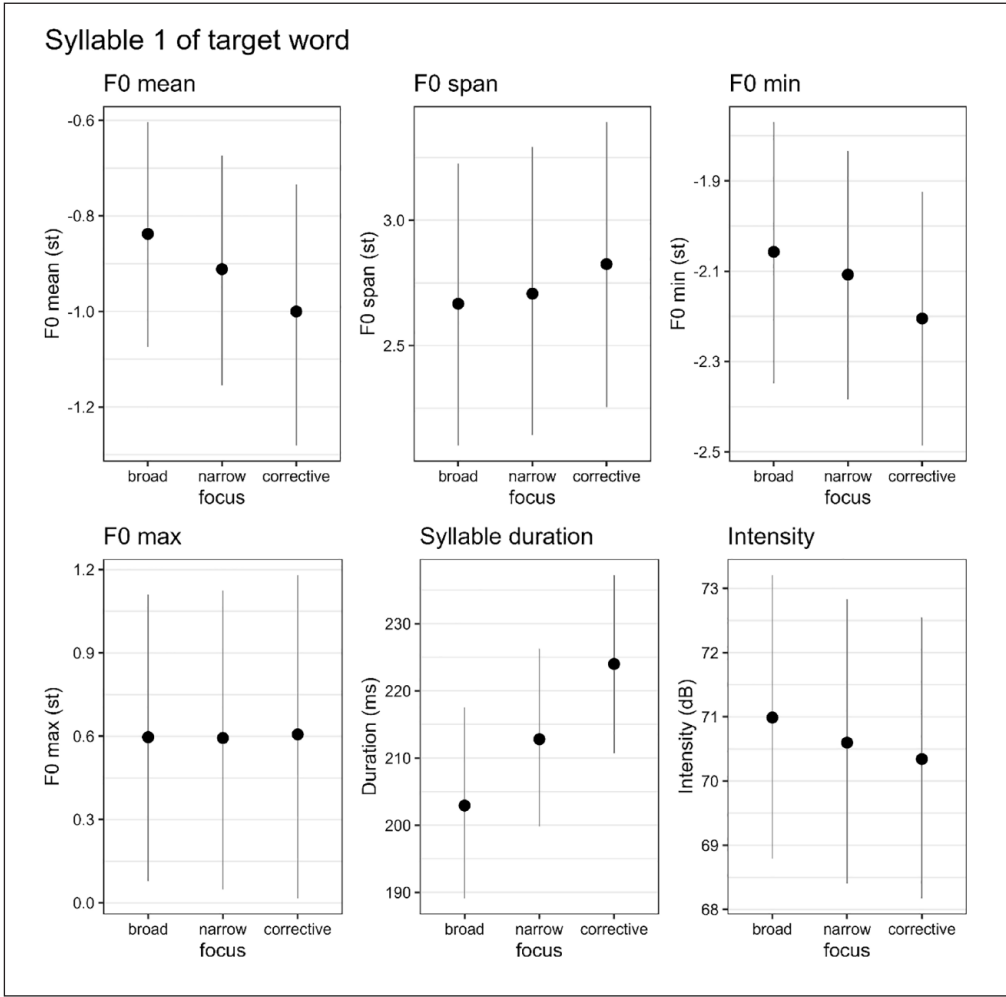
Figure 4. Expected means with 90% credible intervals given the data and the Bayesian regression model for the first syllable in the target word.

4.3 Second syllable of target word

For the second syllable of the target word, the descriptive statistics (means and standard deviations) for all six phonetic parameters are presented in Table 5. The results of the Bayesian regression analysis are presented in Figure 5 (expected means with 90% CIs) and Table 6 (estimated differences). We find compelling evidence for differences in the F0 span between broad and narrow focus as well as between broad and corrective focus: Narrow and corrective focus exhibit larger F0 spans than broad focus (broad < narrow, and broad < corrective). Looking at the F0 minimum and maximum results, it becomes evident that these changes in span are caused by both a lowering of the F0 minimum and a raising of the F0 maximum.

As in the first syllable, focus type also affects duration in the second syllable. We find compelling evidence for differences in all comparisons (the 90% CIs for all differences do not include

Table 4. Estimated Parameters From the Bayesian Regression Model for the First Syllable of the Target Word.

F0 mean (st)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	−0.07	0.09	−0.22	0.07
Corrective − broad	−0.17	0.12	−0.37	0.04
Corrective − narrow	−0.09	0.11	−0.27	0.09

F0 span (st)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	0.04	0.15	−0.21	0.29
Corrective − broad	0.16	0.17	−0.12	0.43
Corrective − narrow	0.11	0.16	−0.14	0.37

F0 min (st)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	−0.05	0.13	−0.27	0.16
Corrective − broad	−0.15	0.14	−0.38	0.09
Corrective − narrow	−0.10	0.14	−0.33	0.13

F0 max (st)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	0.00	0.13	−0.21	0.20
Corrective − broad	0.01	0.18	−0.29	0.30
Corrective − narrow	0.01	0.15	−0.23	0.25

Duration (ms)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	9.85	3.26	4.25	14.84
Corrective − broad	20.91	4.45	13.17	27.70
Corrective − narrow	11.05	4.14	4.04	17.40

Intensity (dB)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	−0.39	0.21	−0.74	−0.05
Corrective − broad	−0.65	0.22	−1.00	−0.30
Corrective − narrow	−0.26	0.25	−0.66	0.15

zero) with shortest durations in broad focus and longest durations in corrective focus while narrow focus ranges found in between the two (broad < narrow < corrective). One should note that the effect sizes for duration are smaller in the second syllable than in the first syllable.

Table 5. Descriptive Means and Standard Deviations for All Phonetic Parameters in the Second Syllable of the Target Word.

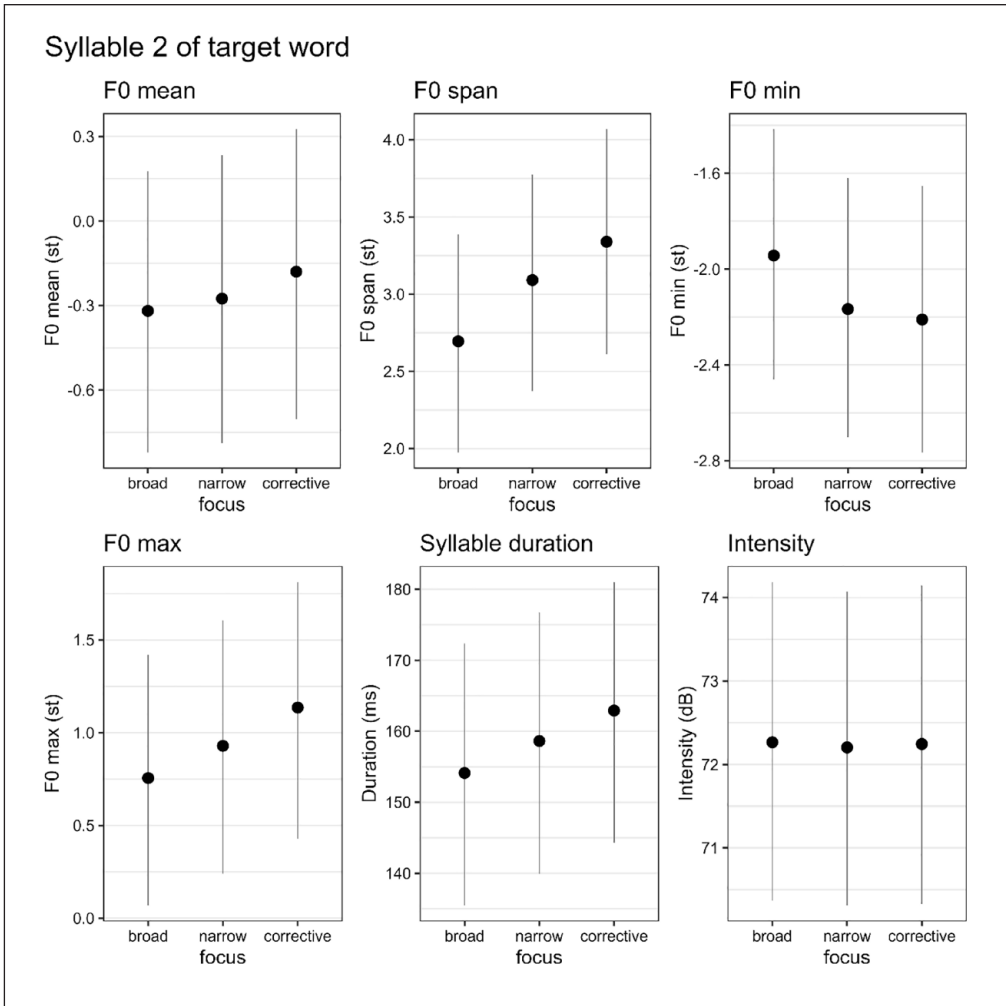| Focus	F0 mean (st)		F0 span (st)		F0 minimum (st)		F0 maximum (st)		Duration (ms)		Intensity (dB)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Broad	−0.33	1.13	2.68	1.44	−1.95	1.28	0.74	1.39	154.28	31.95	72.20	4.00
Narrow	−0.29	1.21	3.08	1.57	−2.18	1.57	0.90	1.47	159.26	31.38	72.17	3.80
Corrective	−0.20	1.06	3.34	1.53	−2.22	1.37	1.12	1.32	163.49	31.91	72.18	3.86

Figure 5. Expected means with 90% credible intervals given the data and the Bayesian regression model for the second syllable in the target word.

Table 6. Estimated Parameters From the Bayesian Regression Model for the Second Syllable of the Target Word.

F0 mean (st)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	0.04	0.08	−0.09	0.17
Corrective − broad	0.14	0.13	−0.07	0.35
Corrective − narrow	0.10	0.12	−0.11	0.30

F0 span (st)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	0.40	0.11	0.23	0.57
Corrective − broad	0.65	0.18	0.36	0.94
Corrective − narrow	0.25	0.18	−0.04	0.54

F0 min (st)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	−0.23	0.12	−0.42	−0.03
Corrective − broad	−0.27	0.15	−0.51	−0.02
Corrective − narrow	−0.04	0.14	−0.27	0.20

F0 max (st)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	0.17	0.09	0.02	0.32
Corrective − broad	0.38	0.18	0.09	0.67
Corrective − narrow	0.21	0.17	−0.07	0.48

Duration (ms)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	4.48	1.64	1.79	7.09
Corrective − broad	8.84	1.76	5.89	11.65
Corrective − narrow	4.35	1.67	1.64	7.05

Intensity (dB)				
Difference	Estimate	*SD*	Q5	Q95
Narrow − broad	−0.07	0.16	−0.34	0.20
Corrective − broad	−0.03	0.16	−0.29	0.24
Corrective − narrow	0.04	0.18	−0.26	0.34

Table 7. Descriptive Means and Standard Deviations for All Phonetic Parameters in the Pre-Target Region.

Focus	F0 mean (st)		F0 span (st)		F0 minimum (st)		F0 maximum (st)		Duration (ms)		Intensity (dB)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Broad	2.04	1.13	7.94	4.13	−1.61	2.66	6.33	3.38	941.36	89.62	68.05	3.88
Narrow	1.77	1.13	7.99	4.40	−2.26	3.72	5.74	2.65	904.03	92.78	67.62	3.85
corrective	1.74	0.93	8.09	4.07	−2.09	2.71	6.00	3.08	911.05	86.34	67.31	3.81

4.4 Pre-target region

Table 7 presents the descriptive statistics (means and standard deviations) for all six phonetic parameters in the pre-target region, that is, the interval from the start of the utterance until the start of the target word. The results of the Bayesian regression analysis are presented in Figure 6 (expected means with 90% CIs) and Table 8 (estimated differences). We find evidence for differences between broad and narrow focus, as well as between broad and corrective focus in terms of F0 mean, F0 minimum, and intensity. In all cases, narrow and corrective focus exhibit lower values than broad focus (broad > narrow, and broad > corrective). In addition, F0 maxima are lower, and durations are shorter in narrow focus than in broad focus (broad > narrow).

Note that in broad focus, the pre-target domain is part of the focus, while it is pre-focal in both narrow and corrective focus. The parameter modulations found here may thus be attributable to a distinction between focal and pre-focal. This distinction is characterized by a decrease in intensity and a general lowering of F0 mean and F0 minimum (note that both F0 minimum and duration are lowered at least in narrow compared to broad focus).

4.5 Post-target region

Table 9 presents the descriptive statistics (means and standard deviations) of the post-target region for all six phonetic parameters, which is the interval from the end of the target word to the end of the utterance. The results of the Bayesian regression analysis are presented in Figure 7 (expected means with 90% CIs) and Table 10 (estimated differences). In this region, we find compelling evidence for modulations of the parameters F0 mean and intensity. F0 means are lower in corrective than in broad and narrow focus (broad > corrective, and narrow > corrective). Intensities are lower in narrow and corrective focus than in broad focus (broad > narrow, broad > corrective).

5 Discussion and conclusion

We have presented data on the prosodic marking of focus in Persian with respect to expressions of broad, narrow, and corrective focus in a multidimensional phonetic analysis. To summarize the main results: In the first syllable of the target word, we find compelling evidence for effects of focus type on duration and intensity. Durations are longer in narrow than in broad focus, and longer in corrective than in narrow focus. While duration exhibits a three-step pattern between all three focus types, intensity only shows differences between broad focus and the other two types, with broad focus having higher intensities than narrow and corrective focus.

In the second syllable of the target word, we find compelling evidence for effects of focus type on duration and F0. Durations are longer in narrow than in broad focus, and longer in corrective than in narrow focus. F0 span only shows differences between broad focus and the other two types,
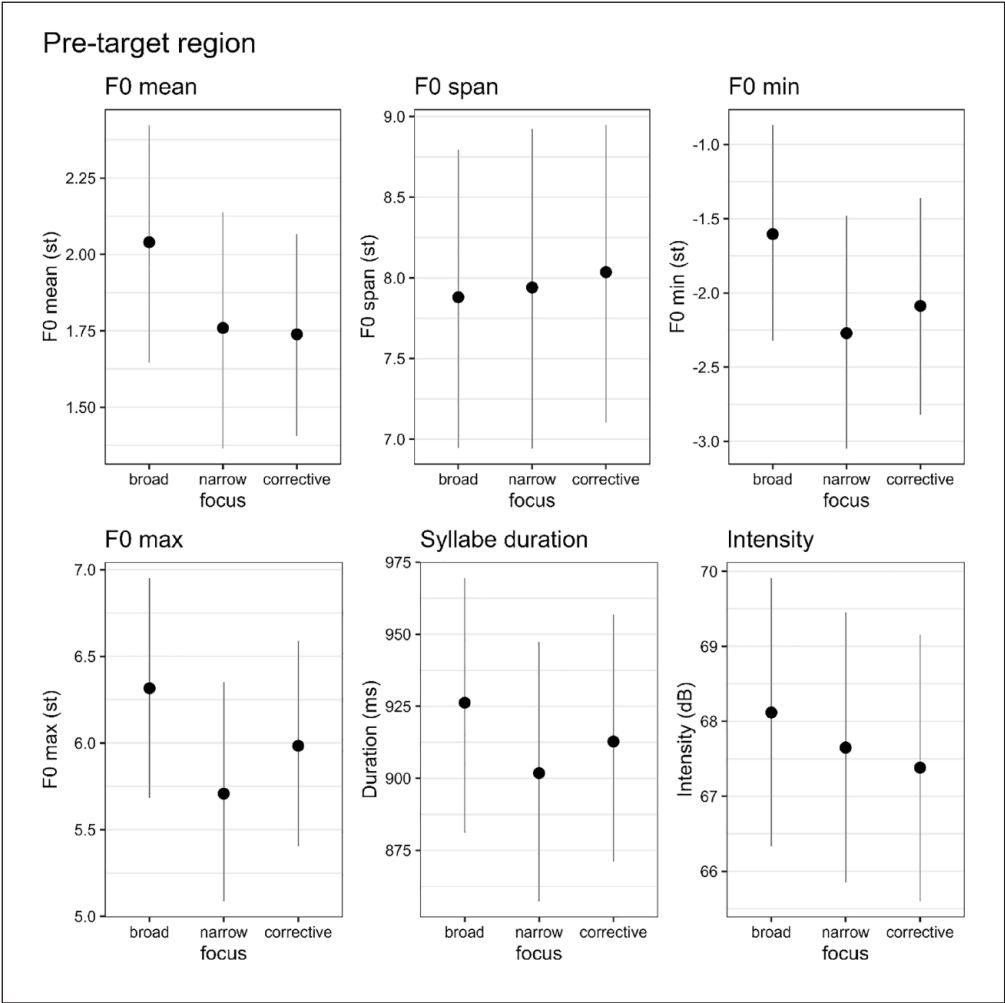
Figure 6. Expected means with 90% credible intervals given the data and the Bayesian regression model for the pre-target region.

with narrow and corrective focus having a larger F0 span than broad focus. This difference in span is caused by a simultaneous lowering of the F0 minimum and raising of the F0 maximum.

In the pre-target region, we find a lower F0 mean and F0 minimum, as well as a lower intensity in narrow and corrective focus than broad focus. F0 maximum and duration are only affected when comparing broad and narrow focus, with shorter durations and a lower F0 maximum in narrow focus. In the post-target region, we find compelling evidence for effects of focus type on F0 mean and intensity. We report lower F0 means in corrective focus than in broad and narrow focus. In addition, the results show lower intensities in narrow and corrective focus than in broad focus. All results are summarized in Table 11.

Taking the local linguistic perspective, that is, zooming in on the target word, we find focus types affecting the second syllable of the target word in terms of F0, intensity, and duration. This is also the locus where we expect these modulations the most, since the second syllable is predicted

Table 8. Estimated Parameters From the Bayesian Regression Model for the Pre-Target Region.

F0 mean (st)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	−0.28	0.08	−0.41	−0.15
Corrective − broad	−0.30	0.14	−0.52	−0.08
Corrective − narrow	−0.02	0.14	−0.24	0.21

F0 span (st)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	0.06	0.38	−0.56	0.67
Corrective − broad	0.15	0.34	−0.42	0.73
Corrective − narrow	0.09	0.40	−0.56	0.74

F0 min (st)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	−0.66	0.27	−1.10	−0.22
Corrective − broad	−0.48	0.26	−0.91	−0.05
Corrective − narrow	0.18	0.29	−0.30	0.64

F0 max (st)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	−0.60	0.26	−1.02	−0.18
Corrective − broad	−0.33	0.26	−0.77	0.10
Corrective − narrow	0.28	0.28	−0.18	0.73

Duration (ms)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	−24.32	6.31	−33.99	−13.17
Corrective − broad	−12.94	7.99	−25.01	0.81
Corrective − narrow	11.38	7.68	−0.60	24.17

Intensity (dB)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	−0.46	0.16	−0.73	−0.20
Corrective − broad	−0.74	0.16	−1.00	−0.48
Corrective − narrow	−0.27	0.18	−0.57	0.01

Table 9. Descriptive Means and Standard Deviations for All Phonetic Parameters in the Post-Target Region.

	F0 mean (st)		F0 span (st)		F0 minimum (st)		F0 maximum (st)		Duration (ms)		Intensity (dB)	
Focus	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Broad	−3.50	2.19	9.84	6.99	−8.60	6.03	1.24	3.32	515.83	161.01	64.15	4.51
Narrow	−3.53	2.36	9.84	6.43	−8.49	5.83	1.35	3.56	513.76	155.04	63.64	4.4
Corrective	−3.83	2.41	10.4	6.86	−8.84	6.06	1.56	3.77	517.12	154.42	63.34	4.61
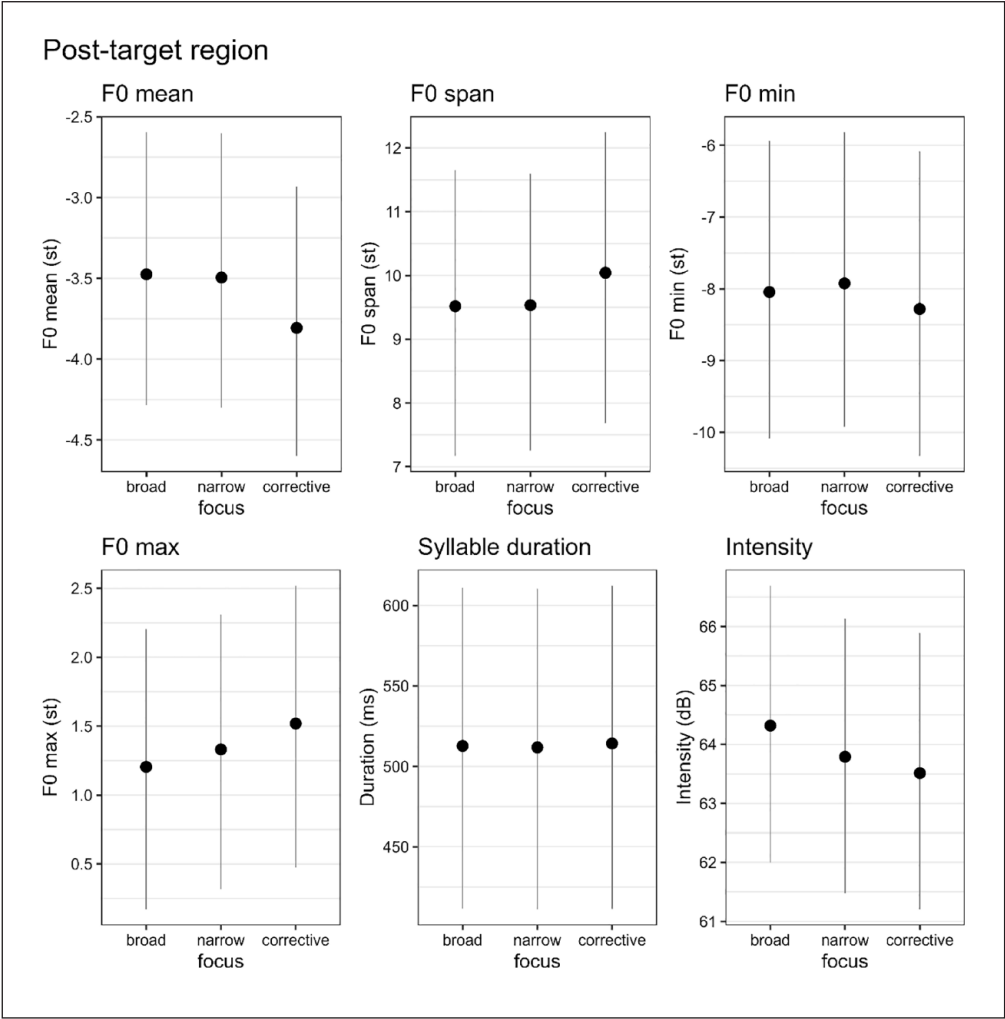
Figure 7. Expected means with 90% credible intervals given the data and the Bayesian regression model for the post-target region.

Table 10. Estimated Parameters From the Bayesian Regression Model for the Post-Target Region.

F0 mean (st)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	−0.02	0.16	−0.29	0.25
Corrective − broad	−0.33	0.17	−0.59	−0.06
Corrective − narrow	−0.31	0.18	−0.61	−0.02

F0 span (st)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	0.01	0.51	−0.83	0.84
Corrective − broad	0.54	0.49	−0.25	1.34
Corrective − narrow	0.53	0.52	−0.33	1.39

F0 min (st)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	0.11	0.42	−0.56	0.80
Corrective − broad	−0.23	0.40	−0.89	0.42
Corrective − narrow	−0.34	0.43	−1.05	0.36

F0 max (st)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	0.13	0.30	−0.36	0.62
Corrective − broad	0.32	0.32	−0.19	0.83
Corrective − narrow	0.19	0.33	−0.34	0.72

Duration (ms)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	−0.71	4.24	−7.68	6.34
Corrective − broad	1.19	3.85	−5.19	7.55
Corrective − narrow	1.90	4.01	−4.67	8.30

Intensity (dB)				
Difference	Estimate	SD	Q5	Q95
Narrow − broad	−0.54	0.18	−0.84	−0.23
Corrective − broad	−0.82	0.19	−1.13	−0.51
Corrective − narrow	−0.28	0.2	−0.6	0.05

phonologically to bear the pitch accent. However, there are also effects on the first syllable of the target word, but they are restricted to the non-tonal domain, that is, duration and intensity. Interestingly, the effect of focus types on duration is even stronger in the first syllable than in the

Table 11. Summary of Findings.

	F0 mean	F0 span	F0 min	F0 max	Duration	Intensity
Syllable 1 of target word					br < na < co	br > na br > co
Syllable 2 of target word		br < na br < co	br > na br > co	br < na br < co	br < na < co	
Pre-target	br > na br > co		br > na br > co	br > na	br > na	br > na br > co
Post-target	br > co na > co					br > na br > co

Note. br = broad focus, na = narrow focus, co = corrective focus.

second syllable. In addition, we only find robust evidence for intensity modulations in the first syllable (not in the second syllable). It is worth noting that the duration is increased from broad to narrow to corrective focus, whereas the intensity is decreased from broad to narrow focus and from broad to corrective focus. The prosodic weakening of elements not only in the first syllable of the target word but also in pre-target and post-target elements may serve to enhance the salience of the second syllable of the target word. Importantly, it was already indicated that PFC is crucial in marking the position of the focused item.

While duration appears to play a systematic role in differentiating between the three focus conditions under investigation, it is worthwhile to take a look at how durational modulations of the textual string interact with F0 cues of the tonal domain. Hosseini (2014) asserts that duration alone cannot signal the difference between pre-nuclear and nuclear accents. In other words, when combined with F0 modulations, durational adjustments support the distinction between the two accents more effectively and efficiently (Hosseini, 2014, p. 178). Thus, the role of duration has been regarded as secondary, with primary emphasis placed on F0 modulations. Our findings from the present production experiment, however, reveal different results, highlighting duration as an important phonetic cue in post-lexical prosody. This is highly relevant because, although F0 changes are limited to the differentiation of broad focus from narrow and corrective focus in the second syllable only, duration varies significantly in a three-step differentiation from broad to narrow to corrective focus.

Taking the global linguistic perspective, that is, looking at the entire utterance, we find that focus structure affects more material of the utterance than the target word. Especially the pre-target region is realized differently when it is part of the focus (in the broad focus condition) compared to when it is pre-focal (in the narrow and corrective focus conditions). This finding is in line with a growing body of evidence for the effect of focus on the pre-nuclear domain in various languages (e.g., Alzamil & Hellmuth, 2021; Andreeva et al., 2017; Baumann et al., 2006; Féry & Kügler, 2008; Kügler, 2008; Roessig, 2023; Royer & Jun, 2019). It is noteworthy that when looking at the parameter F0 mean, we find that broad focus is different from both narrow and corrective focus in the *pre*-target region, but corrective focus is distinct from both broad and narrow in the *post*-target region.

We now turn to the research questions formulated at the beginning of the paper. All in all, we mostly find a distinction between broad focus on the one hand and narrow and corrective focus on the other hand. Only duration is an exception here: In both the first and second syllables of the target word, duration shows a three-step differentiation between all three focus types. For many other parameters, there seems to be a trend for a three-step differentiation, but the evidence for

differences between narrow and corrective focus is not strong enough in the regression models in the way defined in the methodology (90% CI must not include zero). Our answer to question (1) on differentiating broad focus, non-corrective narrow focus, and corrective focus is as follows: It depends on the phonetic parameter considered. In general, we see that speakers do make a difference between non-corrective narrow and corrective focus, but this distinction is less clear and affects fewer phonetic parameters than the one between broad focus and the other two types.

Research question (2) targeted the different prosodic parameters involved in the distinction of focus types. We find that duration plays a major role but that F0-related parameters and intensity are also involved in the differentiation. As mentioned earlier, F0 span emerged as a crucial factor in distinguishing broad and narrow focus on one hand and broad and corrective focus on the other hand in the second syllable of the target word. It is also worth noting that this result shows that F0 is a complex phonetic measurement. Just considering the mean may obscure important differences: A simultaneous, symmetric lowering of the F0 valley and raising of the F0 peak of a rising movement leaves the mean unaltered (while span may reveal this vital difference).

To come back to our research question (3), we can confirm previous studies showing that focus structure affects the sentence globally. In the stretch before the target word (i.e., the pre-target region), we find a differentiation of focal (in broad focus) versus pre-focal (in narrow/corrective focus) condition. We also find post-focal effects (in the post-target region), but these seem to come out most strongly when comparing broad and corrective focus.

Our research question (4) took a closer look at the two syllables of the target word. It is interesting to note that we do not only find a focus-induced modulation of phonetic parameters in the second syllable of the target word, which is expected to carry the pitch accent, but also in the first syllable of the target word. The phonetic parameter that plays the main role in this time window of our analysis appears to be duration. This result is generally in line with previous findings on the domain of accentual lengthening by, among others, Turk and Sawusch (1997) and Dimitrova and Turk (2012), who reported that the durational effect of an accent can extend beyond the stressed syllable in English. However, in those studies, the spillover of an accented syllable to an unaccented syllable seems to be in the forward direction in the temporal domain, that is, syllables produced after the accented syllable are affected. Katsika and Tsai (2019) showed for Greek that the durational spillover effect may affect adjacent syllables in both directions of the accented syllable, which is in line with our results. Recent psycholinguistic research shows that the pre-accentual (or immediately preceding pre-focal) syllable may play a role for focal accents. Clifton et al. (2022) replicated and extended the study of Cutler (1976) and found that a longer pre-accentual syllable contributes to the well-studied phoneme detection advantage associated with focal accents (comparable to our corrective focus). Lengthening of the pre-accentual syllable may thus be an effective tool that helps in the marking of focus. Future research on Persian should look at the domain of prominence-induced lengthening effects more closely.

In conclusion, the present study sheds light on the intricate nature of prosodic prominence in Persian. On the basis of our speech-production data, we conclude that durational modifications play an important role in prominence marking, alongside F0 and intensity as additional but less robust cues. For future research, we need to assess perceptual factors, to observe how Persian listeners respond to these three focus types. Are they able to differentiate the three focus conditions when manipulating the phonetic cues in a systematic way? How might the phonetic realization of each syntactic structure (the one with the verb "to pass" as a compound verb versus the one with "to give" as a simple verb) affect the recognition of prosodic focus in the perception experiment? Given the findings for the pre-target region, what role does this region play in the identification of focused elements? Therefore, we remain cautious, as the three-step pattern observed in certain phonetic cues in Persian will be further investigated through a perception experiment.

From a broader perspective on focus in Persian, there are specific morphosyntactic constructions in which the focused element cannot carry an accent, rendering F0-related cues unreliable. In such contexts, durational effects may become important (for more information, see Rahmani et al., 2018). In addition, the interaction between prosodic focus and morphological focus presents another intriguing area in Persian that currently lacks experimental investigation. It is typically assumed that if focus is indicated by morphological devices, prosodic focus becomes irrelevant. However, as claimed by Rahmani et al. (2018), this is not necessarily the case. They argue that prosodic focus marking remains obligatory, even in the presence of specialized morphological markers of focus, such as adverbs meaning "only." Addressing these issues in future research would be a significant advancement in our understanding of prosodic focus.[2]

Ultimately, we conclude that our study provides an important contribution to prominence marking in speech production by including different phonetic parameters, different locations of modulations, and a comprehensive analysis of fine-grained patterns of focus structure encoding.

Data availability

The data and analysis scripts of this study are openly available in Open Science Framework: https://osf.io/t5r7a/.

ORCID iDs

Mortaza Taheri-Ardali https://orcid.org/0000-0002-1260-6128
Simon Roessig https://orcid.org/0000-0001-6413-1761
Lena Pagel https://orcid.org/0000-0001-6745-8401
Doris Mücke https://orcid.org/0000-0002-6217-3121

Notes

1. Intensity measures are known to be highly influenced by recording settings and speakers. We used a head-mounted microphone and did not change the settings during the recording; hence, the intensity measures within each speaker are comparable. In addition, we used random intercepts and slopes to account for the differences between speakers.
2. We would like to acknowledge the second reviewer for their thoughtful suggestion to include this paragraph.

References

Abolhasanizadeh, V., Bijankhan, M., & Gussenhoven, C. (2012). The Persian pitch accent and its retention after the focus. *Lingua*, *122*(13), 1380–1394.

Alzamil, A., & Hellmuth, S. (2021, December 6–9). The realization of different structural focus conditions in Saudi Arabic dialects. In *Proceedings of the 1st International Conference on Tone and Intonation (TAI)*, 6–9 December 2021, Sonderborg, Denmark, (pp. 196–199). International Speech Communication Association (ISCA).

Andreeva, B., Barry, W. J., & Koreman, J. (2017). Local and global cues in the prosodic realization of broad and narrow focus in Bulgarian. *Phonetica*, *73*(3–4), 256–278.

Anonby, E., Taheri-Ardali, M., et al. (Eds.). (2015–2025). Atlas of the Languages of Iran (ALI). Geomatics and Cartographic Research Centre (GCRC). https://iranatlas.net

Baumann, S., Becker, J., Grice, M., & Mücke, D. (2007). Tonal and articulatory marking of focus in German. In *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)* (pp. 1029–1032). Saarbrücken, Germany: Universität des Saarlandes.

Baumann, S., Grice, M., & Steindamm, S. (2006). Prosodic marking of focus domains—Categorical or gradient? In *Proceedings of the 3rd International Conference on Speech Prosody (SP2006)* (pp. 301–304). Dresden, Germany: International Speech Communication Association (ISCA).

Boersma, P., & Weenink, D. (2001). PRAAT, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345.

Bürkner, P. C. (2018). Advanced Bayesian multilevel modeling with the R Package brms. *The R Journal*, *10*(1), 395–411.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Lee (Ed.), *Subject and topic* (pp. 25–55). Academic Press.

Chen, L., Wang, L., & Yang, Y. (2014). Distinguish between focus and newness: An ERP study. *Journal of Neurolinguistics*, *31*, 28–41.

Clifton, C., Rysling, A., & Bishop, J. (2022). The prosodic accent advantage in phoneme detection: Importance of local context. *Attention, Perception, & Psychophysics*, *84*, 244–259. https://doi.org/10.3758/s13414-021-02371-5

Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America*, *77*(6), 2142–2156.

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge University Press.

Cutler, A. (1976). Phoneme monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics*, *20*, 55–60. https://doi.org/10.3758/BF03198706

Dimitrova, S., & Turk, A. (2012). Patterns of accentual lengthening in English four-syllable words. *Journal of Phonetics*, *40*(3), 403–418.

Eslami, M. (2000). *The prosody of the Persian language and its application in computer-aided speech recognition* [PhD thesis]. University of Tehran [in Persian].

Féry, C., & Ishihara, S. (Eds.). (2015). *The Oxford handbook of information structure*. Oxford University Press.

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new, and focused constituents in German. *Journal of Phonetics*, *36*(4), 680–703.

Franke, M., & Roettger, T. B. (2019). *Bayesian regression modeling (for factorial designs): A tutorial*. PsyArXiv. https://psyarxiv.com/cdxv3/

Frota, S. (2000). *Prosody and focus in European Portuguese: Phonological phrasing and intonation*. Garland Publishing.

Gussenhoven, C. (2007). Types of focus in English. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and focus: Cross-linguistic perspectives on meaning and intonation* (pp. 83–100). Springer.

Hosseini, S. A. (2013). L1 interference in L2 prosody: Contrastive focus in Japanese and Persian. *Language and Information Sciences*, *11*, 55–67.

Hosseini, S. A. (2014). *The phonology and phonetics of prosodic prominence in Persian* [PhD thesis]. Tokyo University of Foreign Studies.

Jadoul, Y., Thompson, B., & de Boer, B. (2018). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, *71*, 1–15.

Kahnemuyipour, A. (2003). Syntactic categories and Persian stress. *Natural Language & Linguistic Theory*, *21*(2), 333–379.

Karimi, S. (2005). *A minimalist approach to scrambling: Evidence from Persian*. Mouton de Gruyter.

Katsika, A., & Tsai, K. (2019). The scope of prominence-induced lengthening in Greek. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 850–854). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, *55*, 243–276.

Kügler, F. (2008). The role of duration as a phonetic correlate of focus. In P. A. Barbosa, S. Madureira, & C. Reis (Eds.), *Proceedings of the 4th International Conference on Speech Prosody (SP2008)* (pp. 591–594). Campinas, Brazil: International Speech Communication Association (ISCA).

Kügler, F., & Calhoun, S. (2020). Prosodic encoding of information structure: A typological perspective. In C. Gussenhoven & A. Chen (Eds.), *The Oxford handbook of language prosody* (pp. 453–467). Oxford University Press.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press.

Ladd, D. R., & Arvaniti, A. (2023). Prosodic prominence across languages. *Annual Review of Linguistics*, *9*(1), 171–193.

Mahjani, B. (2003). *An instrumental study of prosodic features and intonation in modern Farsi (Persian)* [MA thesis]. University of Edinburgh.

Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation—Is it mediated by accentuation? *Journal of Phonetics*, *44*, 47–61.

Rahmani, H. (2019). *An evidence-based new analysis of Persian word prosody* [Doctoral dissertation, Radboud University Nijmegen]. LOT. https://www.lotpublications.nl/an-evidence-based-new-analysis-of-persian-word-prosody

Rahmani, H., Rietveld, T., & Gussenhoven, C. (2015). Stress "deafness" reveals absence of lexical marking of stress or tone in the adult grammar. *PLOS ONE*, *10*(12), 1–16.

Rahmani, H., Rietveld, T., & Gussenhoven, C. (2018). Post-focal and factive deaccentuation in Persian. *Glossa: A Journal of General Linguistics*, *3*(1), 1–36.

Repp, S. (2016). Contrast: Dissecting an elusive information-structural notion and its role in grammar. In C. Féry & S. Ishihara (Eds.), *The Oxford handbook of information structure* (pp. 270–289). Oxford University Press.

Roessig, S. (2023). *The inverse relation of pre-nuclear and nuclear prominences in German*. PsyArXiv. https://psyarxiv.com/upmka/

Roessig, S., & Mücke, D. (2019). Modeling dimensions of prosodic prominence. *Frontiers in Communication*, *4*, Article 44.

Roessig, S., Winter, B., & Mücke, D. (2022). Tracing the phonetic space of prosodic focus marking. *Frontiers in Artificial Intelligence*, *5*, Article 842546.

Royer, A. J., & Jun, S. A. (2019). Prominence marking in Kazan Tatar declaratives. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 3797–3801). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Sadat-Tehrani, N. (2007). *The intonational grammar of Persian* [PhD thesis]. University of Manitoba.

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

Sadeghi, V. (2017a). Word-level prominence in Persian: An experimental study. *Language and Speech*, *60*(4), 571–596.

Sadeghi, V. (2017b). The timing of pre-nuclear pitch accents in Persian. *Journal of the International Phonetic Association*, *49*(3), 305–329.

Sadeghi, V. (2022). The intonation of Persian polar questions. *Speech Communication*, *145*, 47–70.

Scarborough, R. (2007). The intonation of focus in Farsi. *UCLA Working Papers in Phonetics*, *105*, 19–34.

Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2017a). The role of F0 and duration in the identification of wh-in-situ questions in Persian. *Speech Communication*, *93*, 11–19.

Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2017b). The role of prosody in the identification of Persian sentence types: Declarative or wh-question? *Linguistics Vanguard*, *3*(1), 1–11.

Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2018). When is a wh-in-situ question identified in standard Persian? *Language, Cognition and Neuroscience*, *33*(9), 1168–1183.

Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2019). Do Persian native speakers prosodically mark wh-in-situ questions? *Language and Speech*, *62*(2), 229–249.

Taheri-Ardali, M., Rahmani, H., & Xu, Y. (2014). The perception of prosodic focus in Persian. *Proceedings of the 7th International Conference on Speech Prosody (SP2014)*, 20–23 May 2014, Dublin, Ireland, (pp. 515–519). International Speech Communication Association (ISCA).

Taheri-Ardali, M., & Xu, Y. (2012). Phonetic realization of prosodic focus in Persian. *Proceedings of the 6th International Conference on Speech Prosody (SP2012)*, 22–25 May 2012, Shanghai, China (pp. 326–329). International Speech Communication Association (ISCA).

Taheri-Ardali, M., & Xu, Y. (2015). An articulatory-functional approach to modeling Persian focus prosody. In M. Wolters, J. Livingstone, B. Beattie, & R. Smith (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences* (pp. 1–5). International Phonetic Association.

Towhidi, J. (1974). *Studies in the phonetics and phonology of modern Persian: Intonation and related features*. Buske.

Turk, A. E., & Sawusch, J. R. (1997). The domain of accentual lengthening in American English. *Journal of Phonetics*, *25*, 25–41.

Vahidian-Kamyar, T. (1972). *Suprasegmental features in Persian: Stress, intonation and a description of pause* [PhD thesis]. University of Tehran.

Xu, Y. (2011). Post-focus compression: Cross-linguistic distribution and historical origin. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhSXVII)*, 17–21 August 2011, Hong Kong, China (pp. 152–155). City University of Hong Kong.

Zimmermann, M., & Fery, C. (Eds.). (2010). *Information structure: Theoretical, typological, and experimental perspectives*. Oxford University Press.