A Comparative Study of Persian Synthesized Intonation Based on Autosegmental-Metrical and PENTA Models

Fatemeh Lotfi, Imam Khomeini International University, Iran
Vahid Sadeghi, Imam Khomeini International University, Iran

Abstract

This paper presents a comparison study concerning the level of naturalness of the manipulated Persian intonation based on AM and PENTA models of speech prosody. In this piece of study, the implementation rules in accordance with the aimed approaches are based on the pitch contours analyses. Intonation is used in all languages and it is linguistically structured and pragmatically meaningful. PENTA model is based on communicative and functional view of speech. AM model reflects the connection between two subsystems of phonology and views tone on separate tiers. A corpus containing 15 sentences was created with different types of focus on the content words. In production level, 12 male and female native speakers of Persian participants were assigned to produce and record the sentences. The data were analyzed and resynthesized via PRAAT software manually. Finally, 10 Persian native speakers were selected to judge the naturalness of the manipulated sounds of both AM and PENTA Models. After accomplishing the perception level, based on the quantitative results obtained from a chi-squared test (xi2) analyses, we argue that firstly, there is a significant difference between the natural speech intonation produced by the control group and the sampling groups. Although, the other test was conducted on two sampling groups comparing AM and PENTA models. The obtained results of its xi2 test indicated no significant difference between the Persian manipulated speech intonation based on Autosegmental-metrical approach parameters and PENTA.

Keywords: AM, PENTA, Intonation, Prosody, Speech Manipulation, Speech Synthesis

iafor

The International Academic Forum
www.iafor.org

Introduction

This paper is an attempt to discover the naturalness of the Persian synthesized speech intonation. It determines the manipulated prosody of speech intonation based on the two outstanding and state-of-art approaches Autosegmental-Metrical (AM) and Parallel Encoding and Target Approximation (PENTA), respectively. It scrutinizes the difference between the manipulation of the intonation with the help of data gathered and judged by native speakers. This piece of study would assist the better and conscious decision making in NLP and machine learning fields of study of Persian. This study will enhance our understanding of Persian speech intonation synthesis by aligning with international approaches in prosody. For this purpose, the components of the two approaches will be studied and evaluated on data and later the more efficient one on Persian will be determined. On the other hand, Persian language has not been under broader research. There is a substantial body of literature and numerous areas of ongoing research in this field. This study is significant as it validates the approach by demonstrating its effectiveness in an additional language, thereby enhancing its credibility.

In order to synthesize Persian intonation, systematic intonational units should be implemented based on the credited approaches, AM and PENTA, respectively. The primary concern that arises in this context is determining which of these options would yield more natural and higher quality intonation in Persian and Which analytical components in the two approaches take precedence over the other in the context of the Persian language.

Prior to commencing the research, based on previous studies, we hypothesize that intonation manipulated using the AM approach will exhibit superior quality compared to that produced by the PENTA approach.

In the present study, we have elected to implement a two-level investigation. The initial phase will focus on production and synthesis, wherein the intonational phrases of Persian will be analyzed in accordance with the components of both the Autosegmental-Metrical (AM) and PENTA (Parallel Encoding and Target Approximation) frameworks. A corpus of 15 sentences has been constructed. Each sentence incorporates both subjective and objective focus. The words have been carefully selected to include two or three syllables and voiced consonants, ensuring a smooth pitch contour without disruptions. The acoustic analyses of the utterances are conducted using PRAAT (Boersma, 2022). Using this software, we explore and manipulate the pitch changers of Persian intonation utterances by examining pitch tracks on the aimed focus. A pitch track demonstrates the fluctuations in the form of a curve over a span of time.

Intonation involves the systematic and contextually significant variation of fundamental frequency (F0) in speech. Intonation is used in all languages and is specified in phrasal level. The occurrence of pitch movement is based on a number of determining factors such as syntax, pragmatics, prosodic phrases. In linguistics, intonation is utilized to convey meanings and verbal information, while non-verbal features such as age and mood are not taken into account. There are said to be two main patterns for intonation: Rising and Falling patterns. The carry different contextual meanings in various utterances. Informative sentences follow the falling pattern. While in question, surprise, hesitation or emphasis we have the rising pattern. Consider Figure 1 as example from Persian, and observe the pitch contour as extracted using PRAAT. the question is:

A. /Qejmæt in tʃænd æst/? "How much do they cost?"
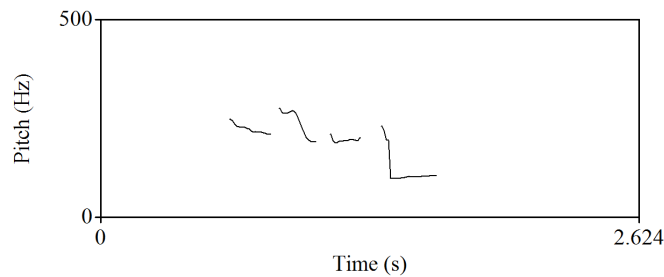B. /Pændʒah hezar Tomæn/. "50 bucks"

Figure 1: Pitch pattern for the Persian sentence: /Pændʒah hezar Tomæn/

Consider Figure 2 as example from Persian, and observe the pitch contour as extracted using PRAAT. the question is:

A. /Qejmæt in tʃænd æst/? "How much do they cost?"
B. /Pændʒah hezar Tomæn/? "50 bucks?"-we are shocked by the price or ask for emphasis.
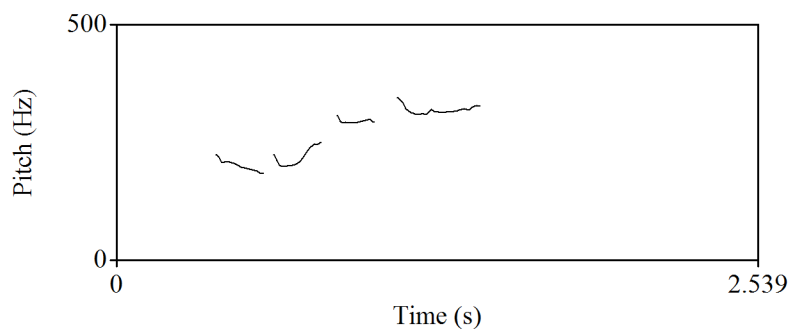

Figure 2: Pitch pattern for the Persian sentence: /Pændʒah hezar Tomæn/

Another feature that causes the pitch curve to alter its shape is prominence. The alteration in pitch accent results in the alteration of the contextual meaning. Consequently, it may contrast with other meanings of the sentence arguments. This point is also significant and merits attention.

There are two main approaches in regard to intonation, level and configurational analysis of speech intonation. The following paragraphs will be assertive of the named Analytic approaches.

AM Theory

The partisans of level-analysis of English intonation strongly believe that the pitch levels are relative (Boliger, 1951). As G.L.Trager and Henry Lee Smith, JR state this point explicitly: "It is relative pitch, not absolute, that is being discussed". However, it has not been enlightened how relative these relative pitches are thought to be (Boliger, 1951). The American structuralists believe that the intonation of a language should be defined using the four tonal levels: very high, high, mid, low (Pike, 1945). One of the scholars who criticized this perspective was Dwight L. Bolinger. He assumes that if pitch pattern had four different tonal realization, the entire pattern may be transferred to the higher or lower frequencies

without interfering inner relation of the four tones (Boliger, 1951). As a result, the pitch range would smoothly expand or compress. However, each tone is to have its specific pitch range. Leben pointed out that the tonal units have different tones from the segmental structure of a speech (Leben, 1976). Then, Goldsmith showed that each tone in phonological tier is independent from the segmental tier and is known as the auto-segmental tier (Goldsmith, 1976). Along with these studies, Liberman also approved his intonational studies and analyses on the auto-segmental tier (Liberman, 1975). These were the prelude for the birth of the Auto-segmental Metrical theory.

The most significant reason for the appearance of Autosegmental-Metrical theory was to define and explain the continuous pitch changes in speech base on the three PhD dissertations Liberman, Bruce (Bruce, 1977) and Pierhummbert (Pierrehumbert, 1980). The initial framework in AM theory started with works such as Liberman (Liberman, 1975), Bruce (Bruce, 1977), and Pierrehumbert (Pierrehumbert, 1980). They were all about speech intonation synthesis. They played a crucial role in the rise of the AM approach, specifically Bruce`s dissertation on Swedish. There were points stated in his work that is worth mentioning. First, the H and L targets of the F0 curve is in pitch contour level. Second, he proved that the whole curve of an utterance consists of two different tones with difference structures, the first group is the pitch accent that occurs on the word and then the boundary tone that happens with the end of the intonational phrase. The third point in his work is that tonal units in words and phrases are all put together to build up the whole contour. Pierhummbert, whose dissertation was based on the initial AM theory of Bruce, provided a comprehensive analysis of the English intonation. The findings of her dissertation were then developed by Beckman- Pierhummbert (Beckman & Pierrehumbert, 1986) and Pierhummbert-Beckman (Pierrehumbert, 1988). In such framework the tonal structure is composed of significant tonal events which are considered to be H (high) and L (low) pitch targets as well as their combination and they are considered way more important that the transition between the target points. This means the whole shape of the curve may not serve an importance throughout the analysis. Tones on the pitch track is phonetically aligned depending on different factors such as a focused element in an utterance. The smallest unit of Persian prosody is the accentual phrase (AP) with the phonological representation of L+H* related to the stressed syllable (Sadat-Tehrani, 2007). Intonational features are focus, tone, phrasing and pitch range. All the mentioned features correlate and the correlation is determined by two factors: prominence and pitch contour. Prominence comes in two terms: weak and strong. Pitch contour comes with two main definitions rising and falling pattern.

L and H tones are the abstract symbolic (i.e. phonological) primitives of intonation. Their identity as Hs and Ls is largely determined by phonetic observation and defined in relative terms: H is used to represent tones deemed to be high in a melody with respect to the speaker's range and other tones in the same contour; L is used to represent tones deemed to be low by the same criteria (Pierrehumbert, 1980). Tones are morphemes with pragmatic meaning; all tones in a melody contribute compositionally to the pragmatic interpretation of an utterance (Pierrehumbert, 1990).

Tones are *autosegments* that is they exist independently of the segmental string. They are phonologically associated to structural positions in the metrical representation of an utterance. It is this connection that has given rise to the term autosegmental-metrical. Phonetically, tones are said to be realized as *tonal targets* that is as specific points in the contour, while the rest of an F0 curve is derived by *interpolation* between these targets. tones associate either with phrasal boundaries or constituent heads. The last and most important of

pitch accents in a phrase is referred to as the *nuclear* accent. Tones that associate with boundaries are collectively known as *edge tone* (Arvaniti, 2022). The primary function of edge tones is to mark the boundaries of phrases. Phonetically, tones are said to be realized as *tonal targets* that is as specific points in the contour, while the rest of an F0 curve is derived by *interpolation* between these targets. Tonal targets are typically, turning points in the contour, such as peaks. It is very essential to understand the intonational structure for modelling.

Focus serves two meanings in linguistics. First, prosodically speaking, it refers to the most prominent element comparing to the others. Second, it refers to a part of the sentence in which an element has a new information and it is emphatic (Ishihara, 2001), (Vallduví, 1990), (Jackendoff, 1972). Focus carries pragmatic and contextual significance, and various languages employ different syntactic, prosodic, or semantic methods to express it and also some other may have all of these methods (Vallduví & Engdahl, 1996). *Focus* is categorized into three types: broad focus, contrastive focus, and narrow focus.

PENTA Theory

Configuration approach has been the center of attention, since it is fully based on the speaker`s intuition. The relation between the form and meaning is simple and natural. In this approach, meaning and communicational functions are depended on pitch curve and directly affects the changes in the pitch track. Accordingly, the whole shape of the curve changes in the pitch accent. Unlike AM, here the target tones and points are not playing the role. Prosodic parameters such as F0, duration, intensity defines the pitch changes. Ladd (Ladd, 2008) names this as parametric approach. As it is believed, in this type of analysis, the entire shape and form of the pitch curve is considered not just the target points in a specific time. In accordance to the configuration approach, there are three computational models of speech intonation synthesis: INTSINT[1] (Hirst et al., 2000), OXIGEN[2] (Grabe et al., 2004) and PENTA[3] (Xu & Xu, 2005). The Parallel Encoding and Target Approximation (PENTA) model of speech prosody was proposed as an attempt to improve the understanding of prosody by putting emphasis on two aspects of speech prosody communicative functions and articulatory mechanisms (Xu & Xu, 2005). The development of PENTA followed a different approach comparing to AM. It focused mostly on the meaning and communicative functions, importantly, it determines how prosody encodes the meaning in a way that one can decode it. Figure 3 is a diagram of PENTA in general. Not only does this schematic diagram show the prosody, but it also represents other aspects of speech.

---

[1] International Transcription System for Intonation
[2] Oxford Intonation Generator
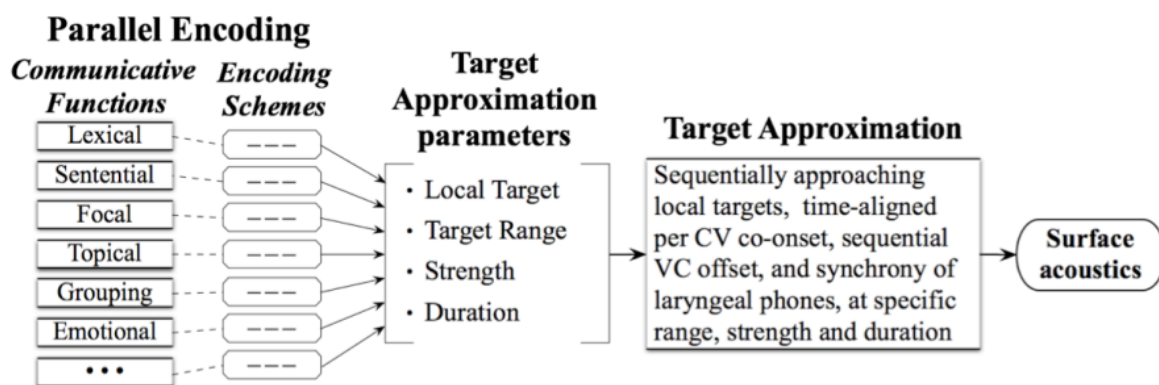[3] Parallel Encoding and Target Approximation

Figure 3: A schematic Sketch of PENTA model

In this model of intonation, the only obligatory melodic primitives are the syllable-sized pitch targets and this resembles the tone in AM theory (Xu et al., 2015). The phonetic characteristics of these targets include height, slope and rate of approximation.

In the subsequent sections, I will examine the existing Persian background studies and conduct a literature review.

This comparative study is the first research on Persian language. In the current work, I have conducted a comparison study on the naturalness of the speech intonation synthesis based on the two main intonational models AM and PENTA. Many Iranian scholars have studied Persian intonation using the AM framework.

Towhidi (Towhidi, 1974) did study on Persian intonation based on English school of thought. He applied the pragmatic meanings for determining the intonation pattern. He divided the speech into smaller tonal units that the smallest is a work and the biggest is a sentence. Each unit has a pitch accent and the boundary tone is defined by the change in tone. Mahootian (Mahootian & Gebhardt, 1997) conducted a study which examines the syntactic view point of the Persian intonation. Hayati (Hayati, 1998) conducted a comparative study on Persian and English intonation to gain a comprehensive understanding of the challenges Persian speakers face in producing English intonation patterns. Eslami (Eslami, 2000) performed an introductory study on Persian intonation based on Autosegmental-Metrical (AM) theory in his PhD dissertation. He has comprehensively defined Persian intonation. Within this study, he introduced *the head avoidance principle*. This principle provides insights into the syntactic and derivational characteristics of an accent within a phrase or utterance. Vahidiyan (Vahidian-Kamyar, 2001) has studied on semantic units of Persian intonation. He does not consider the syntactic features throughout his studies. He categorizes speech intonation into semantic groups as words, phrases, small sentences and utterances. This categorization was applied utilizing English school. Mahjani (Mahjani, 2003) studied Persian intonation based on AM theory in his MA thesis. in his studies, he grouped Persian intonation into three: intonational phrase, accentual phrase and intermediate phrase. Sadat Tehrani (Sadat-Tehrani, 2009) has a more comprehensive PhD study on Persian intonation based on Autosegmental-Metrical theory. In his research, he recorded data of 2100 utterances produced by Persian speakers and analyzed it using AM framework of intonation. In his thesis, it was proposed that the smallest unit of Persian prosody is the Accentual Phrase (AP), with the pitch accent L+H* associating with the stressed syllable. Sadeghi (Sadeghi, 2018) has done valuable research on Prosodic features of Persian. In his book- The prosodic structure of the Persian

language- he has delved into the intonational approaches and frameworks and has analyzed different methods.

In summary, there are two main analytical approaches for intonation of a language that was discussed earlier. Each served as a foundation for the major intonational models, Autosegmental-Metrical (AM) and Parallel Encoding and Target Approximation (PENTA).

Methods

This study is a primary and experimental quantitative work. The subjects were wisely chosen and they were all monolingual native speakers of Persian. They all reside in Tehran.

The first step was to customize a corpus for the study. I designed 15 sentences all including 3 simple arguments, subject, verb and object. All the accented syllables were voiced so as not to have a teared pitch signal. In order to gather the data, questions were designed according to the intended answers that were expected to be produced by the subjects. 15 utterances were articulated in 5 forms of focus: broad focus, subjective contrastive focus, objective contrastive focus, subjective narrow focus and objective narrow focus. All the produced data was uploaded in PRAAT for the manipulation process. There were 7 men and 11 women. I excluded 6 speakers due to the bad quality and noises that was recorded unintentionally. 12 recorded voices were selected. Recording phase was conducted by Boya microphone to denoise the environment. Data preparation, trimming and annotation has been done using PRAAT afterwards. Annotations were according to the framework of the two separate models of intonation AM and PENTA. The files were saved instinctively by special codes in naming so would be used in trainer software of both theories. After creating data text grids for the AM trainer and PENTA trainer, we had to import them into the trainers and start synthesizing.

It was required to create two tiers for annotating the data according to AM model. In the first tier, I marked the syllable boundaries and named them with lower case letters by order. The second tier was created to be a tier for marking the pitch accents on the accented syllable. Figure 4 demonstrates annotated tiers for the Persian sentence: "Mary watered the plants."
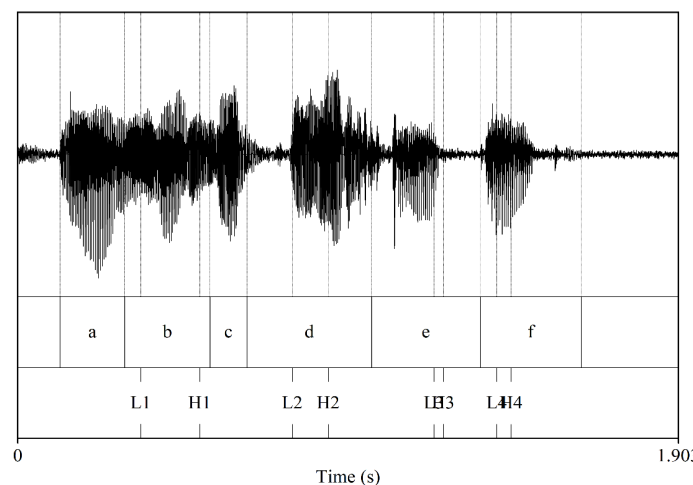

Figure 4: Annotating syllabic and tonal tiers in Persian sentence
/mærjæm be gɔl ab dad/

The files were saved as Wav.

For annotating the data according to PENTA framework, just one syllabic tier is marked. As xu believes in this approach the whole form of the pitch curve serves importance throughout the syllable, it is not just a tonal target in temporal alignment. Figure 5 illustrates the annotated tiers for the Persian sentence: "Ayda is a proficient speaker of German."
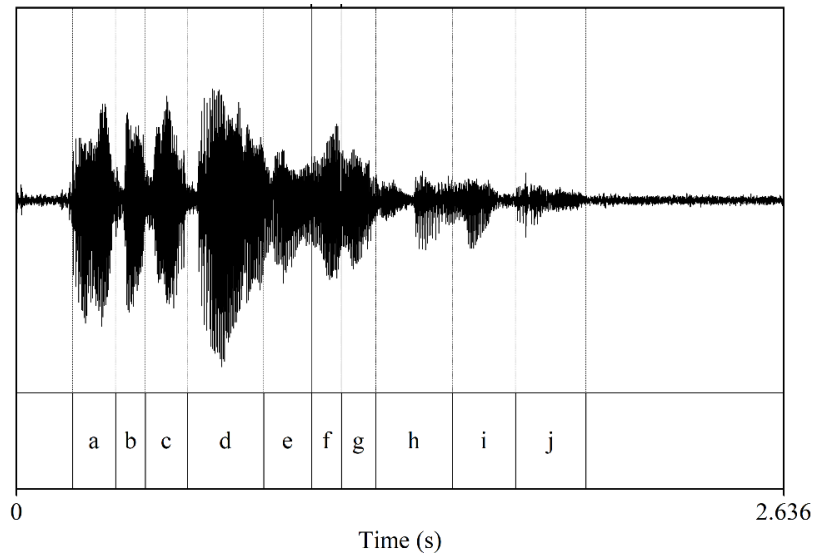

Figure 5: Annotating syllabic and tonal tiers in Persian sentence
/ajda be almani mɔsælæt æst/

The same process was conducted for the second set of data.

Input data was ready to be loaded in trainer programs. On the process some problems came up and avoided the machine syntheses. We changed the plan for manipulation process and synthesized manually in PRAAT. So, we can retreat to the problem and find a solution for the drawbacks.

Having the AM framework in mind, I followed the order below to first stylize the speech signal.

Manipulate sound → View and edit → Pitch → Stylize pitch

The resolution was set as 2 semitones. Following the structure above will help better and smoother transition of the target points on a pitch track (Boersma, 2022). The manipulation is done on the broad focus and the focus is being created artificially by displacing the target tones. All the accents on the other arguments were deaccented where needed. The increase in target point was about 30 Hz on each aimed tone. After each manipulation, I followed the structure below:

Pitch Tab → Resynthesize → Interpolate quadrable

PRAAT is using a linear interpolation between the two starting and final points in all tiers as: Pitch Tier, Intensity Tier, Duration Tier, Amplitude Tier, Formant Grid. The last version of the data was stored as published resyntheses sounds. Figures 6 and 7 show the synthesized signals for the sentence: "Ayda didn't respond the teacher." in two forms of subjective manipulation and objective manipulation.
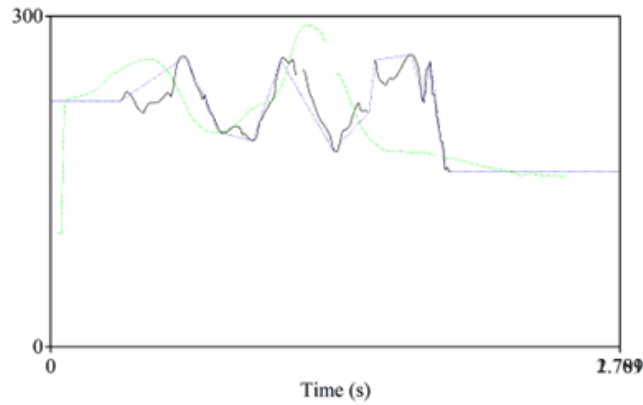
Figure 6: resynthesized informative focus on object in Persian sentence
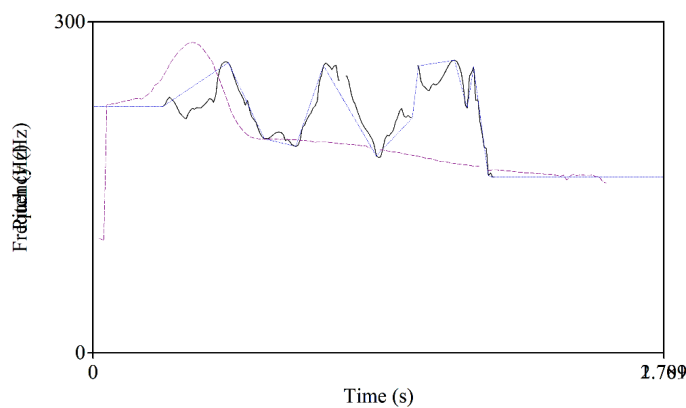/Ajda be mɔælem jævab nædad/

Figure 7: resynthesized informative focus on subject in Persian sentence
/Ajda be mɔælem jævab nædad/

The green dotted line in Figure 6 is the synthesized intonation pitch contour. The purple dotted line in Figure 7 is the synthesized intonation pitch contour.

All the above-mentioned process is done for PENTA model manipulation. There is just one step that is not appropriate and is ignored is to stylize. As PENTA framework is considered, no target points and tones are aimed here but the entire signal in the syllable is of the value.

Figures 8 and 9 are the synthesized sentences according to the PENTA intonation model frameworks.
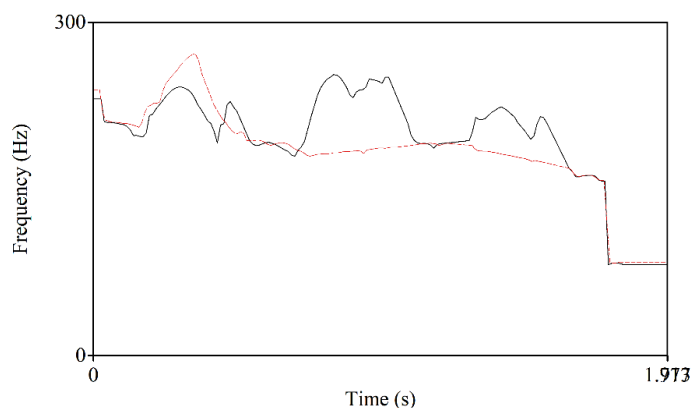
Figure 8: resynthesized contrastive focus on subject in Persian sentence
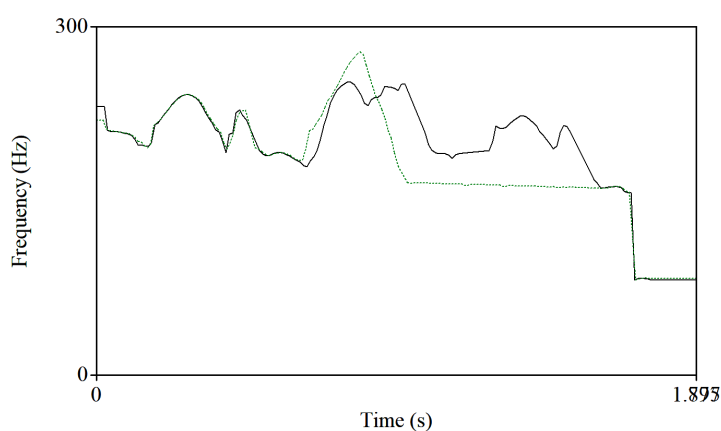/baba æz bazi esteqbal kærd /

Figure 9: resynthesized contrastive focus on object in Persian sentence
/baba æz bazi esteqbal kærd /

In these figures, we only have one colorful dotted line due to the ignorance of the stylizing level according to PENTA framework.

The next level after finishing the manipulation process is to combine the questions that were actually asked in the interview with the resynthesized sounds. Followed:

Select both Sound files in object list using ctrl→ Combine → Concatenate

Perception phase was conducted by *Questionnaire design*. I designed a five scaling LIKERT questionnaire in which the listeners would judge the level of naturalness of the sounds they heard. The evaluated five scales were:

very natural- almost natural- natural- almost unnatural- very unnatural

Ten Persian native speakers were selected as the judgers for the perception phase. Files were played and questionnaire was distributed to be marked.

Results

We compared the results of the different data gathered from both group of questionnaires-AM and PENTA- and graded them using excel. We used the Chi-square statistical test (X^2) for the final analyses of the data. There are three groups to be studied in this paper: the control group is the natural interview that was conducted at the very first step, second group is the synthesized data based on AM and the third group is the synthesized data based on PENTA.

In the first place, I compared each synthesized group with the natural production group. The gained results are as below:

The p level in X^2 test for the speech intonation synthesized according to AM.

X2= 183.74, $p < 0.001$

The p level in X^2 test for the speech intonation synthesized according to PENTA.

X2= 218.28, $p < 0.001$

According to the obtained results, there is a significant difference between the natural intonation of speeches and the synthesized ones.

Sig= 0.001< 0.05

Comparing the two synthesized sentences based on AM and PENTA.

X2= 2.713, $p = 0.52$ (Sig= 0.52> 0.05)

As observing the significant level number of the test, we conclude that there is no considerable difference between the perceived frequency of the data based on AM or PENTA. Generally speaking, the average naturalness of the synthesized intonation in each focused utterance is as below:
- The average of naturalness in the naturally produced utterances: 98.67%
- The average of naturalness based on AM synthesized output: 81.17%
- The average of naturalness based on PENTA synthesized output: 84.85%

Conclusion

This research aimed to determine whether there is a significant difference between Persian synthesized speech intonation manipulated using Autosegmental-metrical approach parameters and PENTA. Additionally, the study sought to assess the naturalness of the manipulated intonation. The findings of this study are intended to encourage further research in Persian language studies within the rapidly advancing field of Natural Language Processing (NLP). Intonation refers to a combination of acoustic parameters including duration, intensity and pitch to distinguish lexical items. The most important parameter, pitch, is determined by the fundamental frequency, or F0, of speech, and is measured in hertz (Hz) or semitones. Intonation here is considered a component of a language's prosody, which is an overarching term that includes interacting elements such as rhythmic structure, prominence, and prosodic phrasing. In order to have a synthesized intonation of a language

we need to follow authentic frameworks. There are models of speech intonation syntheses. In this study, we applied AM and PENTA models of intonation.

So, to run the project, we implemented rules based on pitch contour analyses using specialized software. Initially, a corpus of 15 sentences was created, each with different focuses on content words. The sentences were carefully selected, ensuring that the content words were voiced consonants to maintain the integrity of the pitch contour. This approach aimed to produce precise and fruitful results. Subsequently, 12 native Persian speakers (both male and female) were tasked with producing and recording the sentences. The recorded data were then manually analyzed and resynthesized using PRAAT software. Finally, 10 native Persian speakers evaluated the naturalness of the manipulated sounds for both the AM and PENTA models under identical conditions. The selected subjects were tasked with evaluating whether the responses matched the questions meaningfully and naturally. Following the perception phase, quantitative results from a chi-squared (χ^2) test indicated a significant difference between the natural speech intonation produced by the control group and the sampling groups. However, when comparing the two sampling groups (AM and PENTA), the χ^2 test results revealed no significant difference between the Persian manipulated speech intonation based on Autosegmental-metrical approach parameters and PENTA.

References

Arvaniti, A. (2022). The Autosegmental-Metrical model of intonational phonology.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. Phonology Yearbook, 3, 255–309. *Bolinger, Dwight (1986) Intonation and its parts, Edward Arnold*.

Boersma, P. W., David. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.14. http://www.praat.org/

Boliger, D. (1951). Intonation – levels vs. configurations. 7:199-200.

Bruce, G. (1977). Swedish Word Accents in Sentence Perspective. *Travaux de l'Institute de Linguistique de Lund/CWK Gleerup*.

Eslami, M. (2000). "Šenaxt-e nævay-e goftar-e zæban-e farsi væ karbord-e an dar bazsazi væ bazšenas-yerayan'i-ye goftar." The prosody of the Persian language and its application in computeraided speech recognition] [PhD dissertation].

Goldsmith, J. A. (1976). *Autosegmental phonology* Massachusetts Institute of Technology].

Grabe, E., Kochanski, G., & Coleman, J. (2004). Quantitative modelling of intonational variation.

Hayati, A. M. (1998). A contrastive analysis of English and Persian intonation. *Papers and studies in contrastive linguistics*, *34*, 53-72.

Hirst, D., Di Cristo, A., & Espesser, R. (2000). Levels of representation and levels of analysis for the description of intonation systems. Prosody: Theory and experiment: Studies presented to Gösta Bruce.

Ishihara, S. (2001). Stress, focus, and scrambling in Japanese. *MIT working papers in linguistics*, *39*, 142-175.

Jackendoff, R. S. (1972). Semantic interpretation in generative grammar.

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

Leben, W. R. (1976). The tones in English intonation. *Linguistic analysis*, *2*(1), 69-107.

Liberman, M. Y. (1975). *The intonational system of English* Massachusetts Institute of Technology].

Mahjani, B. (2003). An Instrumental Study of Prosodic Features and Intonation in Modern Farsi (Persian).

Mahootian, S., & Gebhardt, L. (1997). Persian (descriptive grammars). *London: Routledge. doi*, *10*, 9780203192887

Pierrehumbert, J. (1990). Phonological and phonetic representation. *Journal of phonetics*, *18*(3), 375-394.

Pierrehumbert, J., and Mary Beckman. (1988). "Japanese tone structure." Linguistic inquiry monographs 15. 1-282.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* Massachusetts Institute of Technology].

Pike, K. L. (1945). The intonation of American english.

Sadat-Tehrani, N. (2007). The intonational grammar of Persian.

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. (2018). The prosodic structure of Persian intonation.

Towhidi, J. (1974). Studies in the phonetics and phonology of modern Persian: intonation and related features. *(No Title)*.

Vahidian-Kamyar, T. (2001). Næva-ye goftar dær farsi [Melody of speech in Persian]. In: Mashhad: Ferdowsi University Press.

Vallduví, E. (1990). The informational component.(Doctoral dissertation). In: University of Pennsylvania New York.

Vallduví, E., & Engdahl, E. (1996). The linguistic realization of information packaging.

Xu, Y., Lee, A., Prom-On, S., & Liu, F. (2015). Explaining the PENTA model: a reply to Arvaniti and Ladd. *Phonology*, *32*(3), 505-535.

Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of phonetics*, *33*(2), 159-197.

Contact email: Fatemehlotfi1992@gmail.com