THE INTERACTION PATTERN OF FOCAL ACCENT AND DECLARATIVE INTONATION IN MONGOLIAN

Aomin[1] Aijun Li[2]

[1]Inner Mongolia University, China
[2] University of Chinese Academy of Social Sciences (UCASS) and IL, CASS, China
aomin@imu.edu.cn liaj@cass.org.cn

ABSTRACT

Based on a rigorous experimental corpus, this study explores the interaction pattern of focal accent and declarative intonation in Mongolian, spoken in Inner Mongolia Autonomous Region, china. The results suggest that the narrow foci at the initial, middle, and final position of the sentence are all encoded with the F0 top line, but in varied pitch properties and realization domains. The declarative intonation function is rhythmically encoded through F0 bottom line. Meanwhile, the intonation interacts with the focus function during the encoding process.

Index Terms— Mongolian, declarative intonation, focal accent, interaction pattern

1. INTRODUCTION

Focal accent is a central issue in the study of intonation function. The study of focal accent has been a hot issue of common interest. It is a component emphasized to achieve semantic and pragmatic goals [1]. Besides, it is an instructional approach to aggravate or accentuate a part of a sentence relevant to the information structure [2]. Studies on focal accent in many languages prove that focal accent can be realized through phonological, syntactic, and grammatical forms. Moreover, due to the vital role of focus in speech comprehension, a body of psychological research has also explored focus processing and related brain mechanism during speech comprehension [3]. As widely acknowledged, there are two types of focus: broad focus (the sentence focus is multiple words or entire phrases) and narrow focus (the sentence focus is a specific word) [4].

However, there is little research on focal accent in Mongolian, an SOV language, one of three groups within the Altaic language families. Existing research is concerned with phonology and phonetics from a linguistic perspective. Karlsson used the Autosegmental-metrical theory (AM theory) to examine the pattern of focal accent realization and bearing units in the Mongolian Khalha dialect. He found that (1) the focal accent was mainly realized by pitch range change [5]; (2) the unit of focal accent was one word or the first two moras of a phrase [6]. Synthesizing the overall pattern of Post Focus Compression (PFC) in Altaic Uyghur, Mongolian, and Turkish, the study concluded that Mongolian was a language subordinated to PFC [7].

Two main research methods are commonly used in the study of the Mongolian declarative intonation: intonation description based on descriptive linguistics and acoustic analysis. Two perspectives, including phonology and syntax are usually involved as well. Despite the diversity of research methods and perspectives, conclusions are generally consistent, that is, the end of the Mongolian declarative intonation shows a flat or falling pitch pattern [8]. Acoustic analysis of the Mongolian declarative intonation also indicates that the sentence-final pitch range is somewhat narrower than that of the sentence-initial [9].

The limited research of Mongolian intonation so far has not reached a consensus in many key topics, and therefore leads us to the following questions: (1) The difference among various types of focal accent (broad and narrow focus) in terms of acoustic characteristics. Acoustic analysis of focal accent shows that it is rhythmically encoded and realized through the change of pitch, duration, intensity and energy [10], and pitch plays a decisive role in the prosodic realization of focal accent [11]. However, scholars differ in their views on the prosodic encoding of focal accent with regard to duration and energy. (2) Whether the realization of focal accent is subject to the focus type and position in the sentence. Although it has been demonstrated that Mongolian focal accent conforms to the " Tri-zone Realization of Focus " model [6], the prosodic encoding of focal accent among different dialects may display dissimilarities even within the same language [12]. Thus, it is necessary to investigate the case of Mongolian in a broader view. (3) The function scope of focal accent and declarative intonation. Apart from lexical meanings, a sentence also contains intonational information. Intonation can distinguish different tones and prosodic groups, highlight or weaken certain components, convey focus information. These functions can be local functions of the sentence or global functions that affect the whole intonation pattern of the sentence. Therefore, the domain that the interaction between the focal accent and Mongolian declarative intonation affects is a topic worthy of deeper investigation.

Three issues were examined: (1) the pitch encoding mechanism of focal accent; (2) the similarities and differences between the encoding of declarative intonation and focal accent in terms of pitch (F0); and (3) the scope of declarative intonation and focal accent.

2. CORPORA

When designing the experimental sentences of Mongolian focal accent, characteristics of Mongolian as an agglutinative language, as well as the segmental variation of long and short vowels in spoken language, were taken into account. Short vowels of Mongolian are inclined to be weaken and shed, while long vowels maintain a more stable segmental articulatory feature in different syllable positions [13][14]. Thus, four factors were considered in the corpus design: word length (monosyllabic and disyllabic), vowel type (all vowels in the experimental sentences were long vowels), sentence structure (only designed a sentence composed of three words in the Subject-Object-Verb sequence), syllable type (all syllables of all words were open syllables with a Consonant+Vowel structure). Finally, two experimental sentences were designed. Table 1 provides details of the designed corpus. Based on two experimental sentences, a broad/narrow focus (at initial, middle, and final of the sentence, respectively) was elicited by the carrier sentences. Twenty speakers (ten males and ten females) from standard Mongolian-speaking regions were asked to read the sentences three times in a row in a random order. 480 (2 sentences * 4 focus types * 20 speakers * 3 repetitions) sound samples were finally obtained.

Table 1. Experimental sentences

No.	ABBV	Target sentence	IPA	Translation
1	S1	[illegible]	tʰɐː suː sɐː	You Milk
2	S2	[illegible]	nɪmɐː moːk tuːnəː	Nima picks mushrooms

The recording tool was the WeChat application "Jiuzhou Sound Collection"[15]. The recordings were in a question-and-answer matching format. For each recording, the tool displayed a question and a declarative answer to the question on the screen and presented them randomly. The parts to be emphasized were bold in the text. Each speaker was asked to read both the questions and the declarative answers. The PENTA Trainer2 [16] was used to annotate intonation of the sound files. Four tiers were annotated: vowel type within a syllable (L for long vowels, S for short vowels), syllable position in the word (1 for word-initial syllables and 6 for the word-final), focus type (Pre for pre-focus, On for on-focus, Post for post-focus domain, and Broad for broad focus), and sentence type (S for declarative and Q for question). Figure 1 shows an annotated example.

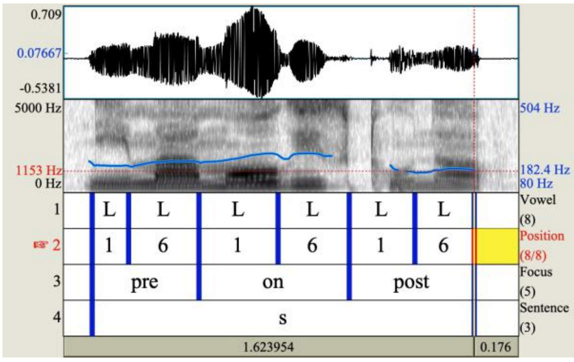
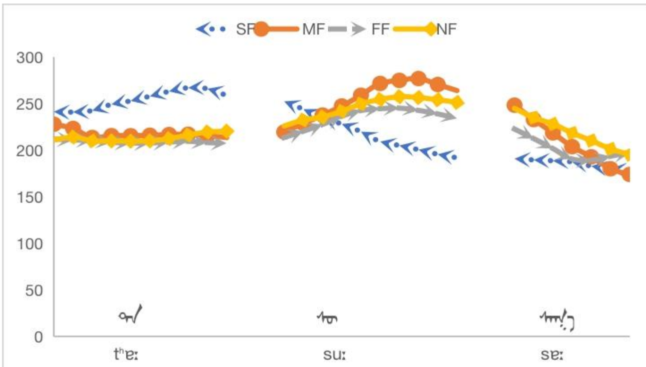


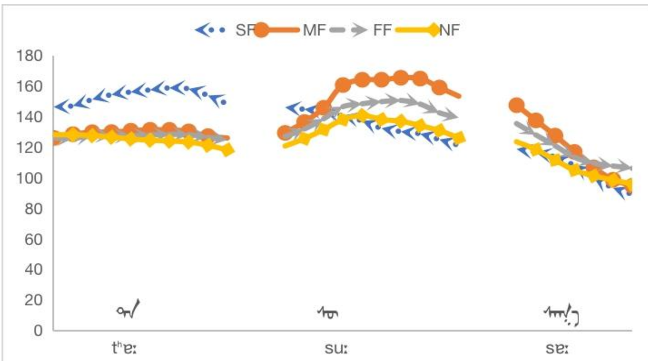
Figure 1.An example of Mongolian intonation annotation

3. ACOUSTIC ANALYSIS

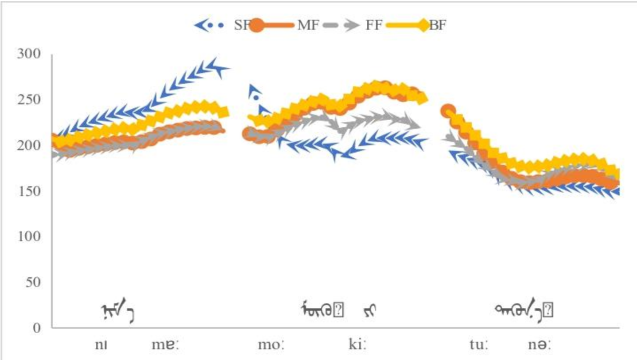
The acoustic analysis was conducted according to gender. We extracted time-normalized F0 value at 10 points per syllable to analyze the pitch characteristics of the focus with the help of PENTAT_trainer 2. The F0 value for each point was the average of all recordings by each gender. Pitch patterns of two experimental sentences under four focus types are plotted in Figure 2. The SF, MF, FF,and NF in Figure 2 represent for initial focus,middle focus,final focus and broad focus respectively. S1 and S2 represent for two sentences as listed in Table 1. When analyzing the pitch properties of the focus function, the sentences were segmented into three parts: pre-focus, on-focus, and post-focus. The pitch pattern of each part was analyzed as well.



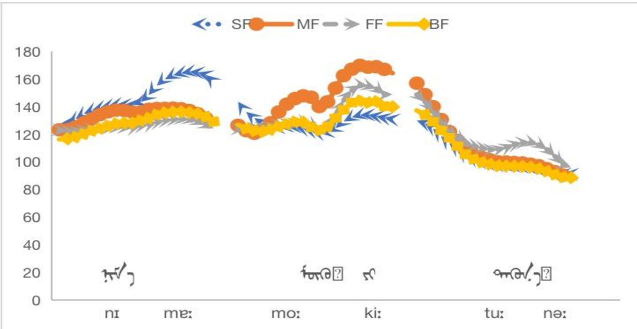
a. S1-Female



b.S1-male



c.S2-Female



d. S2-Male

Figure 2.Pitch patterns of two experimental sentences under four focus types

3.1. On focus

In the focal domain, the pitch line slowly rises from the syllable onset position and peaks near the middle of the on-focus domain. There are two types of pitch range. The first type is foci excluding the sentence-final focus. The pitch range of focal word widens substantially. The degree of pitch expansion is not affected by syllable numbers. Moreover, the pitch value reaches the highest of the whole sentence and decreases sharply, which is also known as the phenomenon of Post Focus Compression (PFC). The second type is the sentence-final focus. When the focal word is located at the end of the sentence, the pitch contour of the monosyllabic word lowers in the first half of the syllable and slightly rises in the second half. This pattern of pitch change is basically congruent in monosyllabic and bi-syllabic words. The only nuance is that when the narrow focus falls on monosyllabic words, the rising of pitch contour occurs in the second half of the syllable; while in the case of bi-syllabic words, pitch contour changes occupies the entire second syllable. However, the pitch change triggered by the sentence-final narrow focus is extraordinarily different from the pitch patterns of the initial and middle focus sentences. With respect to the distribution of MaxF0 in the entire sentence, the MaxF0 of initial and middle focus occurs on the focused words, while in broad focus, the MaxF0 occurred on the second word. No MaxF0 value is reported to occur in the focal domain for final focus, which we suspected, is influenced by the declarative intonation. One-way ANOVA comparisons of MaxF0 values for words on broad and narrow foci at different positions are shown, MaxF0 values of broad focus words and all narrow focus words are significantly different, with p-values < 0.05 or < 0.01 for all focused words except for the sentence-final narrow focus words. MaxF0 values of the second syllable is more significant than that of the first syllable in focal bi-syllabic words.

3.2. Pre focus

Three pre-focus patterns are found: (1) The initial syllable of sentence-middle focus sentences (the bi-syllable of S2); (2) the first two syllables of sentence-final focus sentences (the first four syllables of S2); (3) the first syllable of the broad focus sentence.

As can be seen from Figure 2, the pitch contour of the pre-focus domain is the same as the counterpart of the broad focus sentence. No remarkable change in the pitch pattern is found as well. It is interesting to note that the pitch curve is not affected by the pre-focus domain length (the number of words and syllables). One-way ANOVA was used to examine the MaxF0 values of pre-focus domains in two experimental sentences. The results show no significant difference between MaxF0 in the pre-focus domain of narrow and broad focus sentences. Also, there is no significant difference between the first words in the final focus and its counterpart in the broad focus ($p>0.05, df=39, F=3.486$). However, there is a statistical difference between the second word ($p<0.05, df=39, F=11.508$). The post-hoc analysis of words in the pre-focus domain of S2 indicated no statistical differences ($p>0.05, df=39, F=1.363$) between the pitch values of two syllables within words and between monosyllabic words in different narrow foci.

3.3. Post focus

The post-focus domain can be regarded as three types: (1) the last two words in the initial focus sentence (the last 4 syllables in S2); (2) the last word in the middle focus sentence (the last 2 syllables of S2); (3) the last word in the broad focus sentence (the last 2 syllables of S2). In terms of the overall pitch pattern, the pitch in the post-focus domain showed a decreasing trend. Meanwhile, there are dissimilarities among three post-focus types: the pitch of the final focus sentence shows a changing pattern of LHL. The pitch onset of the target word increased or dipped relatively slow, followed by a drop at the word-middle. One-way ANOVA was used to examine the MaxF0 values of post-focus domains in two experimental sentences.

The result indicates that (1) there is a remarkable post-focus compression (PFC). Regardless of the gender and the sentence length (word length), the MaxF0 of all syllables in narrow focus sentences shows an obvious compression compared to its counterpart in broad focus sentences; (2) the degree of the F0 compression for each syllable in the post-focus domain varied. Generally, the pitch range gradually compressed on each syllable sequentially. In other words, for the syllable in the post-focus domain, the compression degree is correlated to its position. The closer it is to the preceding focus, the greater its pitch range gets compressed. (3) post-focus compression may be word-based. When the post-focus domain is two bi-syllabic words, there is a colossal compression on the first syllable of each word but not on the second syllable.

To scrutinize the acoustic information, bi-syllabic words were converted to Semitones after measuring the MaxF0 values on a word-by-word basis, as shown in Table 2. The values are the average pitch of all utterances we collected after temporal normalization. The reference frequency for a semitone is 1Hz. The post-focus domain is hugely compressed in sentence-initial/middle narrow focus sentences, compared to broad focus sentences.

Table2. Temporal normalized MaxF0 (St)values of the post-focus domain in different focus environments

Sentence	Speaker	Narrow focus		Broad focus
		initial	middle	
S1	Female	98.29	98.53	94.22
	Male	85.84	85.48	83.18
S2	Female	98.03	96.52	95.54
	Male	86.44	85.32	84.14

4. DISCUSSION

This paper presents a statistical analysis of the F0 value based on the focal accent of Mongolian declarative sentences. It shows that the focal accent of Mongolian declarative sentences conforms to a Tri-zone Realization of Focus. The pre-focus, on-focus, and post-focus domains are encoded with the change of intonation to realize the communicative function of focal accents. Based on experimental results, three issues raised in the introduction can be further explored.

4.1. F0 encoding mechanism of different focus types

According to pitch patterns and MaxF0 values of different focus types in Mongolian, it can be seen that the encoding mechanism of narrow focus differed when it appeared at different positions compared to broad focus sentences.

4.1.1. Narrow focus at the sentence-initial

Although the pitch range expansion of the narrow focus at the sentence-initial is more salient than that of the broad focus (12 Hz for male & 14 Hz for female), it is the smallest compared to that of the narrow focus at the middle and final. Among three narrow focus positions in the sentence, the most significant pitch range compression was found in the post-focus domain in the sentence-initial one (varying between 54-93Hz). Therefore, pitch range compression between on-focus and post-focus domains is most likely the best significant F0 feature during prosodic encoding.

4.1.2. Narrow focus at the sentence-middle

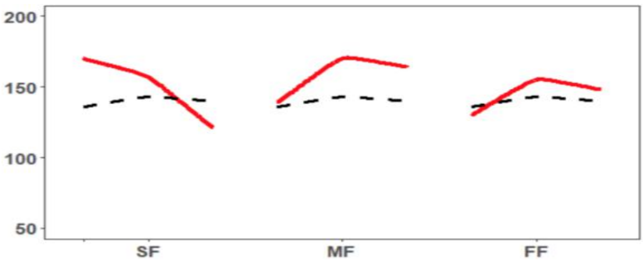
The on-focus domain expands for an average of 36Hz (for monosyllabic words) and 52Hz (for di-syllabic words) for male speakers, and 57Hz (for monosyllabic words) and 53Hz (for di-syllabic words) for females compared to the broad focus. It should be noted that the pitch range of sentence-middle focus is significantly different from the sentence-final focus ($p<0.01$, df=3, F=5.5864). However, in terms of pitch range expansion, no similar difference is found ($p>0.05$, df=3, F=6.2952). In addition, regardless of narrow focus position, the MaxF0 values of on-focus domains were generally high, although the difference between the MaxF0 of initial and middle of the sentence was smaller than final cases. Overall, the most significant feature of narrow focus in the sentence middle was the co-encoding of the MaxF0 value and pitch range expansion.

4.1.3. Narrow focus at the sentence-final

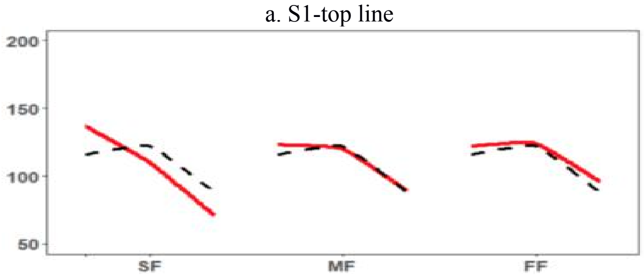
With respect to MaxF0 values of the focal domain, the MaxF0 value of the final focus sentence is significant difference ($p<0.01$,df=39, F=3.175) from those of the initial and middle positions, and it is also significant difference ($p<0.05$,df=39,F=2.124) compared to the broad focus. Meanwhile, when it comes to the overall pitch pattern of the whole sentence, there is little difference among all narrow focus sentences. In terms of pitch range change, the sentence-final focus had the widest pitch range. We must note that the pitch contour of the whole sentence would be regulated by the declarative intonation. And therefore, the sentence-final domain shows a downward slant no matter whether it is focused or not. Under this premise, it is not easy that the whole sentence with a sentence-final focus to shows a pitch pattern of LHL. The listening and discriminating experiments of the sentence-final focus also proved that this pitch encoding scheme made an important contribution to the decoding process of the sentence-final focus [17]. Thus, with regard to pitch properties, the pitch change pattern of LHL in the on-focus domain is the most significant prosodic encoding feature among all encoding features of the sentence-final focus.

4.2. Interaction patterns of declarative intonation and focal accent

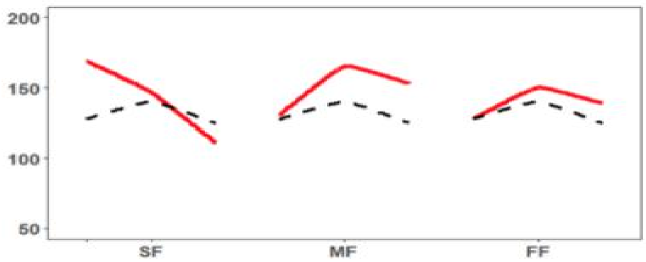
The parallel encoding between declarative intonation and focal function on phonological level is realized as a superficial pitch curve. Thus, we can attribute acoustic features to intonation functions. Acoustic analysis enables us to examine the interaction pattern of multiple intonation functions through pitch curves of the sentence. Acoustic studies on focal accent[18] have concluded that intonation has a modulatory effect on the top and bottom lines of the pitch, that is, the modulatory change in the top line is associated with semantic intensification while in the bottom line is associated with the integrity of the prosodic structure. In order to examine the functional role of the top and bottom lines of Mongolian intonation, Figure 3 presents the pattern of top and bottom lines in two experimental sentences, both of which are under the regulation of the focus.



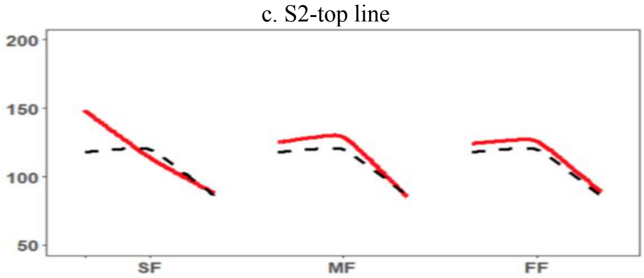
a. S1-top line



b. S1-bottom line



c. S2-top line



d. S2-bottom line

Figure3 Top line and bottom line under the focus function modification for males

Data in the figure was obtained from the average pitch value of all recordings from 10 male speakers. Top lines are plotted based on the MaxF0 value of each syllable in the sentence. Similarly, bottom lines are composed by MinF0 values. The dashed line in the graph represents the intonation of broad focus, and the solid line is the pitch pattern of the narrow focus. SF, MF and FF are the narrow focus in the initial, middle and final position respectively. Patterns of female speakers which are the same as those of males are ignored here due to space limitations, following conclusions can be drawn from Figure 3.

4.2.1. The focal function is mainly encoded by the F0 top line

The F0 top line trend is closely related to the focus, directly reflecting the focus position. Specifically, the pitch curve of the whole sentence showed a downward slant with the narrow focus at the beginning. When the narrow focus was at the sentence-middle or final, the pitch curve presented an LHL pattern. Pitch lines of broad focus sentences also showed the LHL pattern. Additionally, the longer the sentence is, the more salient this pattern is, as illustrated by the comparison of Figure 3a (3-syllable sentences) and Figure 3c (6-syllable sentences). As indicated by the pitch lines of broad and narrow focus, the focus has a lifting effect on the pitch line of the focus range. The initial focus is the widest, followed by the middle and final. In addition, the F0 top line trend in three domains (pre-focus, on-focus, and post-focus) changed systematically in term of focal position. To be specific, the pitch line of broad focus sentences is the flattest, along with the narrowest pitch range. However, with the modulation of the narrow focus, the MaxF0 value of the pre-focus domain is almost the same as the broad focus sentence. But the MaxF0 value of the on-focal domain differs most from the broad focus. Meanwhile, the pitch line of post-focus domains is hugely compressed.

4.2.2. The declarative intonational function is mainly encoded by the F0 bottom line

The F0 bottom lines of both broad and narrow focus sentences show a falling trend from the sentence-middle position. The trend and fluctuation of F0 bottom lines overlaps as well. Meanwhile, there is no significant difference among the pitch values at sentence final points of all focused sentences (the difference varied from 4 to 15Hz). All details are provided in Figures 3b and 3d. Similarly, no significant difference between MinF0 values of the narrow focus at the beginning and middle of the sentence is reported (the difference varied from 1 to 4Hz). The F0 values of sentence-final points are the lowest in each sentence of all focused environments. Also, the F0 bottom lines of all focus-adjusted declarative sentences show a declination. This trend testifies that the focus type and position had no effect on the bottom lines. In other words, the adjustment of broad and narrow focus has no notable effect on F0 bottom line trends of declarative intonation.

4.2.3. Interaction pattern of intonation and focus function

Although the declarative intonational function is mainly realized by the F0 bottom line, and the top line is mainly related to the focus function, there is an interactive relationship between two intonation functions in the actual encoding process. First, the pitch encoding process of the sentence-final narrow focus is more influenced by the declarative intonation. This can be illustrated by ANOVA tests shown in Tables 2 and 3. It is also suggested that the pitch encoding of pre-focus and post-focus differed from that of the narrow-focus. Second, the declarative function is superior to focal function when the declarative and focus functions are encoded simultaneously. The result in Table 3 shows that the second word even in pre-focus has an expanded pitch. It should be noted that the expansion is significantly different from that of the words in other pre-focus domains ($p<0.001$). This acoustic pattern indicates that the modulatory effect of declarative function is greater than that of narrow focus when they are encoded simultaneously. Finally, the bottom line is compressed when the narrow focus falls at the initial position. This post-focus compression is more prominent than those of other positions. The significant pitch compression feature of the SF bottom line is observed in Figure 3b&3d.

5. CONCLUTION

This paper statistically analyzed the F0 change of Mongolian focal accent in the declarative intonation and the interaction patterns of sentence intonation and focal accent. Major findings are summarized as follows. (1) The F0 top line of declarative intonation under the modulation of focal accent could be divided into three parts, including pre-focus, on-focus, and post-focal domain, each of which underwent systematic changes. However, each domain varied in the way the F0 top line was encoded, and was divergent from that of the broad focus sentence. (2) The declarative intonation was mainly realized by the F0 bottom line of the sentence, but it had no significant effect on the F0 top line. The F0 bottom line of the declarative intonation showed a smooth downward slant from the middle to the end of the sentence, which was not influenced by the focus. (3) The declarative intonation and focus had relatively independent encoding function, but they indeed showed some interactive features to some degree. Future studies may consider the intonation typology discussion between Mongolian intonation pattern and other languages.

6. ACKNOELEDGE

This work is supported by Inner Mongolia Autonomous Region Local Language Research Project " Mongolian Intonation Research "(MW-YB-2021022) and the " Four Batches" Talent Project grant awarded to Aijun Li .

7. REFERENCES

[1] Bolinger D.L., " Intonation and grammar, " *Language Learning*, York, pp.31-37, 1958(12).

[2] Halliday.M. Notes on transitivity and theme in English, *Journal of Linguistics*, Cambridge, pp. 199-244, 1967(03).

[3] Karolina Bros, Martin Meyer, Maria Kliesch, Volker Dellwo, "Word stress processing integrates phonological abstraction with lexical access – An ERP study," *Journal of Neurolinguistics*," kidlington, pp.1-18, 2021(57).

[4] D.R.Ladd, Intonational Phonology, Cambridge University Press, Cambridge, 1996.

[5] A. Karlsson, Rhythm and Intonation in Halh Mongolian, Lund university, 2015.

[6] A. Karlsson, "Mongolian Intonation," Journal of the Phonetic Society of Japan, 2007.

[7] Shaobo Sun, "The Prosodic Realization of Focus in Different Communicative Environments: A Comparative Study of Jing and Mongolian, Beijing, " Minzu University of China, 2019.

[8] Qenggeltei, Mongolian Grammar, Inner Mongolia People's Publishing Press, Hohhot, 1991.

[9] Harnud, H, "A Basic Study of Mongolian Prosody," University of Helsinki, 2003.

[10] LI Aijun , YUAN Yi, Phonetic Implementation of Multiple Foci in Mandarin Chinese, *Chinese Journal of Phonetics*, Beijing, pp.1 - 26, 2019(01).

[11] Xu,Y., "Effects of tone and focus on the formation and alignment of F0 contours,"*Journal of Phonetics*,1999,27,55-105.

[12] Liangping Zhong, "The Dialects Contrastive Study About Prosodic Encoding of Focus accent," Nanjing Normal University, 2015.

[13] Sunzhu, "Weakening Vowels in Modern Mongolian," *Minority Language of China*, Beijing, pp.46-54, 1981(01).

[14] Huhe,H., Acoustic Study of Mongolian Phonetics, Social sciences Academic Press, Beijing, 2018.

[15] Ziyu Xiong, Jiuzhou Sound Collection Plat, http://ling.cass.cn/tzgg/201905/t20190527_ 4906085.html, 2019.

[16] Xu, Y., Prom-on, S, "Toward invariant functional representations of variable surface fundamental frequency contours: Synthesizing speech melody via model-based stochastic learning," Speech Communication,Netherland, 2014(57), pp.181-208.

[17] Aomin, "Modeling Mongolian Declarative Intonation based on PENTA," Chinese Journal of Phonetics, Beijing, pp.138 - 146, 2021(01).

[18] Shih, Chilin. 1988. "Tone and intonation in Mandarin. Working Papers, " Cornell Phonetics Laboratory, Ithaca, pp.83-109, 1988(03).

[This paper was published at OCOCOSDA 2022]