Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/13/20230714

A corpus study on the relationship between focus and accent in natural speech in Mandarin Chinese

Ruiting Dong

School of Humanities, University of Amsterdam, Amsterdam, Netherlands

ruitingdong@muc.edu.cn

Abstract. This paper aims to explore the relationship between focus and accent, namely, whether the focus of a sentence possesses some acoustic features of accent. Previous studies suggest that focused words depict different patterns of acoustic prominence embedded in accent. According to previous evidence found in English, pitch and duration are two crucial factors contributing to differentiating focused and unfocused words. Therefore, pitch and duration are the main two factor considered in this study. Studies in aspect of the relationship between focus and accent and their realization in the phonetic level in natural speech are insufficient. Therefore, this study aims to provide support for the connection between focus and accent in natural speech in Mandarin. Based on 24 sentences from three native Mandarin speakers selected from a Chinese corpus, THCHS-30, published by Tsinghua University, this study extracts the narrow-focused words, broad-focused words, and the corresponding unfocused-words as the main subjects and analyses their mutual relationships. Results support the symmetric relationship between focus and accent. Focused words show phonetic features of accent and words with no obvious phonetic prominence of accent are unlikely to be considered as focus. Focus triggers a decrease in F0 values and an increase in duration of focused words. Linear mixed effects tests in R suggest that a major effect of narrow and broad focus in causing F0 and duration changes has not been deducted, but narrow focus conditions can result in duration proportion increase of narrow focused words compared to broad focused ones.

Keywords: focus, accent, Chinese corpus, pitch, duration.

1. Introduction

Focus, as a functional utterance element, shows the semantic stress of the sentence which the speaker intends to express. By means of pitch, length, intensity, and other rhythm elements, emphasized focus forms a distinct and identifiable prominent pattern achieved as accent, and plays a role in the surface formation of intonation. Halliday defined focus as a part of information structure, and stated that the most prominent part of a sentence in intonation accent is the information focus [1]. The information focus is endowed with accent and realized acoustically. During speech interaction, the prosodic feature is an effective form to transmit information. As one of the leading prosodic features, the emphasized accent obtains prominence with the help of phonetic form. Eady et al stated that focus and non-focus words show different patterns in acoustic features and in perception, especially in pitch contour and duration [2]. The process of phonetic prominent includes but not limited to the change of the contour of fundamental frequency, the extension of duration, the increase of intensity or loudness, more sufficient

DOI: 10.54254/2755-2721/13/20230714

acts of pronunciation, and less co-articulates [3]. Therefore, accent is closely related to focus and the prosodic form of focus can be realized by accent.

1.1. Accent realization: pitch and duration

In the studies about English accent, Cooper et al stated that accent shows an increase in the pitch of the emphasized syllables and a decrease in the pitch of the adjacent syllables [4]. Accent also has a significant influence on duration. Different from English, Nooteboom and Kruyt pointed out that in Dutch, although the duration of the emphasized syllables was lengthened, the pitch of the emphasized syllable decreased [5]. In terms of Mandarin Chinese, the acoustic features and pattern of pitch and duration of accent are similar with English. According to Xu and Chen, pitch increase of the emphasized constituents, pitch decrease in adjacent following syllables, and longer duration are prominent features of emphasized accents [6][7]. In addition, Jia et al also suggested that the difference between English and Mandarin Chinese in accent realization was the prosodic unit that the accent embedded in, specifically, syllables in English and prosodic words in Chinese [8]. In other words, in Chinese, the increase and decrease in pitch is an increase or decrease in the pitch of the accent prosodic words.

Mandarin Chinese, as a tone language instead of a stress language, its accent has a complicated communicative function and has unique acoustic features. Lin and Wang stated that the accent syllable normally shows a longer duration, a wider pitch range, and a more complete pitch pattern [9]. Pitch and duration are the main tools to express the accent words in a sentence in Mandarin Chinese. Thus, in this paper, pitch and duration are the two main factors being measured and analyzed.

1.2. Narrow and broad focus in accent realization

Ladd divided focus into two groups: narrow focus and broad focus [10]. Narrow focus refers to one single focused word in a sentence while broad focus can be a phrase or even a whole sentence. In natural language comprehension, speakers express focus by endowing the focus with accent, which gives the focused words more energy. This is the coding process from focus to accent. During comprehension, listeners pay more attention to the information words with longer duration and higher pitch. This is the decoding process from accent back to semantics. Without successfully seizing the informative words, the semantic processing can be problematic.

For narrow focus, the coding and decoding process is reversible straightforwardly. However, for broad focus, due to the accent percolation in coding process (which means that not every word in the broad focus range can be accent, so some words are chosen to be accent and some are not), during decoding process, there should be a focus projection process of projecting one or several accents to the corresponding focus range [3]. Focal accent and the broad focus range are not always symmetric. In broad focus sentence, the words that showed phonetic prominence could be only one or several words inside the focused ranges. Gussenhoven put forward the "focus-to-accent theory" (FTA theory) [11]. He pointed out that accent and focus are not in a one-to-one correspondence relationship and the distribution of focus and accent are contents in two aspects. Whether the focused part can be realized as accent in the surface form is restricted by specific language structures and regulations.

Birch and Clifton's study stated that in narrow focus sentence, the narrow-focused words are accent, while in broad focus sentences, focus can be projected to the verb of the phrase or the noun of the verb phrase so that the verb or noun in the broad focus can be accent, but the projection is selective and inconsistent [12]. Ladd stated that broad focus has a weak markedness in accent while narrow focus has a stronger markedness in accent [10]. Therefore, the accent realization of narrow focus and broad focus can be different, which would cause the difference in phonetic features in different types of focus.

1.3. Research question and hypotheses

Due to the close relationship between focus and accent, accent might be the phonetic manifestation form of focus. Previous studies on relationship between focus and accent mostly conducted control experiments by asking different questions about the same information sentence to control the position and number of focus and accent [13]. Zubizarreta defined focus as the non-presupposition part in

DOI: 10.54254/2755-2721/13/20230714

sentences and stated that focus can be realized through different forms of questions [14]. Rather than questions and dialogues, this paper applies an analysis based on corpus data in natural languages without context to explore a more general relationship pattern between focus and accent.

The main aim of this paper is to investigate whether focus of a sentence possesses the features of accent and to illustrate the acoustic features of the focal accents specifically in pitch and duration. In addition, the different roles that narrow and broad focus play in accent realization in the phonetic level are stated and analyzed.

Based on previous studies, I suggest that focused words are accent and the relationship between focus and accent is in general symmetric. The focused words are realized as accent in the surface form of a sentence, by showing some more prominent phonetic features of accent in terms of fundamental frequency and duration, while unfocused words do not show features of accent. To be more specific, I expect the F0 to be higher and the duration to be longer when the words are focused. I predict that the narrow and the broad focus show different prominent phonetic features, with the narrow focus showing a different contour (higher or lower) of fundamental frequency and a longer duration compared to the broad focus.

2. Methodology

2.1. Corpus

The free corpus THCHS-30, an open Chinese speech database published by Centre for Speech and Language Technology (CSLT) at Tsinghua University, is applied in this paper [15]. Most of the speakers are native Chinese students, aging from 20 to 35 in universities, and all are fluent in standard Mandarin. The sampling rate of the recording is 16, 000 Hz, and the sample size is 16 bits. All the sentences are selected from a large volume of news, journals, and novels. The length of each sentence is 20 to 22 characters. There is no semantic connection between these sentences with no context provided during recording.

2.2. Data collection

Five native Mandarin speakers joined the secondary data collection. They were asked to choose the focal accent words in 40 sentences randomly selected from the corpus after listening to the sound recordings. The recordings can be played several times until the participants decided the focal accent words. Sentences containing answers with high uniformity were chosen to go through the next step. Another two native Mandarin speakers, other than the previous five and I co-decided whether the sentences chosen contain narrow focuses or broad focuses. Sentences with two-syllable prosodic narrow focal accent words (both syllables have the same tone feature, the fourth falling tone) or with two-syllable focal accent words inside broad focuses were chosen as the target sentences analyzed.

Two sets of words were measured: the chosen focal accent words in both narrow and broad focus sentences as the target group and the chosen unfocused words as the control group. The unfocused words were chosen in accord with the following rules: two-syllable prosodic words; not in the focus ranges; same part of speech with the corresponding focal accent words (e.g.: verbs vs. verbs, nouns vs. nouns).

Three speakers' data were randomly chosen from the THCHS-30 open Chinese speech database. Finally, eight sentences (of three female speakers, in total 24 target sentences) were extracted from the database with four sentences with narrow focus and four sentences with broad focus.

2.3. Data analysis

The pitch value of the corresponding focused and unfocused words in narrow or broad focus sentences were measured in Praat after setting word boundaries in textgrid files [16]. For each word interval, 10 uniformly-spaced extracted F0 value were measured and recorded. Two comparisons were investigated in details: the pitch pattern of the focused constituents and the pitch pattern of the unfocused parts in the same sentences; the pitch pattern of the narrow focus and the pitch of the broad focus.

DOI: 10.54254/2755-2721/13/20230714

The duration was another factor being investigated. Both the duration of the words and the proportion of the focused parts in the whole sentences were calculated as the Equation below:

$$Proportion = \frac{duration\ of\ the\ words}{the\ duration\ of\ the\ sentence} \quad (1)$$

Two comparisons were investigated in details: the duration of the focus and non-focus in the same sentences; the duration of the narrow focus and the broad focus (2-syllable prosodic words with the same tone features).

3. Results and analyses

3.1. F0 values

Data analyses are performed by RStudio. LmerTest (Linear mixed effects test) with participants and sentences as two random factors, are conducted on the correlation between narrow and broad focus F0 values, focus and non-focus F0 values, narrow focus and non-focus F0 values, and broad focus and non-focus F0 values. The mean F0 values of narrow focus, broad focus, non-focus, and their corresponding average values are shown in Table 1.

Table 1. The mean and average F0 values of focus (narrow and broad) & non-focus (Hz).

	Sentence No.	Sp1 F0	Sp2 F0	Sp3 F0	Average
NarrowF	1	191.6	188.8	227.5	202.6
	2	270.4	309.5	259	279.6
	3	226	240.6	235.2	233.9
	4	228.3	191.4	207.4	209.0
	Average	229.1	232.6	232.3	**231.3**
Non-Narrow-Focus	1	226.8	251	246.7	241.5
	2	296.2	276.6	286.5	286.4
	3	201.9	222.2	242.5	222.2
	4	246.4	224.8	229.8	233.7
	Average	251.6	241.4	252.6	**248.5**
BroadF	5	223	194.6	217.2	211.6
	6	293.6	254.6	256.5	268.2
	7	210.5	162.3	192.8	188.5
	8	272.5	244.6	251.6	256.2
	Average	249.9	214.0	229.5	**231.2**
Non-Broad-Focus	5	201.9	222.2	242.5	222.2
	6	246.4	224.8	229.8	233.7
	7	237.3	227.9	252.2	239.1
	8	371.6	326.3	287.6	328.5
	Average	264.3	250.3	253.0	**255.9**
Focus Average		239.5	223.3	230.9	**231.2**
Non-Focus Average		258.0	245.9	252.8	**252.2**

Narrow focus vs. Broad focus. The mean F0 value of narrow focus is nearly the same as the mean F0 of broad focus, namely, 231.3 Hz and 231.2 Hz. The effect of narrow and broad is not significant (95% confidence interval= -48.48 .. +48.80, t=0.006, p=0.995>0.05). A main effect of narrow and broad focus has not been found. It cannot be concluded that narrow and broad focus can trigger any F0 changes in words.

The pattern of narrow focus F0 and broad focus F0 for speaker1 is shown in Fig. 1. The difference of F0 pattern between narrow and broad focus is not clear.

DOI: 10.54254/2755-2721/13/20230714

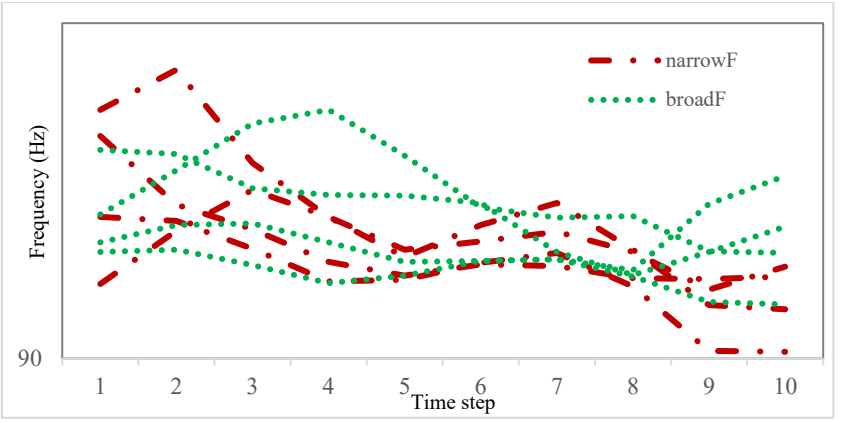

Figure 1. F0 pattern of narrow focus (dash dot) and broad focus (round dot).

Focus vs. Non-focus. From Table 1, the mean F0 value of focus is 21 Hz lower than the mean F0 of non-focus words. The F0 value of non-focus is significantly higher than F0 of the focus words (95% confidence interval= 6.06 .. 35.88, t=2.796>2, p=0.00816<0.05). A main effect of focus and non-focus is found. It can be concluded that focus condition triggers a decrease in F0 in focused words compared to non-focus words.

Narrow vs. Non-focus. The mean F0 value of narrow focus is 17.2 Hz lower than the mean F0 of non-focus words, as shown in Table 1. The F0 value of non-focus words is marginally significantly higher than F0 of the focus words (95% confidence interval= -1.92 .. +36.37, t=2.796>2, p=0.087>0.05&<0.1). A marginally significant effect of narrow focus and non-focus is found. It can be concluded that narrow focus condition triggers a marginal decrease in F0 in narrow focused words compared to non-focus words. The pattern of narrow focus F0 and non-focus F0 for speaker 1 is shown in Fig. 2.

It is clear from Fig. 2 that the F0 pattern of narrow focus shows a pattern of lower F0 at the end of the prosodic words.

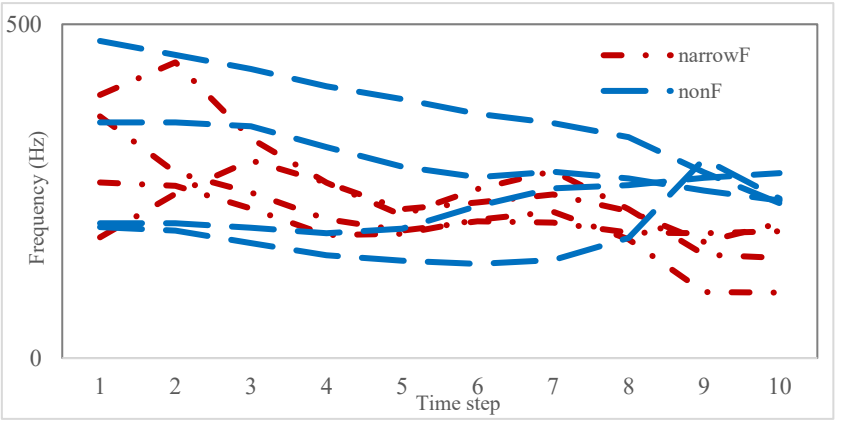

Figure 2. F0 pattern of narrow focus (dash dot) and non-focus (long dash).

Broad vs. Non-focus. The mean F0 value of broad focus is 24.7 Hz lower than the mean F0 of non-focus words. The F0 value of non-focus is significantly higher than F0 of the focus words (95% confidence interval= 1.51 .. 47.94, t=2.15>2, p=0.046<0.05). A significant effect of broad focus and non-focus has been found. It can be concluded that broad focus condition triggers a decrease in F0 in broad focus compared to non-focus words. The pattern of broad focus F0 and non-focus F0 for speaker 1 is shown in Fig. 3.

DOI: 10.54254/2755-2721/13/20230714

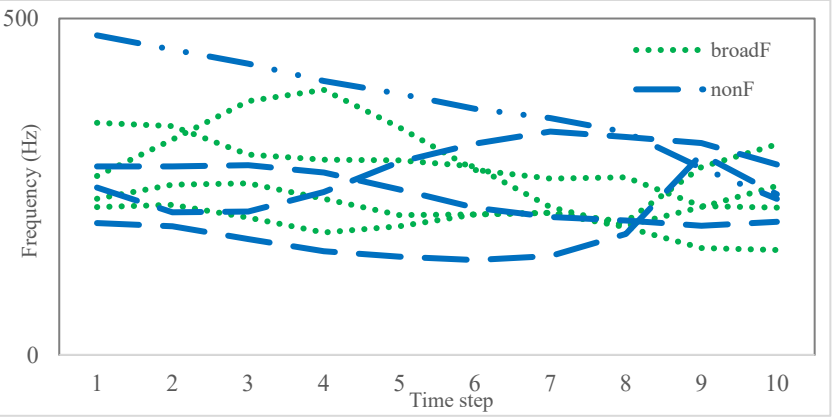

Figure 3. F0 pattern of broad focus (dot) and non-focus (long dash).

The difference between broad focus F0 pattern and non-focus F0 pattern is vague. But one data (with long dash dot) is extremely higher than the others, which may partly cause the significant effect in F0 difference in broad and non-focus words.

3.2. Duration

Data analyses are performed by RStudio. LmerTest with two random factors: participants and sentences, are done on the correlation between narrow and broad focus duration, focus and non-focus duration, narrow focus and non-focus duration, and broad focus and non-focus duration. The mean duration of narrow focus, broad focus, non-focus, and the corresponding average duration are shown in Table 2.

Table 2. The mean and average duration of focus (narrow and broad) & non-focus (seconds).

	Sentence No.	Sp1 dur	Sp2 dur	Sp3 dur	Average
NarrowF	1	0.449	0.56	0.485	0.498
	2	0.586	0.567	0.521	0.558
	3	0.356	0.431	0.408	0.398
	4	0.552	0.479	0.525	0.519
	Average	0.486	0.509	0.485	**0.493**
Non-Narrow-Focus	1	0.382	0.39	0.423	0.398
	2	0.433	0.41	0.453	0.432
	3	0.305	0.343	0.287	0.312
	4	0.404	0.378	0.294	0.359
	Average	0.381	0.380	0.364	**0.375**
BroadF	5	0.373	0.347	0.53	0.417
	6	0.604	0.411	0.446	0.487
	7	0.397	0.431	0.407	0.412
	8	0.522	0.453	0.421	0.465
	Average	0.474	0.411	0.451	**0.445**
Non-Broad-Focus	5	0.353	0.28	0.36	0.331
	6	0.451	0.455	0.462	0.456
	7	0.308	0.363	0.343	0.338
	8	0.427	0.356	0.325	0.369
	Average	0.385	0.364	0.373	**0.373**
Focus Average		0.480	0.460	0.468	**0.469**
Non-Focus Average		0.383	0.372	0.368	**0.374**

DOI: 10.54254/2755-2721/13/20230714

Narrow vs. Broad. The mean duration of narrow focus words is slightly longer than the mean duration of broad focus words, 0.048 second, as shown in Table 2. The effect of narrow and broad is not significant (95% confidence interval= -0.026 .. +0.12, t=1.243, p=0.26>0.05). A main effect of narrow and broad focus has not been found. It cannot be concluded that narrow and broad focus conditions can trigger any duration changes in words.

Focus vs. Non-focus. As can be seen from Table 2, the mean duration of focus words is 0.095 second longer than the mean duration of non-focus words. The duration of focus words is significantly longer than duration of the non-focus words (95% confidence interval= +0.066 .. +0.12, t=6.587>2, p=7.95e-08<0.01). A main effect of focus and non-focus has been found. It can be concluded that focus condition triggers an increase in duration in focus words compared to non-focus words.

Narrow vs. Non-focus. According to Table 2, the mean duration of narrow focus words is 0.118 second longer than the mean duration of non-focus words. The duration of narrow focus words is highly significantly longer than duration of the non-focus words (95% confidence interval= 0.086 .. 0.15, t=7.383>2, p=5.41e-07<0.01). A significant effect of narrow focus and non-focus has been found. It can be concluded that narrow focus condition triggers an increase in duration in focus compared to non-focus.

Broad vs. Non-focus. From Table 2, the mean duration of broad focus words is 0.072 second longer than the mean duration of non-focus words. The duration of broad focus words is significantly longer than duration of the non-focus words (95% confidence interval= 0.025 .. 0.12, t=3.108>2, p=0.00639<0.05). A significant effect of broad focus and non-focus has been found. It can be concluded that broad focus condition triggers an increase in duration in broad focused words compared to non-focus words.

3.3. Duration proportion

To avoid more individual differences in duration pattern. I suppose it can be more objective to testify the duration proportion as the individual difference can be weakened. Data analyses are performed by RStudio with lmerTest with two random factors: participants and items. The mean duration proportion of narrow focus, broad focus, non-focus, and the corresponding average duration proportion are shown in Table 3.

Table 3. The mean and average duration proportion of focus (narrow & broad) & non-focus.

	Sentence No.	Sp1 dur	Sp2 dur	Sp3 dur	Average
NarrowF	1	0.078	0.103	0.077	0.086
	2	0.093	0.108	0.088	0.096
	3	0.075	0.093	0.084	0.084
	4	0.094	0.096	0.090	0.093
	Average	0.085	0.1	0.085	**0.090**
Non-Narrow-Focus	1	0.066	0.072	0.067	0.069
	2	0.069	0.078	0.076	0.075
	3	0.065	0.074	0.059	0.066
	4	0.069	0.076	0.050	0.065
	Average	0.068	0.076	0.062	**0.068**
BroadF	5	0.074	0.064	0.087	0.075
	6	0.090	0.080	0.074	0.081
	7	0.075	0.097	0.083	0.085
	8	0.065	0.077	0.056	0.066
	Average	0.076	0.080	0.075	**0.077**
Non-Broad-Focus	5	0.070	0.052	0.059	0.060
	6	0.067	0.089	0.077	0.078
	7	0.059	0.082	0.070	0.070
	8	0.053	0.061	0.043	0.052

DOI: 10.54254/2755-2721/13/20230714

Table 4. (continued).

Average	0.060	0.077	0.063	**0.067**
Focus Average	0.081	0.090	0.080	**0.083**
Non-Focus Average	0.065	0.073	0.063	**0.067**

Narrow vs. Broad. From Table 3, the mean duration proportion of narrow focus words is 1.3% bigger than the mean duration proportion of broad focus words. The effect of narrow and broad is significant (95% confidence interval= 0.0029 .. 0.023, t=2.54, p=0.044<0.05). A main effect of narrow and broad focus has been found. It can be concluded that narrow focus condition can trigger duration proportion increase in narrow focused words compared to broad focused words.

Focus vs. Non-focus. The mean duration proportion of focus words is 1.6% bigger than the mean duration proportion of non-focus words. The duration proportion of focus words is significantly bigger than duration proportion of the non-focus words (95% confidence interval= +1.2% .. +2.1%, t=6.874>2, p=4.18e-08<0.01). A main effect of focus and non-focus has been found. It can be concluded that focus condition triggers an increase in duration proportion in focus words compared to non-focus words.

Narrow vs. Non-focus. According to Table 3, the mean duration proportion of narrow focus words is 2.2% bigger than the mean duration proportion of non-focus words. The duration proportion of narrow focus words is highly significantly bigger than duration of the non-focus words (95% confidence interval= 1.6% .. 2.6%, t=8.693>2, p=1.16e-07<0.01). A significant effect of narrow focus and non-focus has been found. It can be concluded that narrow focus condition triggers an increase in duration proportion of narrow focus compared to non-focus words.

Broad vs. Non-focus. As shown in Table 3, the mean duration proportion of broad focus words is 1.0% bigger than the mean duration proportion of non-focus words. The duration proportion of broad focus words is significantly bigger than duration proportion of the non-focus words (95% confidence interval= 0.38% .. 1.9%, t=2.988>2, p=0.0083<0.05). A significant effect of broad focus and non-focus has been found. It can be concluded that broad focus condition triggers an increase in induration proportion in broad focused words compared to non-focus words.

4. Discussion

4.1. F0 values

Focus and non-focus. The first research aim is to investigate the relationship between focus and accent. Similar with initial hypothesis, generally, the focused words are realized as accents by indicating more prominent phonetic features. However, the change in F0 indicates an opposite result different from initial hypothesis. Compared to unfocused words, focused words show a decrease in F0, which contradicts with the previous research conducted by Xu and Chen [6][7]. There are two possibilities. Firstly, the data in this paper is naturalistic language, which could cause different accent features in F0. In other words, it is possible that F0 follows a different phonetic realization pattern in natural speech and some condition resulting in widening the pitch range in an opposite direction. Secondly, the focused words can be unaccented or the phonetic realization of focus does not embed in the focal accent, which could be caused by the unsymmetric relationship between focus and accent. Previous studies also stated that the adjacent words of the focused words also indicated a change in F0 [8], but this phenomenon is not studied in this paper.

Narrow focus and non-focus. Comparison between narrow focus and non-focused words indicates a marginal difference in F0 change. From Figure 2, narrow focus condition triggered a decrease in F0 in focused parts, especially in the final parts of the words. One possible reason can be the falling tone feature of the focused words. The falling contour could cause the final pitch to decrease while the unfocused words chosen in this paper do not always have falling tones. Another possibility can be related to previous studies. Jia and Li's study stated that the focal accent shows an increase in pitch [17]. However, the adjacent followed words (focused or unfocused) showed a decrease in pitch. Therefore, it

DOI: 10.54254/2755-2721/13/20230714

is possible that the pitch decrease in word final can be a preparation for the decrease in pitch in adjacent words. However, this hypothesis still needs further support.

Broad focus and non-focus. Comparison between broad focus and non-focused words is problematic. Although in the data analysis, a clear correlation has been found indicating that the words in broad focus have lower pitch values compared to unfocused words, there exists an extreme data in unfocused words with extra high pitch values. The unfocused word has not been recognized as focus nor accent by native speakers. One reason why the unfocused word has such a high pitch could be its position, sentence initial. The other reason could be due to its falling tone contour, which could make the word initial entail a high pitch.

After getting rid of the extreme data, the lmertest shows that the effect of broad focus and non-focus in F0 is not significant (95% confidence interval= -11.28 .. +33.72, t= 1.015, p=0.329>0.05). Therefore, it cannot be concluded that broad focus and non-focus can trigger any F0 changes in words.

By looking at Figure 3, the broad focus words do not show a falling contour in F0 at the end of the words like narrow focus words do. This is surprisingly consistent with the study by Jia and Li that in sentences with broad focus, the function of accent is to prevent the falling contour of F0 [17].

Narrow and broad focus. The comparison between narrow and broad focused words illustrates a contradicting result compared with the initial hypothesis. No effect of focus type (narrow and broad) has been found in aspect of F0 values. One possibility could be that there is no difference in pitch value between narrow and broad focused words. Another factor that might play a role is part of speech. Both nouns and verbs are included in this paper. The phonetic realization of these sentence parts could be different [18].

4.2. Duration

In terms of duration, the results are neater and more consistent with hypothesis. Duration plays an important role in distinguishing focus and non-focus words. Broad focus is on average 0.07 second longer than non-focused words while narrow focus is on average 0.12 second longer than non-focused words. Consistent with previous studies and expectation, focused words show a significant longer duration in words length compared to unfocused words. Longer duration of focused words has gained one more cross language support.

When it comes to the comparison between duration of narrow and broad focus, the duration has not been significantly influenced by the type of focus. Sentence length variance and speakers' individual differences can be the causes. Therefore, the final factor duration proportion has been brought into the analysis.

4.3. Duration proportion

Similar with results in duration, focus has longer duration proportion in whole sentences compared to the non-focus. After comparing narrow focus and broad focus separately with unfocused words, narrow focus words show longer duration proportion compared to broad focus. Comparison between narrow and broad focus supports that duration proportion is significantly influenced by the focus type. Narrow focus words have a larger duration proportion than broad focus words. A possible explanation is the weak markedness in broad focus [10]. Moreover, the phonetic realization of broad focus could be a general reinforcement of phonetic prominence across the whole focus range. Therefore, the accent and energy can be distributed to all words in the focused range and the phonetic prominence for each single word is limited.

5. Conclusion

This paper mainly investigated the phonetic realization of focal accent in natural speech in Mandarin Chinese. Consistent with hypothesis, focused words have phonetic features of accent, including lower F0 values, longer duration, and longer duration proportion of the whole sentences. In general, focus and accent are symmetric in sentences with narrow and broad focus. Specifically, if a word is focused semantically, then it is highly likely that the word depicts a decrease in F0 values and an extension of

DOI: 10.54254/2755-2721/13/20230714

both duration and duration proportion in the whole sentence. A major effect of narrow and broad focus in causing F0 and duration changes cannot be concluded according to the data collected. However, data in this paper supports that narrow focus accent words occupy larger duration proportion in whole sentences than broad focus.

This is a pilot study on phonetic realization of focal accents in natural speech. The current research can be supplemented and improved with further studies. Firstly, the tone contour features for narrow and broad focus are consistent but not for unfocused words. The significant difference in focused and unfocused words could partly result from the difference in different tone features in some cases. Secondly, the sentence structure in this dataset is not always consistent. As different sentence structures could cause different strategies in accent distribution and phonetics realization, in further studies, sentence with identical structures can be grouped together and analyzed to see if the findings are consistent with this paper. Thirdly, in this research, only the focused words are analyzed. In further studies, the adjacent pairs of the focused words can also be measured and discussed to see if the phonetic realization of focal accents only work for the focused ranges. Fourthly, this study only includes 8 sentences of 3 speakers, the dataset is limited. More data can be included in further studies to achieve a more representative support.

References

[1] Halliday, M. A. K.: Notes on transitivity and theme in English. Journal of Linguistics 3(1), 37-81 (1967).
[2] Eady, S. J., Cooper, W. E., Klouda, G. V., Mueller, P. R., Lotts, D. W.: Acoustical characterization of sentential focus: narrow vs. broad and single vs dual focus environments. Language & Speech 29(3), 233-251 (1986).
[3] Wang, D., Yang, Y.: An advancement in studies on the relationship between the Focus and the Accent. Journal of Shanxi Normal University 33(4), 117-122 (2004).
[4] Cooper, W. E., Eady, S. J., Mueller, P. R.: Acoustical aspects of contrastive stress in question-answer contexts. Journal of the Acoustical Society of America 77(6), 2142-2156 (1985).
[5] Nooteboom, S. G., Kruyt, J. G.: Accents, focus distribution, and the perceived distribution of given and new information: an experiment. Journal of the Acoustical Society of America 82(5), 1512-1524 (1987).
[6] Xu, Y.: Effects of tone and focus on the formation and alignment of F0 contours. Journal of Phonetics 27(1), 55-105 (1999).
[7] Chen, Y.: Durational adjustment under corrective focus in standard Chinese. Journal of Phonetics 34(2), 176-201 (2006).
[8] Jia, Y., Xiong, Z., Li, A.: Phonetic and phonological analysis of focal accents of disyllabic words in standard Chinese. In: 5th International Symposium on Chinese Spoken Language Processing, pp. 55-66. Springer, Berlin, Heidelberg (2006).
[9] Lin, X., Wang, L.: Phonetics. Beijing Linguistics College Press, Beijing (1992).
[10] Ladd, R. D.: Intonational Phonology. Cambridge University Press, Cambridge (1996).
[11] Gussenhoven, C.: Focus, mode and the nucleus. Journal of Linguistics 19(2), 377-417 (1983).
[12] Birch, S., Clifton, C.: Focus, accent, and argument structure: effects on language comprehension. Language and Speech 38(4), 365-391 (1995).
[13] Eady, S. J., Stephen, J.: Speech intonation and focus location in matched statements and questions. Journal of the Acoustical Society of America 80(2), 402-415 (1986).
[14] Zubizarreta, M. L.: Prosody, Focus, and Word Order. MIT Press (1998).
[15] Wang, D., Zhang, X., Zhang, Z.: THCHS-30: A free Chinese speech corpus. http://arxiv.org/abs/1512.01882, last accessed 2015/12/10.
[16] Boersma, P., Weenink, D.: Praat: doing phonetics by computer (Version 4.3.12) [Computer Program]. http://www.praat.org, last accessed 2005/05/11.
[17] Jia, Y., Li, A.: The Hierarchical Structure of Chinese Accents. Journal of Chinese Phonetics 3(1), 30-39 (2012).

DOI: 10.54254/2755-2721/13/20230714

[18] Hoskins, S. R.: The prosody of broad and narrow focus within noun phrases. Journal of the Acoustical Society of America 101(5), 3193-3194 (1997).