Production and Perception of New, Given, and Contrastive Information in Taiwan Mandarin

Ho-hsien Pan, Wan-ting Huang,Ying-Hui Huang

Institute of Foreign Literatures and Linguistics
National Chiao Tung University, TAIWAN
hhpan@faculty.nctu.edu.tw

Abstract

This paper investigates the production and perception of an NP in utterance initial and medial position produced with given, new, and contrastive information structures in spontaneous Taiwan Mandarin speech. Production data showed that duration of NP were lengthened when carrying new and contrastive information. However, results of a dialogue history perception experiment found that listeners were unable to distinguish new and given information, but were able to distinguish contrastive and given information.

1. Introduction

A sentence can carry given information that is shared by a listener and a speaker, and new or contrastive information that adds more knowledge to dialogues. In terms of discourse, new information is usually elicited after WH questions, while contrastive information is usually elicited through correction of inaccurate information provided in previous utterances.

In production, the difference between given, new and contrastive information can be observed either in acoustic variation, or in the distribution of sentence accents through the utterance. In the case of acoustic variation, it has been revealed that syllables carrying new information are longer in duration, have expanded f0 range, and are produced with more peripheral vowels than syllables carrying given information in English [1, 2], Japanese [3], Peking Mandarin [4, 5], and Taiwan Mandarin [6]. In the case of accent distribution, narrow focused syllables carrying new or contrastive information attract sentence accents in languages such as Dutch and English [7]. Languages that vary the placement of accents to indicate focus are called plastic languages, while languages that modify already existing acoustic properties, including duration, F0 expansion, and vowel quality are called non-plastic languages. Thus intonation structure varies according to focus location in plastic language. This is not the case in non-plastic languages such as Italian, where placement of focus on syllables carrying new and contrastive information does not attract sentence accent. In non-plastic languages, the intonation structure remains the same regardless of the variation on the domain of focus in a sentence [8]. The duration elongation, f0 range expansion, and peripheral vowels observed on focused syllables were not profound enough to the extent of modifying tonal distribution of a sentence in non-plastic language. The location of sentence accent remains the same in non-plastic language.

Studies on broad versus narrow focus in tone languages revealed that narrow focused syllables carrying new information are marked by duration lengthening, f0 range expansion, and have more peripheral vowel nuclei. That is,

for a high front vowel /i/, the vowel quality becomes fronter and higher under narrow focus, while a high back vowel /u/ becomes higher and backer under narrow focus [3, 6]. Moreover, dialectal differences are observed between Peking and Taiwan Mandarin in terms of relative importance of the three acoustical parameters associated with narrow focus. For Peking Mandarin, f0 range expansion is a more salient acoustical cue than duration lengthening [4, 5], while duration elongation is a more important cue than f0 range expansion in Taiwan Mandarin [6]. The more salient role that duration elongation plays in realization of narrow focus in Taiwan Mandarin may be due to phonetic interference form Taiwan Min, a language spoken by the majority of people in Taiwan. In Taiwan Min, it has been found that the differences between broad and narrow focused syllables vary more dramatically in terms of duration lengthening, than in terms of f0 range expansion [9].

In terms of perception, in plastic languages such as English and Dutch, listeners decipher the information structure of a sentence through identification and location of sentence accents [10]. In a dialogue history experiment, Dutch listeners were able to reconstruct preceding utterances based on placement of sentence accent which signaled the parts of the sentence that carried new and contrastive information. After identifying new and contrastive information carried by syllables with a sentence accent, listeners were then able to deduce the type of information that was not present in the preceding utterance, thus identifying the previous utterance in the dialogue history experimental paradigm [8, 11]. On the other hand, in non-plastic languages such as Italian, the location of sentence accent does not vary with changes in information structure. Thus, the lack of accentual cues in the sentence posed more difficulties to listeners when trying to determine the information structure of an utterance. Not being able to identify which part of the sentence carried new information, listeners were unable to identify the type of information missing from the previous utterance and thus had more difficulty retrieving the preceding utterance in an dialogue history experiment [11].

It is unknown whether tone languages such as Mandarin and Taiwan Min are non-plastic or plastic languages. However, because tonal targets for syllables in a sentence are specified at a lexical level and do not vary according to information structure in Taiwan Mandarin, it is hypothesized that Taiwan Mandarin is more similar to Italian, a non-plastic language, of which the acoustic variation observed on focused syllables were not strong enough to cause shift in sentence accent or tonal distribution.

This study expands the study of focus in tone languages from broad and narrow focus to contrastive focus. Both production and perceptual data will be analyzed. Acoustic and

perception measures are used to address the following: (1) the effect of contrastive focus on acoustical variation; (2) the perception of given, new, and contrastive information in a dialogue history experiment experimental paradigm; and (3) typological differences between plastic, non-plastic and tone language.

2. Method

2.1. Production data

2.1.1. Speaker

A female native Taiwan Mandarin speaker, WHZ, produced the corpus in Taiwan Min. She received 10 years of ESL education and was unable to speak Taiwan Min or Hakka, two of the most prevalent languages used at home in Taiwan. WXT was a student at the National Chiao Tung University at the time of recording and was paid for her participation.

2.1.2. Corpus

Four Mandarin color terms, and four disyllabic nouns were chosen, with each syllable produced with each of four lexical tones: high level /˥/, low rising /˩˥/, falling rising /˩˩˧/, and high falling /˥˩/, as shown in Table 1.

Table 1: Four color terms and four nouns used in the corpus

Tone	Color	Noun
1 ˥	/dʒin1/ 'gold'	/tsə1 sĩã1/ 'cabinet'
2 ˩˥	/tsəŋ2/ 'orange'	/mən2 pai2/ 'plate on door'
3 ˩˩˧	/fən 3/ 'pink'	/ fən3 bi3/ 'chalk'
4 ˥˩	/xə4/ 'brown'	/dʐəŋ4 dʒian4/ 'ID'

By matching the four color terms with four nouns, 16 noun phrases (NP) were created. The 16 NPs were placed at both the initial and medial position of a sentences with a structure as shown in (1).

- (1) *NP1* tsai4 *NP2* de0 sia4 mian4

- *NP1* [VP tsai4 [*NP2* de0 [N sia4 mian4]]]

- '*NP*1 is under *NP2*'.

By matching each of the 16 noun phrases at the initial position with the other 15 noun phrases at the middle position, 240 (16 NP1 × 15 NP2) sentences were created.

In Experiment session 1, each of the 240 sentences could be produced with narrow focus carrying new information on NP1 or NP2. All together 1440 (16 NP1 × 15 NP2 × 2 position ×3 repetitions) sentences were elicited in Experiment 1.

In Experiment 2, the 16 NPs at the initial (NP1) and medial position (NP2) carried either contrastive information on both the adjective and noun (CC), contrastive information on the adjective and given information on the noun (CG), given information on the adjective and contrastive information on the noun (GC), or given information on both the adjective and noun (GG). All together 384 (16 NP × 4 information structures × 2 positions ×3 repetitions) sentences were elicited by correcting or confirming the previous utterance.

2.1.3. Instrumentation

Two AKG HSD 200 earphones with head-mounted microphones were used to collect speech from experimenters and WHZ on separated channels that were recorded by a SONY CDR-W66 compact disc recorder deck.

2.1.4. Experimental procedure

The recording was conducted in the phonetics lab at National Chiao Tung University, with the experimenter sitting outside of a sound treated room in a quiet lab setting, while the speaker sat inside a sound treated room. The experimenter and speaker could only hear each other through the headphones.

In order to elicit spontaneous speech, during experiment 1, WHZ were presented with 16 colored objects (4 colors × 4 nouns) arrayed in six different 4 × 4 matrixes, with each matrix on a separate page. The order of sentences elicited from the same matrix was randomized. During the recording, the experimenter asked precursor questions such as 'What is under / on top of / at the right of / at the left of NP2?' to elicit utterances with focus on the NP1, or 'NP1 is under / on top of / at the right of / at the left of what?' to elicit utterances with focus on NP2. Then the speaker searched the matrix to answer the questions. If the speaker made a mistake and provided the wrong answer then the question was repeated again. After asking all the questions designed for the first matrix, the speaker was presented with the next matrix on a separate page.

During Experiment 2, only the first matrix was used. The sentences were further divided into two sections, with one section containing sentences with contrastive focus on NP1, and the other containing sentences with focus on NP2. During the first session subjects were instructed to pay attention to the description of NP1 and see if the experimenter's description matched with the object on the matrix. If the experimenter's description matched the objects on the matrix, then the speaker would reply by saying 'yes' then repeat the sentence again. In this case, given information was elicited on both the color and noun (GG) of NP1. If the experimenter's description disagreed with the object on the matrix then the speaker answered 'no' then provided the correct information. For example, a speaker replied with 'no, the orange chalk is at the right of the golden cart' to correct the experimenter's precursor description, that was 'the orange chalk is at the right of the orange ID.' In this case, both the color and the noun of the NP2 were incorrect, thus the speaker provided contrastive information for the color and the noun (CC) in NP2. There were cases in which either only the color or the noun were incorrect, thus contrastive information was elicited only on the color (CG) or the noun (GC).

2.1.5. Data analysis

The recorded signals were digitized at 22 kHz and 16 bits. F0 tracks and formants were generated with a pitch and formant tool function in EMU (http://emu.sourceforge.net/). Vowel nuclei for each of the five syllables in the NP1 and NP2 were segmented using spectrographic cues. For each of the five vowels in NP1 or NP2, the time at the onset and offset of vowel nucleus were queried, as were the F0 values at 10 % intervals in time into the vowel nucleus.

Durations for the vowel nucleus of each of the five syllables were derived by subtracting the time at the onset of the vowel from the time at the offset of vowel.

2.2. Dialogue history experiment

2.2.1. Listeners

Six native Taiwan Mandarin speakers participated in the experiment. They were students at the national Chiao Tung University at the time of test.

2.2.2. Corpus

During the learning session of dialogue history experiment 1, sentences with the same tone on both color and noun in the NP were selected from production experiment 1. The four NPs that carried the same tone on both color and noun are 'golden cabinet, orange door-plate, pink chalk and brown ID'. By matching the 4 NP1s with 3 NP2s, there were 12 sentences. Using both second and third repetitions for each sentence from the production experiment, there were 48 (4 NP1* 3 NP2 * 2 focus positions * 2 repititions) selected for the learning session in dialogue history experiment 1. For the test session of dialogue history experiment 1, the first repetition from production Experiment 1 was used. There were all together 480 (16 NP1*15 NP2 * 2 focus positions) sentences used in test session. The order in which the 480 sentences were presented was randomized and divided into eight sessions, with 60 sentences presented in each session. The ISI was 7 seconds. Each session lasted 8 minutes. After finishing two sessions, listeners took a five minute break.

The learning session of dialogue history experiment 2 consisted of second and third repetitions of four NPs with the same tone on both color term and object, such as 'golden car', 'orange door-plate', 'pink chalk' and 'brown ID', at the initial (NP1) and medial (NP2) positions from production experiment 2. By matching the four information structures, contrastive information on color and object (CC), contrastive information on color and given information on object (CG), given information on color and contrastive information on object (GC), given information on both color and object (GG), with the four NPs carrying the same tone at the initial and medial positions, there were 32 (4 NP * 4 information structures *2 position) sentences. Using both repetitions 2 and 3 for each sentence from production experiment 2, 64 (4*2*4* 2) sentences were selected. During the test session of dialogue history experiment 2, the first repetition of sentences from production experiment 2 were used. There were128 sentences that were further divided into two sections, with 64 sentences in each section. The ISI was 10 seconds. Listeners took a 5 minutes break between each section.

2.2.3. Dialogue history experimental procedure

During the learning session of dialogue history experiment 1 and 2, listeners sat in front of the computer, listening to sentences played through speakers in randomized order, while boxes, containing preceding utterances were displayed on the monitor. After listening to the sentence, subject selected one of the boxes containing preceding utterances, and then a box with the correct answer lighted up as feedback. For dialogue history experiment 1, there were two choices on the monitor. For example, for the sentence 'Golden cabinet is on the left of pink chalk,' the two choices on the monitor were 'What is on the left of the pink chalk?' and 'The golden cabinet is on the left of what?' As for dialogue history experiment 2, for the sentence 'The *golden cabinet* is on the left of the pink chalk,' the four choices shown on the monitor were 'The *pink* cabinet is on the left of the pink chalk.' (CG on NP1), 'The golden *ID* is on the left of the pink chalk,' (GC on NP1), 'The *pink ID* is on the left of the pink chalk,' (CC on NP1), or 'The *golden cabinet* is on the left of the pink chalk,' (GG on NP1). Again, after the sentence was played, listeners selected one of the four boxes, and then the box containing the right choice lit up to provide feedback. During the test session the same procedure continued for dialogue history experiment 1 and 2, except that there was no feedback.

3. Results

3.1. Production experiment

Duration of NP2, but not NP1, was longer when carrying new information in production experiment 1. Similarly, duration of NP1 and NP2 were longer when carrying contrastive than given information in production experiment 2, as shown in Figure 1.
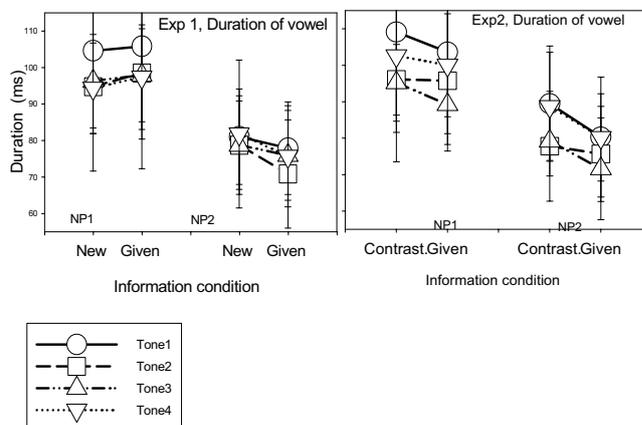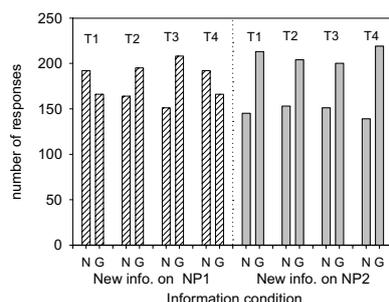

Figure 1. Duration of vowel in Exp1 and 2

Figure 2, Results of dialogue hisotry exp 1.
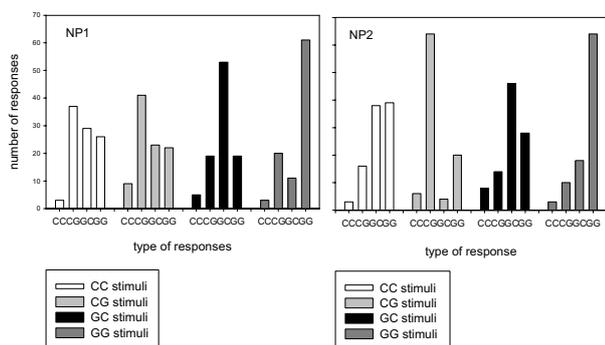
on new and given informaiton

Figure 3, Results of dialogue history exp, 2 on
contrastive and given information

Results of dialogue history experiment 1 showed that listeners' identified new information on NP2 as given information (Figure 2). Although, new information on NP1 carrying tone 1 and 4 were identified more often as new information than as given information, however, even in these cases, the percentage of responses as new information was around 50%. In other words, listeners were not able to distinguish between new and given information in dialogue history experiment 1. During dialogue history experiment 2 besides NP1 and NP2 stimuli carrying CC information, listeners responded to CG, GC and GG information structures with majority CG, GC, and GG responses respectively. In other words, listeners were able to differentiate CG, GC, and GG information structures, but not on CC information structure (Figure 3).

4. Discussion

In the production experiment, duration elongation is the major acoustical cue to narrow focused syllables with new information on NP2 in experiment 1 and contrastive information on both NP1 and NP2 in experiment 2. Perception wise, listeners are not able to retrieve previous utterance during dialogue history experiment 1 with distinction between new and given information. This is due to the indistinctive duration difference between new and given information in NP1. However, listeners were able to distinguish between contrastive and given information in CG, GC, and CC information structures during dialogue history experiment 2, due to the clear duration distinction.

Typological wise, Taiwan Mandarin is not entirely similar to non-plastic language, since listeners can perceive information structure differences based on duration elongation, however, Taiwan Mandarin is not a plastic language either, since narrow focused syllable does not attract pitch accent, instead is marked by duration elongation. For future study, the thresholds for acoustical variations required to perceive narrow focus with new and contrastive information should be investigated with synthesized stimuli.

5. Acknowledgements

6. References

[1] Cooper, W.E., S.J. Eady, and Muller, "Acoustical aspects of contrastive stress in question-answer contexts." *Journal of Acoustical Society of America*, 77: 2142-2156., 1985.

[2] Eady, S.J., et al., "Acoustic characteristics of sentential focus: narrow vs. broad and single vs. dual focus environments", *Language and Speech*, 29: 233-251, 1986.

[3] Maekawa, K., Effects of focus on duration and vowel formant frequency in Japanese, in *Computing prosody*, Y. Sagisaka, N. Campberll, and M. Higuchi, Editors. 1997, Springer: New York. pp. 129-153.

[4] Jin, S., An Acoustic Study of Sentence Stress in Mandarin Chinese. 1996, The Ohio State University.: Columbus.

[5] Xu, Y., "Effects of Tone and Focus on the Formation and Alignment of f0 Contours", *Journal of Phonetics*, 27: 55-107, 1999.

[6] Huang, Y.-H., Focus condition in Spontaneous Taiwanese Mandarin. 2004, M. A. thesis, National Chiao Tung University: Hsinchu, TAIWAN.

[7] Krahmer, E. and M. Swerts, "On the alleged existence of contrastive accents", *Speech Communication,* 34: 391-405, 2001.

[8] D'Imperio, M., "Focus and tonal structure in Neapolitan Italian", *Speech Communication,* 33: 339-356, 2001.

[9] Pan, H.-h., Focus and Taiwanese Unchecked Tones*, in Topic an*d *Focus: A Cross-Linguistic Perspective,* C.-M. Lee, M. Gordon, and D. Buering, Editors. in press, Kluwer Academic Publishing: Dordrecht.

[10] Cutler, A., Stress and accent in Language production and understanding*, in Intonati*on, a*ccent and Rhythm. studies in discourse phonology, D. Gibbon a*nd H. Richter, Editors. 1984, de Gruyter: Berlin, New York. p. 77-90.

[11] Swerts, M., E. Krahmer, and C. Avesani, "Prosodic marking of information status in Dutch and Italian: a comparative analysis", *Journal of Phonetics*, 30(629-654), 2002.