ACOUSTIC CUES TO SYNTACTIC AMBIGUITY IN TAIWAN MANDARIN

Yuh-jen Pan[a] *& Ho-hsien Pan*[a,b]

[a]Graduate Institute of Foreign Literatures and Linguistics, National Chiao Tung University, Taiwan;
[b]Department of Foreign Languages and Literatures, National Chiao Tung University, Taiwan
yuhjenpan@gmail.com; hhpan@faculty.nctu.edu.tw

ABSTRACT

This study investigated how acoustic cues signal the different syntactic structures of ambiguous sentences in language production. Acoustic cues such as intensity, f0, pause and duration vary according to different syntactic boundaries. Intensity of post-boundary syllable, f0 reset range of syllables around the corresponding target boundary, pause, and final lengthening of pre-boundary syllable were compared. Results showed that the boundaries across higher branching syntactic structures are marked by larger cross boundary f0 reset range, the existence of pause, and longer pre-boundary final lengthening. Acoustic cues can facilitate top down information during nature language parsing to identify intended meanings.

Keywords: Taiwan Mandarin, syntactically ambiguous sentences, boundary, F0 reset range

1. INTRODUCTION

Syntactic ambiguities found in written texts are often absent from speech. Allbritton, McKoon, Ratcliff [1] indicated that speakers are able to produce sufficient cues to help listeners to disambiguate intended meaning. These include longer syllable duration and lower f0 in words before major boundary than before minor boundary [3]. This study observes the role that acoustic cues play in the production of Taiwan Mandarin to signal different syntactic boundaries in syntactically ambiguous sentences.

Final lengthening is one of important cues to signal boundary strength. The final syllable of phrases tend to be longer than the same syllable would be in a phrase-medial position [3]. Previous research has shown that pre-boundary lengthening is organized hierarchically; durations are longer in the proximity of stronger boundary [6]. As to the weaker boundary, the vowel duration tends to be reduced [5].

In Swedish, pitch reset is different according to different syntactic structures. For example, there is a weaker reset in between phrases than between utterances. In addition, speakers can use pitch reset range to signal different boundary strength [4]. In tone languages, lexical tone is a property of a syllable; therefore, f0 variation is more complicated because of different tonal contexts. For instance, In Taiwanese, final-lowering and final-lengthening are contributed by word positions, and can provide cues for phrasal juncture and termination of sentences. Though pitch range were not affected by tonal context, f0 is affected by anticipatory and assimilation tonal coarticulation [2].

Taiwan Mandarin syntactically ambiguous sentences are utterances which have identical texts but different syntactic structures, causing a potential problem in spoken language processing. This study seeks to explore how native Taiwan Mandarin speakers use acoustic features (duration, f0, and intensity) to mark syntactic boundary in their speech and disambiguate syntactically ambiguous sentences.

2. METHOD

2.1. Subjects

Six female native Mandarin speakers studying at National Chiao Tung University participated in the production experiment.

2.2. Corpus

From the United Daily News database, eight types of syntactically ambiguous sentence pairs with structures contrasting at a boundary location were selected. For example, in sentence (1a), the target boundary is located between main verb “care” and adjective phrase (AdjP) “Chinese.”. In sentence (1b), the target boundary is located within the embedded relative clauses to modify noun. In (1b), within the relative clause verb phrase, the verb /kuan55ɕin55/ “care” and the following country names, e.g. “China”, were syntactically more

closely related to each other than the corresponding verb and adjective phrase in (1a). The closer the two words were in syntactic relationship, the weaker the boundaries became between the two words. Thus the boundary in (1b) is weaker than the boundary in (1a) relatively.

(1a)
[kuan55ɕin55][[ʐoŋ55kuo35 tə0][ɕue35ʂeŋ55]]
[care][[China adj-marker][student]]
V * [AdjP+ N]
關心 中國 的 學生
"(He) cares about Chinese students."

(1b)
[kuan55ɕin55[ʐoŋ55kuo35] tə0][ɕue35ʂeŋ55]
[care[China]adj-marker][student]
[V *NP] +N
關心中國 的 學生
"A student who cares about China"

The ambiguities of the sentences were checked by 48 subjects in a pretest. All together 192 sentences (8 types * 2 boundary strengths * 4 lexical items with different tones* 3 repetitions) were elicited from each speakers. Eight types of ambiguous sentences which are represented using grammatical category abbreviation and symbols are shown in Table 1.

Table 1: Eight types of syntactic ambiguous sentences. *: target boundary location, [ ]: phrase domain.

Type	Boundary strength	Symbolization
1	stronger	V*[AdjP+N]
	weaker	[V*NP]+N
2	stronger	[V+NP]*V
	weaker	V+[NP*V]
3	stronger	[AdjP+N]*ConjP+N
	weaker	AdjP+[N*ConjP+N]
4	stronger	[NP]*V
	weaker	[σ*σ]
5	stronger	N*[V+Neg+V]
	weaker	[σ*σ]
6	stronger	N*[NegP]
	weaker	[N*NegP]
7	stronger	P*[N+N]
	weaker	[P*N]+V
8	stronger	AdvP*[NP+I]
	weaker	[AdvP*NP]+I

2.3. Instrument

Because of updates in lab, recordings were made with two different microphones and recorders. An AKG HSD200 microphone and Shure Professional microphone were used to pick up speech which was then recorded with a SONY compact disc recorder CDR-W66 or Korg MR1000 1-bit Professional mobile recorder.

2.4. Procedure

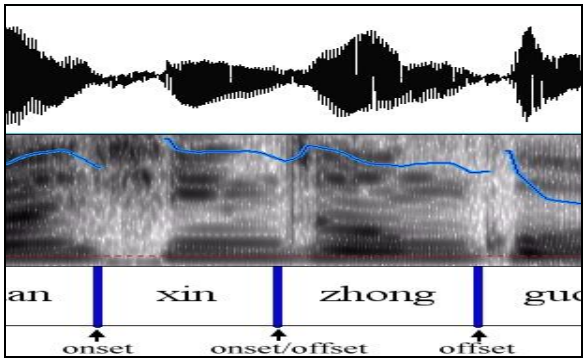
The recording was made in a sound treated room at phonetic lab. The target sentence was shown on a computer screen, and then the experimenter read a small passage describing a scenario in which the displayed sentence could be used. After describing the scenario, the experimenter asked a question to elicit the target sentence with a meaning intended for the specific scenario.

For example, to elicit the target sentence (1a) "(He) cares (about) Chinese students." The experimenter would first read a scenario, "He is a retired principal, and regularly made donations to Chinese students in a suburban school", then the experimenter asked a question, "What did the principal care?" The speaker then answered with the target sentence "(He) cares about Chinese students." To elicit target sentence (1b), the experimenter would first describe a scenario in which a Chinese student immigrated to Canada many years earlier but still cared about Chinese news. Then the experimenter would ask a question "What kind of student is he?" to elicit the sentence "A student who cares about China."

2.5. Data analysis

After the recording, the data was transcribed using Praat to mark the syllable onsets and offsets. The example is as Figure 1. Then the acoustic data was querried with use of Voicesauce. The intensity of syllables around boundaries, the minimum f0 in the pre-boundary syllable, the maximum f0 in the post-boundary syllable, the duration of pauses, and the duration of syllables around boundaries were measured.

Figure 1: Syllable onset boundaries and offset boundaries of syllables.



Due to the possibility of existence of pauses at the target boundary, each production was listened to by three native listeners to decide whether there were pauses in the utterance. If two out of the three transcribers agreed on the presence of pause at target boundary, the duration of pause was tagged.

The raw values of intensity, f0, and duration were adjusted using function (1) to calculate the percentile of the duration of a target syllable within the duration range of the same syllables produced by the same speaker.

$$Adjust\ value_{speaker\ X} = \frac{value_{speaker\ X} - MIN\ value_{speaker\ X}}{MAX\ value_{speaker\ X} - MIN\ value_{speaker\ X}} \quad (1)$$

To control for speech rate, the portion of each syllable duration within each utterance was calculated. Function (2) was used to calculate the portion of the target syllable duration within an utterance.

$$Normalize\ duration_{speaker\ X} = \frac{Adjust\ duration_{speaker\ X}}{\sum_{k=1}^{n}\left(Adjust\ duration_{speaker\ X}\right)_k} \quad (2)$$

where *n* is the syllable number of a sentence, and *k* is the syllable order in the sentence.

3. RESULTS

3.1. Intensity

The intensity of pre-boundary syllables and post-boundary syllables were measured for each sentence. Figure 2 shows the adjusted intensity of pre-boundary syllables and the post-boundary syllables. Paired Sample T-tests were performed for the adjusted syllable intensity. The result revealed that the intensity of post-boundary syllables were stronger after a stronger boundary than after a weaker boundary for sentences types 2, 4 and 6 (type 2: t(71) = 2.159, p <.05; type 4: t(71) = 4.141, p <.001; type 6: t(71) = 3.979, p <.001).

3.2. F0 reset range

Figure 3 shows the normalized f0 reset range in different types of sentences. Paired Sample T-tests showed that the adjusted f0 reset range is significantly higher across stronger boundaries than across weaker boundaries for all sentence types. (type 1: t(71) = 7.015, p < 0.001; type 2: t(71) = 4.137, p < 0.001; type 3: t(71) = 5.803, p < 0.001; type 4: t(71) = 5.231, p < 0.001; type 5: t(71) = 2.659, p < 0.01; type 6: t(71) = 2.687, p < 0.01; type 7: t(71) = 6.007 p < 0.001; type 8: t(71) = 4.898, p < 0.001).

Figure 2: Adjusted intensity of pre-boundary and post-boundary syllables in 8 types of ambiguous sentences.

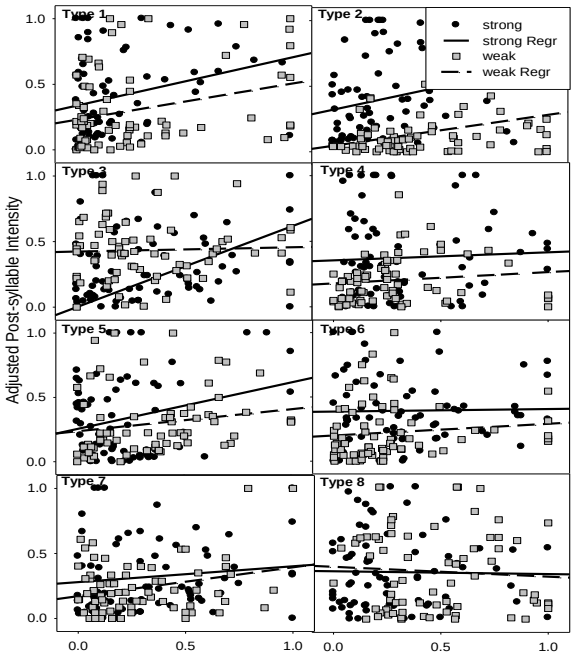
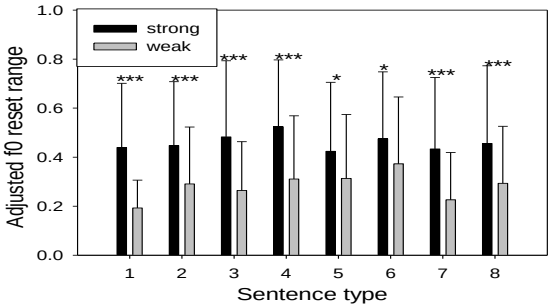


Figure 3: Adjusted f0 reset range in 8 types of ambiguous sentences. *: p<0.05, **: p<0.01, ***: p<0.001.



3.3. Pause

266 out of the 576 sentences with stronger boundaries have pauses at target boundaries. Table 2 lists the percentage and mean duration of pauses for each sentence type. Pauses only existed at the stronger boundaries. There were no pauses after weaker boundaries.

3.4. Duration

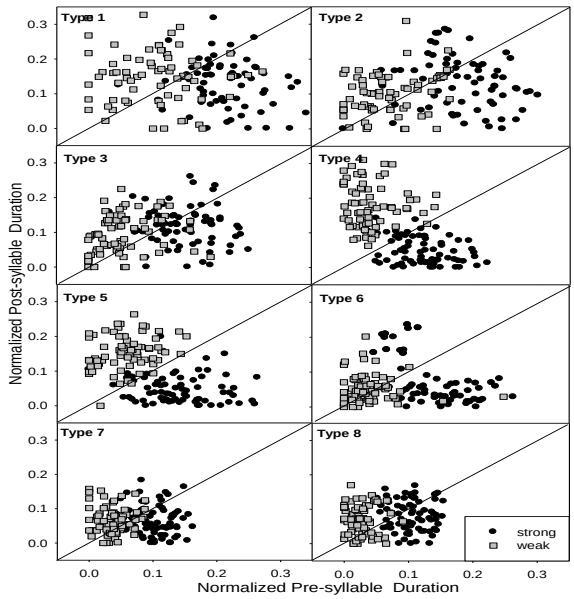
Figure 4 shows the normalized durations of pre-boundary syllables and post-boundary syllables. Paired Sample T-tests was performed for the

normalized syllable duration. This showed that the duration of the pre-boundary syllable is significantly longer before a stronger boundary than before a weaker boundary (type 1: $t(71) = 11.412$, $p < .001$; type 2: $t(71) = 11.501$, $p < .001$; type 3: $t(71) = 12.595$, $p < .001$; type 4: $t(71) = 13.289$, $p < .001$; type 5: $t(71) = 12.615$, $p < .001$; type 6: $t(71) = 11.226$, $p < .001$; type 7: $t(71) = 16.615$ $p < .001$; type 8: $t(71) = 19.710$, $p < .001$).

Table 2: The percentage of sentences with pause and the pause duration means. S: Stronger boundary. W: Weaker boundary.

Type	Boundary strength	Percentage	Pause Duration (ms)
1	S	6.0%	50.98
	W	0%	0
2	S	9.0%	84.27
	W	0%	0
3	S	10.6%	126.09
	W	0%	0
4	S	19.4%	116.32
	W	0%	0
5	S	16.4%	101.11
	W	0%	0
6	S	15.0%	105.20
	W	0%	0
7	S	11.2%	109.97
	W	0%	0
8	S	11.9%	128.78
	W	0%	0

Figure 4: Normalized duration of pre-boundary syllables and post-boundary syllables in 8 types of ambiguous sentences.



4. DISCUSSION

This paper examined the effect of syntactic boundary strength on intensity, f0 reset range, pause, and duration. We found that initial strengthening pattern was not prevalent because only some sentence types had the effect of the intensity of the post-boundary syllable increased. The f0 reset range was significantly different between stronger and weaker boundaries. The greater extent of f0 reset range indicated a stronger boundary. Pause only showed after the stronger boundary. More final lengthening in pre-boundary syllables indicates a syntactic branching of higher syntactic hierarchy. In sum, Taiwan Mandarin syntactically ambiguous sentences can be distinguished with acoustical cues such as f0 range reset across boundaries, pauses, and the duration of pre-boundary syllables.

Further perceptual studies are necessary to identify how listeners use these acoustic cues in processing ambiguous sentences.

5. ACKNOWLEDGEMENTS

The study was supported by a grant from National Science Council from TAIWAN (NSC 98-2410-H-009-038).

6. REFERENCES

[1] Allbritton, D.W., McKoon, G., Ratcliff, R. 1996. Reliability of prosodic cues for resolving syntactic ambiguity. *J. of Experimental Psychology: Learning, Memory, and Cognition* 22, 714-735.

[2] Peng, S.H. 1997. Production and perception of Taiwanese tones in different tonal and prosodic contexts. *J. of Phonetics* 25, 371-400.

[3] Price, P.J., Ostendorf, M., Shattuck-hufnage, S., Fong, C. 1991. The use of prosody in syntactic disambiguation. *J. Acoust. Soc. Am*. 90, 2956-2970.

[4] Swerts, M. 1997. Prosodic features at discourse boundaries of different strength. *J. Acoust. Soc. Am*. 101, 514-521.

[5] Tabain, M. 2003. Effects of prosodic boundary on /aC/ sequences: Acoustic results. *J. Acoust. Soc. Am*. 113, 516-531.

[6] Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M., Price, P.J. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *J. Acoust. Soc. Am*. 91, 1707-1717.