Effects of Prosody in the Processing of Temporarily Ambiguous Japanese Sentences

Yuka Naito-Billen, Robert Fiorentino, Utako Minai
University of Kansas

1. Introduction

The role of prosody in sentence processing has recently been explored in various languages (e.g., Carlson, 2009; Kjelgaard & Speer, 1999). However, it remains unclear how prosody affects the resolution of structural ambiguity in a head-final, free-word-order language like Japanese, while reanalysis of clause boundaries in such a language is a matter of debate (e.g., Miyamoto, 2003). Previous studies on Japanese prosody using globally ambiguous sentences have shown effects of prosody on sentence interpretation, but have argued that these effects are limited, rather than fully deterministic (Misono et al., 1997; Eda et al., 2009). This study further examines native Japanese speakers’ processing of sentences, employing temporarily ambiguous sentences in two auditory timed tasks and one visual untimed task.

2. Method

Eighty-four pairs of sentences (embedded-clause-bias sentences (ES) and main-clause-bias sentences (MS)) were constructed. Examples of the sentences are shown below:

(1) ES: Takasi-wa (#E) tegami-o (#M) yondeiru atarasi sensee-ni sotto eshakusita.
Takasi (name)-Top letter-Acc reading new teacher-Dat softly bowed
“Takashi softly bowed to the new teacher who was reading a letter.”

(2) MS: Takasi-wa (#E) tegami-o (#M) yondeiru atarasi kyookasho-ni sotto hasanda.
Takasi (name)-Top letter-Acc reading new textbook-Dat softly put in
“Takashi softly put the letter between the new textbook pages that he was reading.”

A female native speaker of Japanese who has a Tokyo accent read these sentences in a sound-proof room. She read each sentence with two different prosodies (embedded prosody (EP); prosodic break at (#E); and main prosody (MP); prosodic break at (#M)). This yielded 336 temporarily ambiguous sentences. 168 globally ambiguous sentences and 168 ungrammatical sentences were also created to serve as fillers, half of which were read with EP and the other half with MP by the same speaker. These sentences were divided into four different lists, each being comprised of an equal distribution of sentence type.

In Experiment 1, native speakers of Japanese (N=24) listened to target sentences with congruent prosody (ES-EP and MS-MP), and those with incongruent prosody (ES-EP and MS-EP). Participants rated how good each sentence was using a scale of 1 (very bad) to 5 (very good) as quickly as possible (within 3 seconds). In Experiment 2, another 24 native Japanese speakers listened to these stimuli, and responded to a visually-presented “Yes/No” question asking whether a particular person is the agent of a particular action in the sentence as quickly as possible (within 5 seconds). Judgments and response times (RTs) were recorded for both experiments. In Experiment 3, another group of native Japanese speakers (N=20) read 84 visually-presented sentences and provided untimed goodness ratings, using the same scale as that used in Experiment 1.

3. Results

3.1. Experiment 1: Auditory Acceptability Rating Task (Timed)

Paired comparisons were conducted to test (1) the effect of prosody within the same sentence structure and (2) the effect of sentence structure within the same prosodic congruency. The paired t-test results between the sentences with congruent and incongruent prosodies within the same structures revealed that both ES and MS structures with the congruent prosody had significantly higher rating scores than those with incongruent prosody, both by subjects and items (ES-EP vs. ES-MP: t_(23) = 6.10, p < .001; t_(83) = 10.82, p < .001; MS-MP vs. MS-EP: t_(23) = 6.92, p < .001; t_(83) = 13.19, p < .001). When we examine whether these effects were also evident in RTs, those for the ES-EP were also significantly shorter than ES-MP sentences both by subjects and items (t_(23) = -3.74, p = .001; t_(83) = -2.96, p = .004); however, RTs for the MS-MP and MS-EP sentences were not different (t_(23) = -1.05, p > .30; t_(83) = 1.13, p > .26). It was observed that RTs for MS sentences were overall the longest regardless of the prosody. The acceptability rating and RT results together suggest that incongruent prosody will make sentence processing difficult regardless of sentence structures.

When the congruent prosodies were compared within the same sentence structure, ES-EP yielded significantly higher scores: t_(23) = 8.83, p < .001; t_(83) = 12.48, p < .001) and shorter RTs (ES-EP vs. MS-MP:
\[ t(23) = -4.21, p < .001; t(83) = -3.15, p < .003 \] compared to MS-MP. When incongruent conditions were compared, ES-MP was scored significantly higher than MS-EP \((t(23) = 10.76, p < .001; t(83) = 16.14, p < .001)\), but no difference was found in the RTs (ES-MP vs. MS-EP: \(t(23) = -1.10, p = .28; t(83) = -1.10, p = .28\)). These results together suggest that ES structure is overall preferred over MS structure, though incongruent prosody will increase the processing difficulty of even the ES structure (suggested by the longer RTs for ES-MP compared to ES-EP).

### 3.2. Experiment 2: Auditory Comprehension Task (Timed)

The same sets of paired \(t\)-tests were conducted for the comprehension task results. The comparison of the effect of prosodic congruency within the same sentence structures revealed that MS-MP accuracy was significantly higher than MS-EP \((t(23) = 9.47, p < .001; t(83) = 7.08, p < .001)\), and RT for MS-MP was also significantly shorter than its incongruent counterpart, MS-EP \((t(23) = -4.59, p < .001; t(83) = -4.45, p < .001)\). On the other hand, the accuracy rate for ES-EP was not significantly higher than ES-MP \((t(23) = 1.40, p > .17; t(83) = 1.37, p > .17\), although RT for ES-EP was significantly shorter than its incongruent counterpart ES-MP \((t(23) = -4.61, p < .001; t(83) = -6.74, p < .001)\). It should be noted that ES sentences yielded high accuracy rates regardless of the prosodic congruency.

When different sentence structures were compared within the same prosodic congruency, there were significant differences both in the accuracy rate (ES-EP vs. MS-MP: \(t(23) = 7.07, p < .001; t(83) = 7.39, p < .001\)) and the RTs \((t(23) = 11.38, p < .001; t(83) = 11.87, p < .001\) in the congruent conditions. In the incongruent conditions, there were also significant differences in both the accuracy rate (ES-EP vs. MS-MP: \(t(23) = 12.79, p < .001; t(83) = 13.24, p < .001\)) and the RTs \((t(23) = 7.58, p < .001; t(83) = 12.31, p < .001\). Similar to Experiment 1, it was evident that ES structure was preferred over MS structure regardless of the prosodic congruency.

### 3.3. Experiment 3: Visual Acceptability Rating Task (Untimed)

Paired \(t\)-tests were conducted for the two types of sentences (ES and MS). It was revealed that embedded-clause sentences were rated significantly higher than main-clause sentences \((t(19) = 7.21, p < .001; t(83) = 25.66, p < .001)\). These results are consistent with Experiment 1 and Experiment 2 results, showing that even with sufficient time for sentence analyses, native Japanese speakers have preference for ES structure over MS structure.

### 4. Discussion

All three experiments revealed a strong preference toward ES structure; that is, analyzing the NP-acc as a part of the embedded-clause is more favored than as a part of the main-clause. Also, Experiments 1-2 revealed a preference for prosody-structure congruent conditions over incongruent conditions. These results revealed both effects of structure and prosodic congruency in processing temporarily ambiguous Japanese sentences. Congruent prosody indeed facilitates processing, but prosody alone does not fully neutralize the effect of structure. These results together suggest a significant but limited role for prosody in resolving ambiguous sentence structure during on-line processing.

Furthermore, judgment and RT results are not always as informative as each other across experiments. It is suggested that employing cross-method investigations collecting various measures will give a more complete picture of the role of prosody in sentence processing.

### References


### List of technical Terms

- Prosody: Suprasegmental features of speech, such as rhythm, pauses, accents, amplitude, and pitch variation.
- Global Ambiguity: Ambiguity that is not solved by the end of the sentence. A globally ambiguous sentence has more than one interpretation at the end of the sentence.
- Temporary Ambiguity: Local ambiguity which is resolved by the end of the sentence.