

Ambiguity in Broad Focus and Narrow Focus Interpretation in Japanese

Kiwako Ito

Department of Linguistics
University of Illinois, at Urbana-Champaign
k-ito@uiuc.edu

Abstract

In languages such as English, German, and Dutch, accentuation on an argument is claimed to be obligatory for the expression of broad focus for the entire phrase carrying new information. In Japanese, which has lexical pitch accents, focus is marked by pitch range expansion instead. This paper examines whether prosodic prominence on just the argument leads to a broad focus interpretation in Japanese as well. Listeners' response time and conscious evaluation were measured while they listened to short dialogues. The results show interesting resemblance to those in English reported by Birch and Clifton (1995): as the answer to broad focus questions, listeners judged utterances with pitch range expansion for both the object and the verb as more appropriate than those whose pitch range was expanded for only the object or the verb. However, unlike in Birch and Clifton's results, all three prosodic patterns were accepted equally quickly in the discourse comprehension task. These findings suggest that focus may project from one prominent word to the entire verb phrase in Japanese, but the use of intonation to contrast focus may be restricted in Japanese due to the primary function of pitch accents to contrast lexical meanings.

1. Introduction

One primary function of prosody is to provide cues about the informational structure of discourse. In general, words carrying new or important information in a discourse become foci of the utterance, and thus tend to attract accents (Bolinger, 1972; Halliday, 1967). However, a specific accentuation pattern does not always lead to a single interpretation of the informational status of the utterance. Consider, for example, the relation between the accentuation pattern and the interpretation of what is important or under focus in the utterance in example (1).

She cleans our **BEDROOM**. (1)
(The capitalized word indicates the location of the primary accent in the sentence.)

Sentence (1) may serve as a good answer to *Does your mother-in-law clean your bathroom every morning?* but it may also be a possible answer to *What does your mother-in-law do to annoy you so much?* In the former case, the primary accent in (1) marks narrow focus on *BEDROOM*, whereas in the latter case, the whole verb phrase *cleans our BEDROOM* carries focus. That is, a single accent on the object noun may lead to two different interpretations of focus depending on the preceding context. What this example demonstrates is an ambiguity between narrow focus and broad focus interpretations of an utterance with a single accent in a predicate phrase (i.e., VP). Such ambiguity in pragmatic interpretation of single accent has been known traditionally as

the *focus projection* phenomenon, emphasizing that focus expressed by a single accent seems to project to a larger linguistic constituent than just the word with the accent. Researchers such as Selkirk (1984, 1994) and Gussenhoven (1983, 1992, 1994, 1999) have developed linguistic theories that describe the relation between the syntactic/argument structure and regularities in focus-related accentuation (mostly in English, German and Dutch).

Birch and Clifton (1995) tested a set of hypotheses derived from the above theories about the relation between intonational focus marking and the interpretation of utterances in English using two different tasks. First, listeners were asked to decide as quickly as possible whether each question-answer pair made sense as a conversation. They accepted singly accented utterances such as *She teaches MATH* as quickly as they accepted doubly accented utterances such as *She TEACHES MATH* as answers to broad focus questions such as *Isn't Kerry pretty smart?* Second, listeners were asked to numerically rate the appropriateness of the intonational pattern of the answer in each dialogue. In this task, listeners rated doubly accented utterances higher than singly accented utterances when they were heard as answers to broad focus questions. Thus, Birch and Clifton experimentally demonstrated that although double accenting on both the verb and the object NP is preferred for the expression of broad focus over the VP, single accentuation on the object NP is processed equally quickly during speech comprehension.

The present study investigates whether there is a similar ambiguity between broad and narrow focus interpretation of an utterance with a single intonational prominence in Japanese. Japanese is an ideal language to test the generality of Birch and Clifton's results because the use of pitch accents is functionally different from English. Unlike English in which pitch accents are placed according to the pragmatic structure of an utterance, pitch accents in Japanese are lexically assigned and thus contribute to meaning distinctions between otherwise identical word forms (e.g., *Ha na* a girl's name Hana vs. *hana* nose: an apostrophe indicates a pitch accent on the preceding mora).¹ A pitch accent in Japanese is realized as a sharp F0 fall from the accented mora to the following mora.² Thus the melody of each word is lexically determined and it does not change according to discourse

¹ The Japanese lexicon contains accented and unaccented words and they behave differently in prosodic phrasing (Pierrehumbert and Beckman, 1988). The present study does not contain unaccented words in the target experimental phrases.

² Example productions of accentual minimal pairs (i.e. homophonous words differentiated only by accentual patterns) can be heard alongside the graphic f0 track at <http://www.sip.uiuc.edu/j-hualde/japanese/accnt-jpse.html>
Accentual contrasts in Japanese by Hualde, J.I & Ito, K.

pragmatics. Focus in Japanese is therefore expressed by enhancing the lexically predetermined pitch movement for the word under focus (Ito, 2001; Pierrehumbert and Beckman, 1988). It is important to test whether such pitch enhancement in Japanese functions in the same manner as the single pitch accent placement in English for the expression of broad focus.

2. Experiments

The methodology used by Birch and Clifton (1995) was adopted in the present study with Japanese listeners. First, subjects listened to question-answer pairs and made acceptability judgments about each one of them. Second, they performed conscious ratings on the intonational patterns of the answer of each dialogue. Twenty-one sets of short dialogues were prepared by combining a broad focus question with three types of answers differentiated only by the intonational pattern, as exemplified in Table 1.

Table 1: Example experimental item set.

Question	Three answers
Q. <i>yoko yamakun-wa bo onasu moratta ra do osuru-no?</i>	A1. <i>ka re-wa DA IBINGU- O HAJIMERU -N-DA- YO.</i>
What will Mr. Yokoyama do when he gets a bonus?	He STARTS (scuba) DIVING.
(Note: A1's pitch contour showed two clear tonal humps with a downstep, whereas A3's verb had a higher peak than the object. A2 had a clear tonal hump over the object followed by a low tail to the end of sentence. See Figure 1 for examples.)	A2. <i>ka re-wa DA IBINGU- O hajimeru -n-da-yo.</i>
	He starts DIVING.
	A3. <i>ka re-wa da ibingu-o HAJIMERU -N-DA- YO.</i>
	He STARTS diving.

2.1. Experiment I

In Experiment I, 21 target dialogues were randomly intermixed with 21 other experimental dialogues and 84 filler dialogues.³ Three lists consisting of 126 question-answer (q-a) pairs were prepared, so that each list contains seven q-a pairs in each of the three q-a combinations described above. Half of the fillers did not make sense as conversations. The purpose of these fillers was to prevent listeners from paying attention only to intonation during the acceptability judgment task. Dialogue (2) is an example of nonsensical question-answer pairs.

³ The additional 21 experimental items were prepared to investigate the interpretation of the intonational pattern for given information. Those data are not discussed in the present paper.

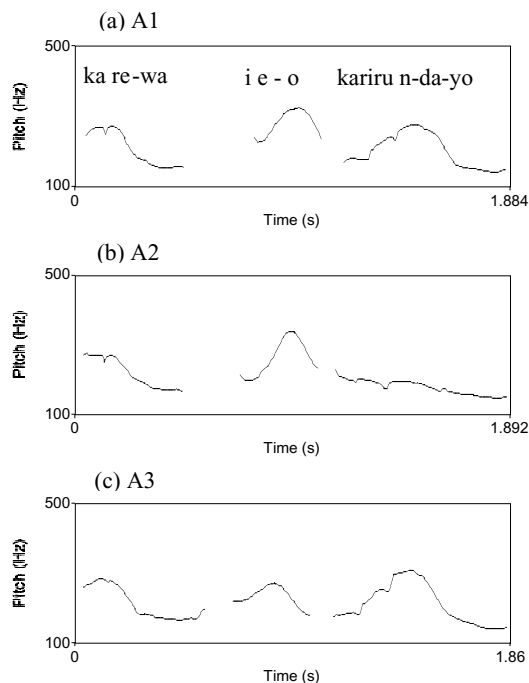


Figure 1: Example F0 tracks of stimuli: *ka re-wa, ie -o kariru -n-da-yo.* gloss: *he-TOP house-ACC rent-COMP-COP-EMPH He rents a house.* (a) Broad focus reading: A1, (b) Object focus reading: A2, (c) Verb focus reading: A3.

Q: *sono ka sa, do ko-de katta -no?*
that umbrella where-LOC bought-Q?
Where did you buy that umbrella?

A: *kinoo A ME-GA fu tta-n-da-yo.*
yesterday rain-NOM fell-COMP-COP-EMPH
It rained yesterday. (2)

Subjects were asked to judge whether each dialogue made sense as quickly as possible. Reaction Times (RTs) and acceptability rates (i.e., %YES) were measured. If Japanese listeners interpret intonational enhancement in the same manner as English listeners interpret pitch accents, A1 and A2 should be responded to equally more quickly than A3, and the acceptability rate of A1 and A2 should both be higher than that of A3, which violates the rule that the internal argument under focus must be intonationally prominent (Selkirk, 1984, 1994; Gussenhoven, 1983, 1999).

2.2. Experiment II

In Experiment II, subjects listened to the same experimental dialogues randomly intermixed with 21 filler dialogues. None of the filler dialogues in this second study were nonsensical. Subjects were asked to consciously rate the appropriateness of the intonational pattern of the answer sentence for its context on a scale from 1 (totally inappropriate) to 5 (totally appropriate). If full intonational prominence is preferred over partial prominence for broad focus utterances, A1 should be rated higher than A2 and A3.

2.3. Subjects

A total of 64 native Japanese speakers participated in both experiments. Half of the subjects were recruited at University of Illinois at Urbana-Champaign, while the other half were recruited at Ohio State University. Each subject participated in Experiment I first. The dialogues were recorded by two female speakers (i.e. one native speaker of Tokyo Japanese and the author) in standard (Tokyo) Japanese.

2.4. Material Preparation and Experiment Procedure

The materials were first recorded on a DAT tape at 48 KHz using a SONY DTC-2E700, and then down-sampled at 16 KHz using *Sound Designer II*. The auditory stimuli were presented through headphones attached to a computer. In Experiment I, subjects responded to each dialogue by pressing a button on a button box. In Experiment II, they gave their ratings by pressing a number key on the computer keyboard.

3. Results

In Experiment I, the mean reaction time (RT) for each condition was calculated for YES (i.e. acceptable) responses. NO responses were excluded from the calculation of RT. For each subject, RTs exceeding 2.5 standard deviations (*SD*) above or below the subject's mean RT were replaced with the value equal to the mean \pm 2.5 *SD*. Ten subjects' data were excluded either because too many RT replacements were required or because there were too few YES responses. The grand mean RT, %YES (Experiment I) and the rating (Experiment II) across the three experimental conditions were 287ms, 93%, and 3.9, respectively. The mean RT, %YES, and the rating for the sensible fillers were 322ms, 94%, and 4.0, respectively, whereas the nonsensical fillers had a mean RT of 634ms and %YES of 13%. Table 2 shows the mean RT, %YES, and rating for each condition (A1-A3).

Table 2: Mean RTs & %YES from Experiment I, and mean ratings from Experiment II. Standard errors are in the parentheses.

Condition (answer type)	Reaction Time (ms): Exp I	% YES: Exp I	Rating (1 - 5): Exp II
1. A1 With prominent object & verb	286 (33)	94 (1.4)	4.1 (.1)
2. A2 With prominent object	289 (32)	92 (1.3)	3.8 (.09)
3. A3 With prominent verb	299 (33)	92 (1.4)	3.9 (.08)

The mean RTs, %YES and ratings were obtained both by subjects and by items and submitted to repeated measures ANOVA. Although no effect of answer type was found in either RT or %YES measurement, there was a significant main effect of answer type on ratings, $F_1(2,106)=5.04, p<.01$; $F_2(2,82)=3.94, p<.05$. To test the hypothesis that utterances with a single prominence on the object were equally acceptable as those with dual prominence on the object and the verb, conditions A1 and A2 were compared on all three measurements. Paired *t*-tests show no difference between the

two in RT and %YES. However, the difference of .3 in mean ratings was statistically significant, $t_1(53)=3.22, p<.01$; $t_2(41)=2.73, p<.01$.

Surprisingly, the utterances with single prominence on the verb did not lead to difficulty during dialogue comprehension. Neither RT nor %YES of A3 differed from A1 and A2. Also, the mean rating of A3 was comparable with that of A2 (A3's rating was significantly lower than that of A1, $t_1(53)=2.08, p<.05$; $t_2(41)=2.21, p<.05$), suggesting that the utterances with a single intonational prominence on the verb were judged as relatively good answers to broad questions.

4. Discussion

The above results suggest that though dual prominence is preferred for answers to broad focus questions, utterances with a single intonational prominence on the object may be comprehended equally quickly as those with dual prominence. The observed difference between the RT and the conscious rating measurements was also reported in Birch and Clifton (1995). However, in the present study, utterances with pitch expansion only for the verb (A3) were also responded to equally quickly as those with dual prominence. This finding contrasts with the results in Birch & Clifton, in which listeners were reliably slower in accepting dialogues with A3 type answers.

As shown by the examples in Figure 1, the target utterances used in the present study had clearly distinct intonational patterns across the three conditions. In fact, the A3 utterances had a significantly lower mean pitch peak for the object and a significantly higher mean pitch peak for the verb than the pitch peaks averaged across A1 and A2 (Diff_{obj}= 51Hz: $t_{obj}(41)=21.37, p<.000$; Diff_{verb}= 70Hz: $t_{verb}(41)=37.79, p<.000$). Thus the lack of differences in RT and %YES between A3 and the other two conditions was not due to acoustic similarities across conditions. The current findings rather seem to suggest that the intonational enhancement or pitch range expansion for the internal argument (i.e. object NP) is not obligatory for the expression of broad focus in Japanese.

It is possible that the relation between argument structure and intonational focus marking is not universal, so that the prominence of an internal argument is required in some languages but not in others. However, we should not disregard the important difference in the intonational contour of a verb focus utterance between English and Japanese. While the object noun following the focused verb (e.g. *math* in *She TEACHES math*) may be completely deaccented in English, the object preceding the verb may never lose its lexical pitch movement in Japanese, as shown in Figure 1(c).⁴ Therefore, the Japanese listeners in the present study may have rated the verb focus utterances as highly as other answers to broad focus questions because the internal argument (i.e. object NP) did not completely lack intonational prominence. Since pre-focal words are often dropped in Japanese, it is also possible that the presence of the pre-focal object itself reduced the perception of the narrow focus on the verb, thus making those utterances be interpreted as broad focus answers. The broad focus questions that provided a discourse context for each dialogue mentioned the topic (i.e.

⁴ The pitch movement of post-focal words may instead be extremely reduced in Japanese, as has been found in many other languages (see Figure 1b).

the subject) but neither the object nor the verb. Thus the object in the answer utterances was introduced always as a new entity in the discourse. Listeners may have paid more attention to the word meaning than to its prosodic prominence of a new discourse entity during the comprehension task. This suggests that listeners' sensitivity to focus-related intonation may be conditioned by discourse contexts. A parallel study is under way in order to test whether the informational status of a word (i.e. new vs. given) affects listeners' attention to its prosodic prominence.

5. Conclusion

The present study demonstrates that a single intonational prominence in a predicate phrase may lead to an ambiguity between broad focus and narrow focus interpretation in Japanese. Japanese listeners showed preference for dual prominence for broad focus answers similarly to English listeners in Birch and Clifton, but utterances with a single intonational prominence were comprehended as broad focus answers just as quickly as utterances with dual prominence. Unlike Birch and Clifton's results, Japanese utterances with intonational prominence on the verb were also accepted as good answers to broad focus questions. This may indicate limitations in the use of intonational focus marking in Japanese due to the primary function of pitch accents to contrast lexical meanings. Such language-specific prosodic characteristics may interact with syntactic features such as pro-dropping, and informational status of words may also influence the way prosody is processed during speech comprehension in Japanese. The relationship among the various factors influencing intonational focus marking in Japanese remains as an open question for future research.

6. References

- [1] Birch, S.; Clifton, C., 1995. Focus, Accent, and Argument Structure: Effects on Language Comprehension. *Language and Speech*, 38(4), 365-391.
- [2] Bolinger, D., 1972. Accent is predictable (if you're a mind-reader). *Language*, 48, 633-644.
- [3] Halliday, M., 1967. Notes on transitivity and theme in English, part 2. *Journal of Linguistics*, 3, 199-214.
- [4] Selkirk, E., 1984. *Phonology and Syntax. The relation between sound and structure*. Cambridge, MA: MIT Press.
- [5] Selkirk, E., 1994. Sentence Prosody: Intonation, stress and phrasing. In *Handbook of phonological theory*, Goldsmith, J. ed. Oxford, U.K.: Blackwell, 550-569.
- [6] Gussenhoven, C., 1983. Testing the reality of focus domains. *Language and Speech*, 26, 61-80.
- [7] Gussenhoven, C., 1994. Focus and sentence accents in English. In *Focus and natural language processing*, Vol. 3, Bosch, P.; van der Sandt, R., eds. Heidelberg: IBM Deutschland, 83-92.
- [8] Gussenhoven, C., 1999. On the Limits of Focus Projection in English. In *Focus Linguistic, Cognitive, and Computational Perspectives*, Bosch, P.; van der Sandt, R., eds. Cambridge University Press, 43-55.
- [9] Ito, K. 2001. ms. The effect of pragmatic focus on lexical and phrasal accents in Japanese.
- [10] Pierrehumbert, J.; Beckman, M., 1988. *Japanese tone structure*. Linguistic Inquiry Monograph Series, 15, Cambridge, MA: MIT Press.