Broad and Narrow Focus Perception in Japanese Learners of English before and after a Short-Term Study Tour

Pinter, Gabor
Kihara, Emiko

(Citation)
神戸大学国際コミュニケーションセンター論集,12:30-39

(Issue Date)
2015

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/81009378

(URL)
https://hdl.handle.net/20.500.14094/81009378

Broad and Narrow Focus Perception in Japanese Learners of English before and after a Short-Term Study Tour

Gábor Pintér[1], Emiko Kihara[2]

1. Introduction

Pronunciation training is a notoriously difficult domain in foreign language teaching. Appropriate training requires well-trained instructors, individual attention, careful feedback and long training on behalf of the students[3]. Unlike reading comprehension exercises, pronunciation training cannot be done effectively in an autonomous manner. Unlike vocabulary practice (cf. Quizlet®, Memrise®), it cannot be automatized easily either. Although commercial Computer Aided Pronunciation Training (CAPT) solutions are available (e.g., AmiVoice® CALL, Rosetta Stone®), they are few in number and—partly for financial, partly for technical reasons—scarce in availability (Pinter 2010). Partly for these reasons, Japanese EFL learners most often do not receive thorough pronunciation training. Without explicit training, learners are in a difficult position to develop *phonological awareness*, which is a known facilitator in the training of pronunciation, reading and listening skills (Li et al. 2012).

The current study investigated a relatively little taught and researched aspect of phonological awareness: the perception of focus. The foremost goal was to investigate if Japanese learners of English could perceptually discriminate focus patterns differing in locations and scope. Another, more pedagogically motivated goal was to test if perception strategies changed after a three-week long exposure to native language. Our department, the *School of Languages and Communication*, organizes a study abroad program for undergraduate students every year. The curriculum of this program usually does not feature pronunciation related training, so we were interested to see if perceptual strategies can improve or change *spontaneously*, merely through exposure to the target language, with no explicit training. For this investigation, prosodic features, such as focus, were considered to be better candidates than segment or syllable level structures, because awareness about the latter ones were not expected to improve in the short span without explicit training.

The results of pre-departure and post-arrival experiments revealed that the participants successfully differentiated four different focus scopes both prior to and following the study tour. Furthermore, the results showed that on average, focus categories were differentiated slightly more reliably, that is, with larger margin after the study tour.

---

[1] 神戸大学国際コミュニケーションセンター　g-pinter@port.kobe-u.ac.jp
[2] 神戸大学国際コミュニケーションセンター　emiwamoto@aquamarine.kobe-u.ac.jp
[3] Acknowledgements: Special thanks to Marian Wang and Gregory Sholdt for their help in the preparation of this article. This work was supported by JSPS KAKENHI Grant Number 15K02480 awarded to Shinobu Mizuguchi.

2. Prosodic Marking of Focus

2.1 Prosodic Focus Marking in English and Japanese

From an information-structural point of view the term focus is defined as the part of the sentence that introduces novel information to the discourse (Lambrecht 1994). Prosody is only one measure languages can utilize to express focus. Morphological markers (e.g., the Japanese *ga* particle), or specific syntactic arrangements (e.g., the Hungarian pre-verbal focus position) can also express focus.

Studies concentrating on the relation between focus and prosody are typically concerned with the following questions (Breen et al. 2010:1044). First, if focused elements in the discourse are marked and perceived prosodically, that is through acoustic features. Second, if different focus types (e.g., contrastive versus informational focus) have different acoustic and perceptual realizations. Third, if there are prosodic differences between different focus scopes. This latter is the main concern of the present paper.

The scope of the focus refers to the portion of the utterance[4] that carries new information. For example, the sentence "I bought a motorcycle" can express various assertions—it can function as an answer to at least the following three different questions (Bishop 2012).

(1) Focus scopes

	Question		*Response*	*Focus Scope*	
a.	What happened?	→	[I bought a motorcycle]	sentence	(broad)
b.	What did you do?	→	I [bought a motorcycle]	VP	
b.	What did you buy?	→	I bought a [motorcycle]	object	(narrow)

In case the focused element is a single element (e.g., the object), it is referred to as *narrow focus*. In contrast, *broad focus* refers to cases when focus extends over several words. The different focus scopes are marked prosodically, through acoustic features. Narrow focus is typically produced with higher pitch, stronger intensity, and longer vowel duration than broad focus (Breen et. al. 2010: 1065). For narrow focus, F0 and amplitude peaks are higher and steeper than for broad focus. In general: "the narrower a focus […] the more relatively prominent it should be" (Bishop 2011:314; see also Baumann et al. 2006).

The acoustic distinction carries over to perception. Starting from Guessenhoven (1983), several studies showed that narrow focus is reliably distinguished from neutral and broadly focused utterances in perception. Although the inverse correlation between the focus breadth and acoustic prominence is widely accepted, the readings of acoustic variables and the size of focus scope do not necessarily covary in a monotonic fashion. For example, the convincingly identified acoustic correlates of focus domains in Baumann et al. (2006) show some irregularities when comparing different broad focus types. As for perception, Bishop (2011) did not find a statistically significant perceptual difference between different levels of broad scopes: there is no statistical difference between responses to *VP focus* versus *sentence focus* stimuli.

These observations, concerning the prosodic marking of focus, are drawn from studies targeting mainly English and West-Germanic (e.g., German, Dutch) languages. Focus marking is somewhat different in Japanese, as narrow focus is marked morphologically with the particle *ga* (Kuno 1973), but prosodic features are also used. For Tokyo Japanese, Pierrehumbert and Beckman (1988) showed that Japanese focus also has prosodic marking,

[4] In more formal terminology *focus scope* can also refer to the size of the *focus constituent* or the size of the *focus domain*.

which is expressed acoustically through expanded F0 and compressed post-focus components. Hwang (2016) showed that not only post-focus, but pre-focus elements also undergo compression, which is also used in focus perception.

2.2 Prosodic marking in Japanese listeners of English

Research on L2 prosody perception is not in the mainstream of either prosody or L2 studies. Accordingly, literature on L2 focus perception is somewhat scarce. One research direction that has recently gained considerable attention is hallmarked by Jennifer Cole (University of Illinois) and her co-operators, and uses the experiment design called Rapid Prosody Transcription (RPT). In RPT studies the non-trained participants are asked to carry out boundary and prominence perception tasks. The stimuli used for these experiments typically are created from spontaneous speech. The RPT paradigm is a good candidate for adaptation to L2 studies, as it does not require trained transcribers and explicit knowledge of prosody. Using the RPT method, Pinter et al. (2014b) found that boundary perception patterns in Japanese listeners correlate relatively strongly with native patterns. In case of prominence detection, the L1-L2 correlation is poor, and the perception is hardly consistent (*Fleiss' kappa* = 0.272).

These results are in contrast with findings of Yamane et al. (2016). Using the RPT method in a three-way design of production, perception and comprehension of prominence, Yamane and her colleagues found that Japanese learners of English performed consistently and better than native speakers in prominence detection tasks. This result is also in contrast with some results of the same study reporting poor results for Japanese speakers in production tasks. Japanese learners in the production experiments did not seem to utilize prosodic features for focus marking.

2.3 Research Questions

The poor and non-consistent performance of prominence detection in Japanese learners of English in Pinter et al. (2014) is in contrast with their reportedly good performance in lexical stress (Sugahara 2016:62) and focus detection (Yamane et al. 2016). One possible explanation for this discrepancy centers on the different types of stimuli used in the experiment. Short, highly controlled laboratory speech stimuli in Yamane et al. (2016) can—to some extent—account for the better perceptual performance, compared to the tasks with spontaneous stimuli of Pinter et al. (2014b). Leaving aside the discussion of asymmetric language processing demands these stimuli types exert on the participants, differences in terms of focus scopes could also contribute to the seemingly contradicting results. The utterances extracted from the directed mini-dialogues in Yamane et al. (2016) featured more narrowly focused elements than the spontaneous utterances in Pinter et al. (2014b). Since narrowly focused words are acoustically more prominent, their perception is expected to be easier and more consistent across listeners. This explanation implies that Japanese listeners—similarly to native speakers—can identify narrow foci more easily than broad ones.

The current study aimed to investigate this line of reasoning by conducting a prominence perception experiment using variable domains of focus. By and large, the experiment was created as a variation on the prominence marking perceptual task presented in Bishop (2011). For Japanese listeners it was expected that narrowly focused words would be identified more correctly and more uniformly than broadly focused elements.

Also, following the results of Bishop (2011), Japanese listeners were expected to show low performance on differentiating within broad focus cases.

Since the participants for this study were to participate in a three-week study program at a university in the USA, it was a good opportunity to test if exposure to native language improves or changes learners' focus perception strategies. A post-arrival experiment result showing decreased confusion rates and variation would point to increased levels of prosodic awareness.

3. Experiment

Two experiments were conducted framing a three-week study tour to Washington University taking place in August 2015. The pre-test was conducted 2 weeks prior to the departure. The post-test was conducted within two weeks the students' return to Japan. The curriculum of the language training program did not feature explicitly any pronunciation related courses or activities.

3.1 Participants

Twenty-three undergraduate students, signed up for the US study tour, were invited to participate in the experiment. Out of the 23 students, 22 participated in both experiments. There was one student who participated only in the post-test, her responses were removed from the data.

3.2 Stimuli

The utterances for the experiment were excerpts from mini-dialogues consisting of pairs of simple questions and answers. An adult female native speaker of North-American English read the answer parts of the dialogues in response for the questions read by the author. The utterances were recorded in a quiet room using a Tascam DR-5 audio recorder.

(2) Dialogues			*Focus scope*	
a.	**Q**	*What did you do yesterday?*	`s1:`	`scope 1`
	A	I [saw a bad film]$_{VP}$ yesterday.	→	Verb Phrase
b.	**Q**	*What did you see yesterday?*	`s2:`	`scope 2`
	A	I saw [a bad film]$_{NP}$ yesterday.	→	Object
c.	**Q**	*What film did you see yesterday?*	`s3:`	`scope 3`
	A	I saw a [bad]$_{Adj}$ film yesterday.	→	Modifier
d.	**Q**	*When did you see a bad film?*	`s4:`	`scope 4`
	A	I saw a bad film [yesterday]$_{Adjunct}$.	→	Adjunct

The questions in the mini-dialogues were designed in a way that they elicited answers with four different focus scopes overlaid on identical sequences of words. The four different focus patterns had scopes over: the whole verb phrase (`s1:` VP); the noun phrase corresponding to the object (`s2:` Obj); the adjective modifier (`s3:` Mod); and the adjunct (`s4:` Adj). The focused terms are represented by [brackets] in (2). The three focus conditions of VP > Obj > Mod form proper subsets of scopes (see `[saw [a [bad ]`$_{Mod}$ `film]`$_{Obj}$`]`$_{VP}$).

The four types of focus patterns were combined with five different sentences resulting in 20 utterances.

(3) Stimuli sentences		*Subject*	*Verb*		*Modifier*	*Object*	*Adjunct*
	1.	I	saw	a	a bad	film	yesterday

2.	bought		a pink	hat
3.	found		a large	book
4.	made		a sweet	cake
5.	learned		a mean	trick

3.3 Design

The answer parts of the mini-dialogues—in sum (4 *focus type* × 5 *sentence*=) 20 utterances—were extracted and presented in a pseudo-random order over a loudspeaker in a classroom. The participants were asked to listen to the sentences and rate how much "stress" or "prominence" they perceive on each word. A sheet with sentences printed in the pre-defined random order were handed out, and the participants were asked to rate the perceived stress by writing a number between 1 and 5, where 5 corresponded to the strongest stress. The participants were not required to mark every word in the sentence. One sentence from the stimuli set was used as practice. There were no filler sentences in the session. The same procedure was used for both the pre-test and the post-test.

(4) Sample task with responses

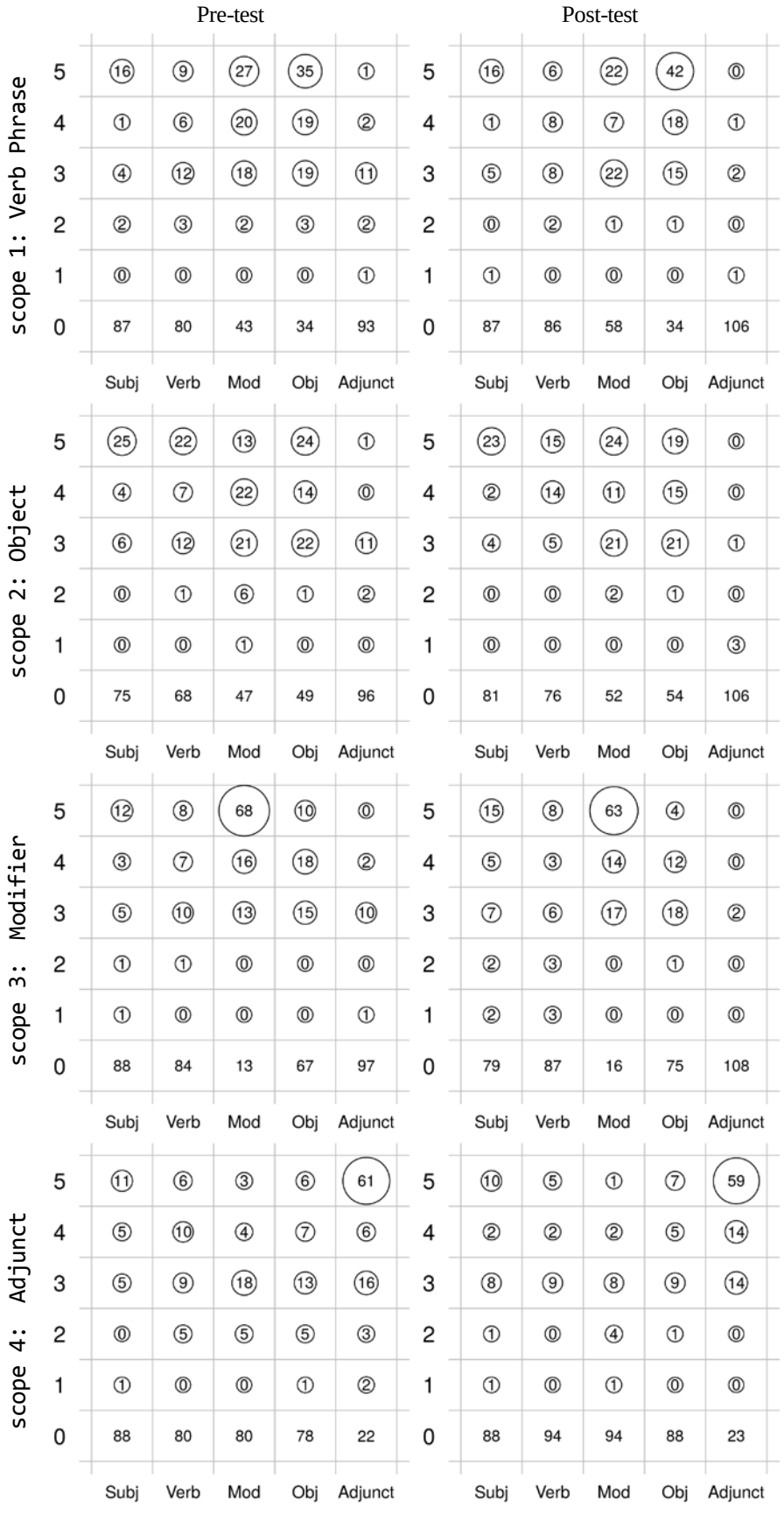
		5		2
I	*saw*	*a bad*	*film*	*yesterday*

The procedure of the experiment was inspired by Bishop (2011), but it diverged from the original work at several points. First, Bishop (2011) used only subject-verb-object sentences with no modifiers or adjuncts. Second, in his experiment, only verbs and objects were used for stress evaluation. Third, stress marking was compulsory for both words.

4. Results

Table 1 summarizes the responses in the pre- and post-tests. Corresponding pre- and post-test results are aligned horizontally. The four rows are for the four focus scope conditions: VP, Obj, Mod, Adj. The vertical axes in each table display a 0 to 5 stress perception scale. Zero means no response. The circled numbers represents the number of received marks. For example, in the modifier-scope condition of the pre-test, the modifiers (*bad, pink, large, sweet, mean*) received 68 ratings of 5s in sum over the five sentences. The size of the circles expresses the number of ratings.

Table 1. Distribution of perceived stress levels



4.1 Focus scope perception

From a visual inspection it is immediately apparent that narrowly focused elements, that is, the modifiers in modifier scope and the adjunct in the adjunct scope utterances, received higher stress marks in both the pre- and the post-test. This visual presentation does not tell us, however, if VP and Object scope conditions are discriminated or not. Figure 1 and 2 present a more compact visualization of the data using the average scores over each word—pooled over 5 sentences within the each scope condition. In these line charts, similarly to the tabular data, the narrowly focused elements stand out, but the broad focuses contours of `s1` and `s2` are rather similar, especially in the pre-test.

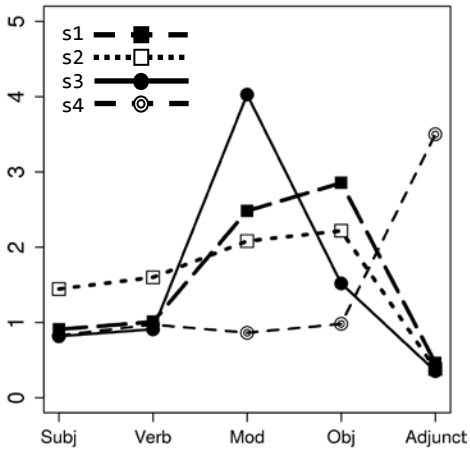


Figure 1. Average stress scores in pre-test

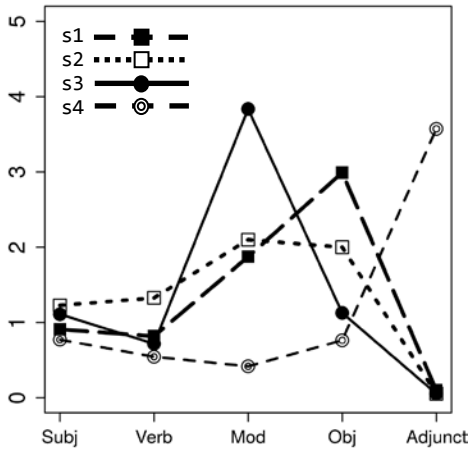


Figure 2. Average stress scores in post-test

In order to verify the perceptual difference between `s1` and `s2` conditions, a Wilcoxon test was carried out using prominence ratings over the object constituent.[5] The results show that markings on the object are statistically different in both pre-test (p = .0239) and post-test (p = .0003), and the difference is more significant in the post-test ($p < .05$ versus $p < .001$).

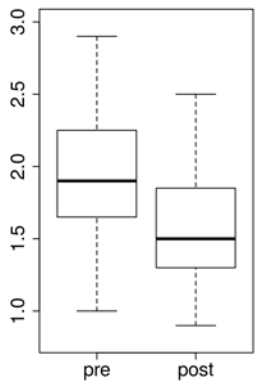


Figure 3. Average number of stress marks per sentence

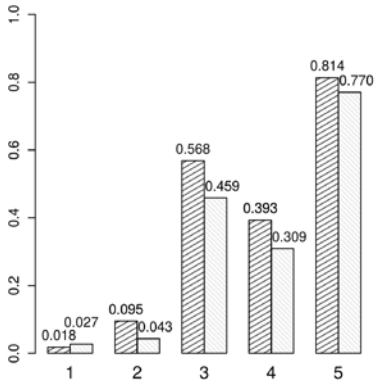


Figure 4. Average distributions of the 5 stress levels per sentence

[5] Simple t-test was not appropriate as the distributions of ratings were highly skewed.

4.2 General differences between pre- and post-test

One of the most apparent differences between the two runs of the experiment was that the number of assigned marks was significantly lower in the post-test. Figure 3 displays the average number of stress marks per sentence in the pre-test and the post-test. On average, 1.84 non-zero marks were assigned to every sentence in the pre-test, but only 1.61 marks in the post-test. Running a paired t-test on the number of marks on each of the 20 sentences verifies that this decrease is statistically significant (*df*=19, *p* = .0492). Figure 4 presents further details about the distributions of the marks. It can be seen that the general decrease in marks is due to fewer marks for levels of 2, 3, and 4. Ratings of 1 actually increased.

Not only the number of marks, but also rating variances decreased between the two runs. Figure 5 shows the marks' standard deviations for each element in the four scope conditions. A paired Wilcoxon signed rank test over the SD values verifies that the decrease is statistically significant at an alpha level of 0.05 (V=44, *p* = .0108).

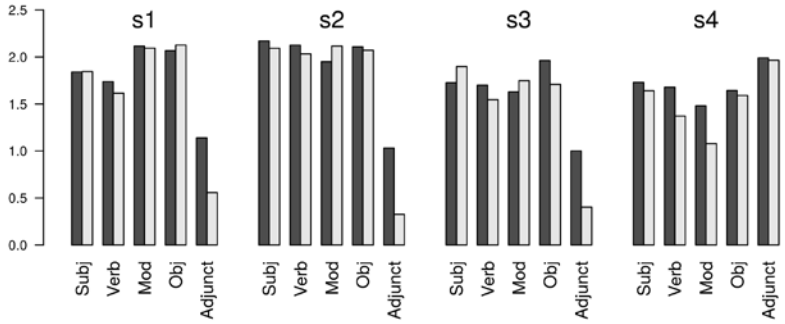


Figure 5. Changes in standard deviations pooled over 5 sentences and 22 participants

4.3 Scope conditions

No pre-test/post-test changes were expected in the perception of narrow focus elements as Japanese learners' performance of detecting English narrow focus was reportedly reliable (Yamane et al. 2016). As it is apparent in Figure 6, the expectation was correct. No changes were found in perceived prominence levels in case of the two narrow focus conditions (`s3`: modifier focus; `s4`: adjunct focus).

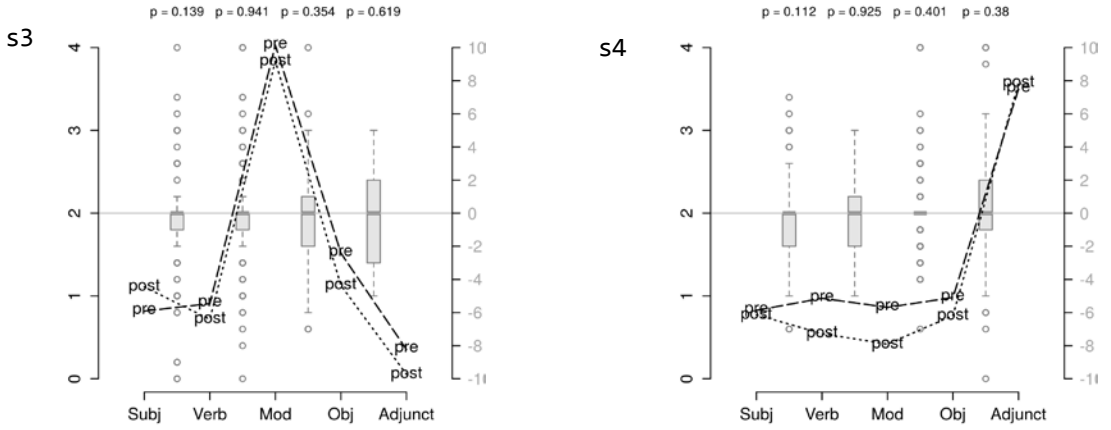


Figure 6. Changes between pre- and post-test in stress mark averages in *narrow focus* conditions

In Figure 6, the boxes between words represent the distribution of differences in the deltas of the stress scores of neighboring words. Note that these values are not the differences between neighboring stress scores. They are the differences between pre-test delta scores and post-test delta scores. For example, in the `s4` condition, the

difference of scores between the object and the adjunct increased in the post-test. That is why the box between those words is skewed upward, in a positive direction. This enhanced difference is due to the lower values on the object in the post-test. None of these changes were found to be significant.

A more interesting question was if responses had changed in the broad focus conditions. The only statistically significant change was an enhanced contrast between the modifier and the object in the VP focus (`s1`) condition (p = .041). This change was due to lower scores for the modifier in the post-test. The perceptual salience of the object was also increased by an enhanced contrast with the following adjunct. The adjunct—similarly to the modifier—received lower scores in the post-test leading to a greater perceptual contrast with the object. The change on the right side of the object, however, was not significant over the repetitions (p = .094).

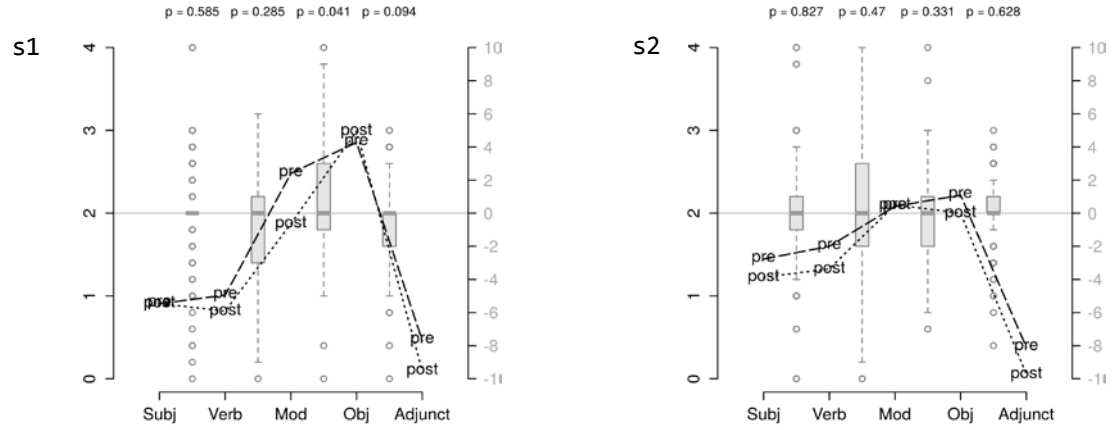


Figure 7 Changes between pre- and post-test in stress mark averages in *broad focus* conditions

5. Discussion

The results of the perceptual experiment imply that Japanese learners of English can easily identify narrow focus sentence elements. Broad focus elements are perceptually less salient. Although the difference in perceived prominence over the object nouns was statistically significant (p = .0239), the difference is significant only at alpha level of 0.05. In general, responses to broad focus utterances (VP focus, object NP focus) yielded rather similar perceptual profiles. Based on these facts it can be claimed that Japanese learners of English have better performance in narrow focus perception than in broad focus perception. This result can explain the seemingly contradicting results of Yamane et al. (2016) and Pinter et al. (2014b). In the former case, the consistent and accurate detection of prominence by Japanese EFL speakers is attributable to the use of narrowly focus utterances in the stimuli. Since narrowly focused elements occur less frequently in spontaneous speech, and the perception of broad focus is less consistent, it is not surprising that Japanese EFL speakers achieved low agreement in a prominence perception in Pinter et al. (2014b).

As expected, differences between results of the pre-test and the post-test were not remarkable. One statistically significant change is that fewer marks were assigned in the pre-test. Also, the responses show less variation. The perceptual salience of the focused modifier is enhanced in the post-test. All of these changes, together with others that did not pass statistical significance, imply that the group became more consistent in focus perception, which can be an implication of increased phonological awareness. The result is not conclusive

though.

The current study is only a preliminary investigation in the topic of L2 stress scope perception. A more complete study should involve a control group that is not exposed to extensive English between the two experiment runs. It is possible, that results discussed above are only effects due to experiment repetition. Also the stimuli needs further investigation and more careful selection, as it comes from a single informant. Re-running the experiment with native speakers could also provide a direct way to compare prosody perception in L1 versus L2 listeners.

References

Baumann, Stefan, Martine Grice and Susanne Steindamm (2006). Prosodic marking of focus domains: Categorical or gradient? *Proceedings SpeechProsody* 2006, 301–304.

Breen, Mara, Evelina Fedorenko, Michael Wagner and Edward Gibson (2010). Acoustic correlates of information structure. *Language and Cognitive Processes* 25, 1044–1098.

Bishop, Jason (2011). English Listeners' knowledge of broad versus narrow focus contrast. *Proceedings of the International Congress of Phonetic Sciences* 18, 312–315.

Bishop, Jason (2012). Information structural expectations in the perception of prosodic prominence. *Prosody and Meaning*, ed. by G. Elordieta and P. Prieto, Berlin: Mouton de Gruyter. 239–269.

Hwang, Hyun Kyung (2016). The role of pre-focus compression in Tokyo Japanese. *Phonological studies* 19, 117–124.

Gussenhoven, Carlos (1983). Testing the reality of focus domains. *Language and Speech* 26, 61–80.

Lambrecht, Knud (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse*. Cambridge: Cambridge University Press.

Li, Miao, Liying Cheng and John R. Kirby (2012). Phonological awareness and listening comprehension among Chinese English-immersion students. *International Education* 41.2, 46–65.

Pierrehumbert, Janet and Mary Beckman (1988). *Japanese Tone Structure*. Linguistic Inquiry Monograph 15, Cambridge: MIT Press.

Pintér, Gábor (2010). Technical report about the current limitations of computer assisted pronunciation training systems. *Bulletin of School of Languages and Communication*, 7, 1–6.

Pintér, Gábor, Shinobu Mizuguchi and Kazuhito Yamato (2014). Boundary and Prominence Perception by Japanese Learners of English: A preliminary study. *Phonological Studies* 17, 59–66.

Pintér, Gábor, Shinobu Mizuguchi and Koichi Tateishi (2014). Perception of prosodic prominence and boundaries by L1 and L2 speakers of English. *Proceedings of InterSpeech* 2014, 544–547.

Sugahara, Mariko (2016). Is Japanese listeners' perception of English stress influenced by the antepenultimate accent in Japanese? Comparison with English and Korean listeners. *Doshisha Studies in English* 96, 61–111.