

Toward processing of prosody in spontaneous Japanese

Shinobu Mizuguchi & Koichi Tateishi*

Abstract. This paper considers how prosody in spontaneous Japanese is processed. We have conducted Rapid Prosody Transcription (RPT) perception experiments on the *Corpus of Spontaneous Japanese* (CSJ) and investigated how boundaries and prominences are perceived. We recruited three groups of participants from different Japanese dialects and found that (i) F0 is not a strong prominence cue in Japanese, contra Japanese literature on focus prominence (Pierrehumbert & Beckman (P&B) 1988; Kori 1989; Ishihara 2016) and (ii) Japanese allows multi-headed and headless intonation phrases, and P&B's reset theory, i.e. focal prominence resets boundary phrases, faces empirical difficulties. We also found that (iii) both content words and function morphemes get highlighted in Japanese, and (iv) perception strategies vary cross-dialectally and listeners from different dialects perceive boundaries and prominences differently.

Keywords. prominence; prosodic boundary; Japanese; spontaneous speech; Rapid Prosody Transcription

1. Introduction. Prosody plays an important role in determining the meaning of the utterance. The studies on stress languages show that prominence is considered to be highly tied to the changes in F0 and that prominence perception is affected by the excursion size of the F0 peak (cf. Ladd et al. 1985 for English, Rietveld & Gussenhoven 1985 for Dutch, Baumann & Winter 2018 for German, Cole et al. 2019 for English, French, and Spanish). This paper considers Japanese. It is a pitch language, and its intonation arises compositionally from the lexical pitch accents (H*L, LH), the phrasal tones, and the boundary tones (L%, (L)HL%, H%) (cf. Venditti et al. 2008, Féry 2017). In the traditional analysis of Japanese, focal prominence is characterized by F0 boost and post-focal compression (cf. Kori 1989, Ishihara 2016, among others) and is considered to introduce a new prosodic phrase (cf. Pierrehumbert & Beckman (P&B) 1988). These studies are based on the production studies on the word or phrase levels at most. The processing of prosodic boundaries and prominences in spontaneous Japanese is understudied. Our research questions are (i) what are boundary and prominence cues in Japanese spontaneous speech? More concretely, is F0 a prominence cue in Japanese, as reported in the previous studies? (ii) Does prominence reset a phrase boundary, as the Japanese focus literature claims? We will conduct Rapid Prosody Transcription (RPT) experiments on spontaneous Japanese, showing that F0 is not a prominence cue so that P&B's reset theory faces empirical problems.

2. RPT perception experiment. We have conducted RPT perception experiments on Japanese spontaneous speech. We will present a general view of RPT first and then discuss our experiments below.

* We thank our experiment participants and the audience of our presentation at the 97th Annual Meeting of the LSA 2023 in Denver. Our deepest gratitude goes to Timothy Mahrt for his help and suggestions in the course of our experiments. This work is partially supported by a collaborative research project fund 2021 of the National Institute of the Japanese Language and Linguistics (NINJAL), given to the first author. All errors are our own. Authors: Shinobu Mizuguchi, Kobe University (mizuguti@kobe-u.ac.jp) & Koichi Tateishi, Kobe College (tateishi@mail.kobe-c.ac.jp).

2.1. RAPID PROSODY TRANSCRIPTION. RPT was developed by Cole and her colleagues and aims to broaden a research target to cover spontaneous speech. They consider that listeners differ in how they perceive prosody for the same utterance. They used prosodic transcription as a tool for prosody research. Listeners identify and mark prominences and boundaries based on an auditory impression of an utterance. Transcribers are given minimal instructions such as “mark prominent words that the speaker has highlighted for the listener, to make them stand out” and “mark boundaries between words that belong to different chunks that serve to group words in a way that helps listeners interpret the utterance” (cf. Cole et al. 2010). They are given no example transcriptions and no feedback on their transcription. The boundary (b-) score and the prominence (p-) score ranging from 0 to 1 are calculated (cf. Table 3). They indicate the proportion of participants who underscore the respective word. The higher values indicate strong perceptual salience of the prosodic element.

The main findings of the previous RPT studies are (i) boundaries are cued by higher syntactic categories, i.e. S & SBAR, Conjunction, and also by non-syntactic categories of Discourse Markers (DM) and Disfluencies (DISF) in English (cf. Cole et al. 2010), (ii) prominences are cued by acoustic features of F0, duration, intensity, pitch movement LH in German, and the final position of a boundary phrase in English (cf. Baumann & Winter 2018; Baumann & Schumacher 2020; Cole et al. 2019), and (iii) the cross-linguistic difference in prominence perception is statistically significant among English, French, and Spanish (cf. Cole et al. 2019).

2.2. OUR RPT EXPERIMENT.

2.2.1. METHODOLOGY. We conducted RPT experiments. We used 13 excerpts of *Corpus of Spontaneous Japanese* (CSJ) released by the National Institute for Japanese Language and Linguistics (NINJAL). CSJ is a corpus of monologues and dialogues by more than 1,400 Tokyo Japanese native speakers, with the total recorded time being about 660 hours. We used 6 lecture-type monologues and 7 pseudo-lecture-type monologues. The monologues are 16 to 41 seconds long for our experiments. Since Japanese is an agglutinative language, we segmented our data set on the morpheme level (cf. (1)); our materials contain 490 content words and 490 function morphemes.

We recruited three groups of Japanese listeners (35 Tokyo Japanese (TJ) listeners (mean age 24.8, SD=0.5), 27 Osaka Japanese (OJ) listeners (mean age 25.3, SD=3), and 11 Northern Kanto (NK) listeners (mean age 23.2, SD=4)).¹ The prosody varies among dialects in Japanese; on the word (ω) level, TJ and OJ have accented (H*L) and unaccented (LH) words, while NK has only unaccented words. On the phrase (φ) level, TJ and NK allow dephrasing, i.e., the process of deleting a φ -phrase when syntactic words form a prosodic phrase, while OJ does not (cf. Igarashi 2014). On the intonation phrase (i) level, TJ is downward (L%), NK is upward (H%), and OJ is both downward and upward. We aim to examine whether listeners from different dialects have the same perception tendencies or not.

We conducted online perception experiments via Yahoo! Crowd-Sourcing Service. After the exercise session, our participants listened to each material twice via PC, while marking boundaries and prominences on the text by clicking a mouse. Their responses were saved on the

¹ We recruited 50 participants for each group. Due to the mismatch between a self-claimed native dialect and the target dialect, we regret that we had to dismiss the mismatched data.

computer via LMEDS, developed by Mahrt 2016. It took about 30 minutes for our participants to complete the task. Our participants were paid in Yahoo! points. Our experiments were approved by the Ethical Committee of NINJAL.

For the syntactic analysis of boundary-marking, syntactic categories of S, S-bar, Conjunction, XP, Particle, and non-syntactic categories of Disfluency and Discourse Marker are assigned at the left edge ('[X'] and the right edge (']Y') of each morpheme following Cole et al. 2010, as in (1).

- (1) [S [Adv mazu]Adv [NP [N watashi]-[Part no]]NP
first I-GEN
[NP [v sukina [N mono]]]NP [[N inu]-[v desu]]VP]S
like thing dog-COPULA
'First, what I like is a dog.'

(N.B. In the experiment, the text is written in Japanese *Hiragana*, i.e. Japanese cursive character, without punctuation.)

We predicted that due to the downtrend in production and the F0 normalization in perception (cf. Pierrehumbert 1979), Tokyo Japanese does not employ F0 as a prominence cue. Another prediction is that in case Maximal F0 is not a crucial prominence cue, prominence is not expected to appear at the initial position of a boundary phrase. It, therefore, does not reset a prosodic phrase.

2.2.2. RESULTS. We will list four main results of our experiments.

- (i) The overall inter-listener agreements are κ 0.638 on b(oundary)-score and κ 0.359 on p(rominence)-score on Fleiss's kappa. Like stressed languages, boundaries are easier to share than prominences in the perception of Japanese.
- (ii) The correlation between the b-score and p-score is weak ($r=0.12$ on Pearson's Correlation).
- (iii) The numbers of markings vary cross-dialectally, as given in Table 1, which gives the numbers of b-/p-markings larger than 0.2.²
- (iv) Boundaries are cued by syntactic categories of S, ConjP, NP, and VP and non-syntactic categories of TopP and DM (cf. Figure 1).

Dialect	Boundary-marking (b-score>0.2)	Prominence-marking (p-score>0.2)			
		IP-initial	IP-mid	IP-final	total
Tokyo	167	74	157(31)	53(45)	284(76)
Osaka	181	63	95(11)	26(23)	182(34)
Northern Kanto	163	52	55(6)	22(15)	129(21)

Table 1: Number of boundary and prominence markings

(N.B. The number in the parenthesis is the number of p-markings on a function morpheme)

One-way ANOVA between the b-score and three dialects was $F(2, 2940)=0.48$, $p=0.01$, and further Bonferroni is significant only between TJ and OJ ($p=0.01$). The story is different in

² We decided to choose b/p-score >0.2 because 0.2 is a little bit above the average score.

prominence perception. One-way ANOVA between p-score and three dialects is $F(2,2940)=1.55$, $p<0.001$, and further Bonferroni is significant between TJ and OJ ($p=0.001$), TJ and OJ as well as OJ and NK ($p<0.001$, for both).

Table 2 shows the results of regression analyses between b-/p-scores and acoustic features, and we can see that acoustic features of range F0 or duration do not predict b-scores. P-scores are more or less predicted by range F0, but not by Max(imal) F0, duration, or intensity.

Dialect	Adjusted r^2 (b-score X)		Adjusted r^2 (p-score X)			
	bXrange F0	bXduration	pXMaxF0	pXrange F0	pXduration	pXintensity
Tokyo	0.009**	0.0036***	0.159***	0.237***	0.147***	0.051***
Osaka	0.015***	0.056***	0.172***	0.252***	0.146***	0.051***
Northern Kanto	0.011***	0.041***	0.121***	0.199***	0.097***	0.04***

Table 2: Results of regression analyses between b-/p-scores X acoustic features, where *** $p<0.001$, ** $p<0.01$

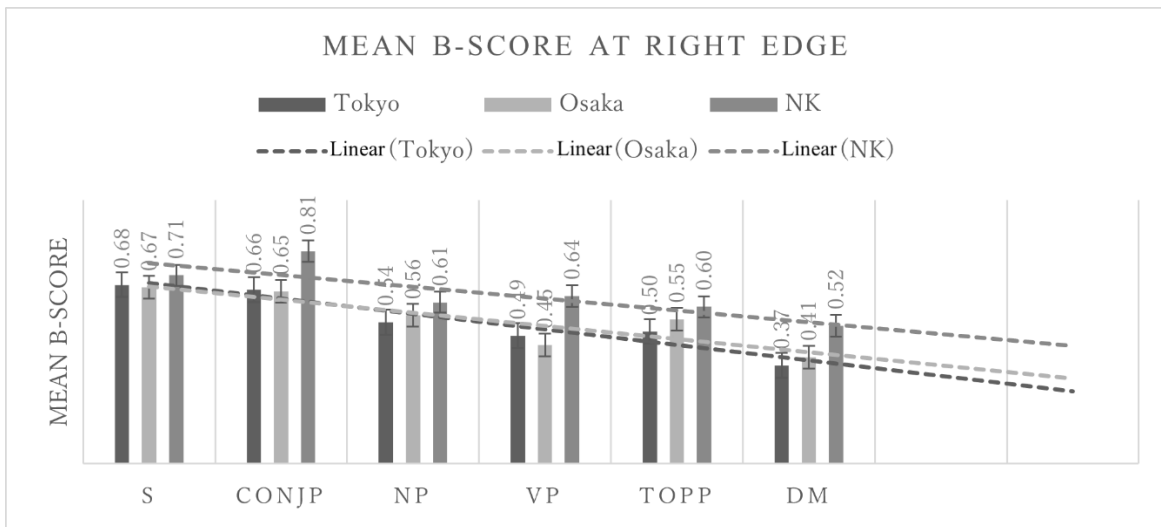


Figure 1. Mean b-scores per category at the right edge, where S=Sentence, CONJP=Conjunction Phrase, NP=Noun Phrase, VP=Verb Phrase, TOPP=Topic Phrase, DM=Discourse Marker

3. Discussion.

3.1. ANSWERS TO RQS. We had two research questions when we started this study; RQ 1 was ‘What are boundary and prominence cues in Japanese spontaneous speech? More concretely, is F0 a prominence cue in Japanese, as reported in the previous studies?’ RQ2 was ‘Does prominence reset a phrase boundary, as the Japanese focus literature suggested?’ Our answer to RQ 1 is that boundaries are cued not acoustically (cf. Table 2) but are cued by syntactic and non-syntactic categories (cf. Figure 1). Cole et al. (2010) claim that English uses higher syntactic categories of S & SBAR and Conjunction as boundary cues (cf. Figure 2). Comparing Figure 1 with Figure 2, we see that Japanese uses lower syntactic categories of NP and VP as boundary cues in addition to higher syntactic categories. We also observe that non-syntactic categories of

Disfluencies (DISF) and Discourse Marker (DM) are less effective boundary cues in Japanese than in English.

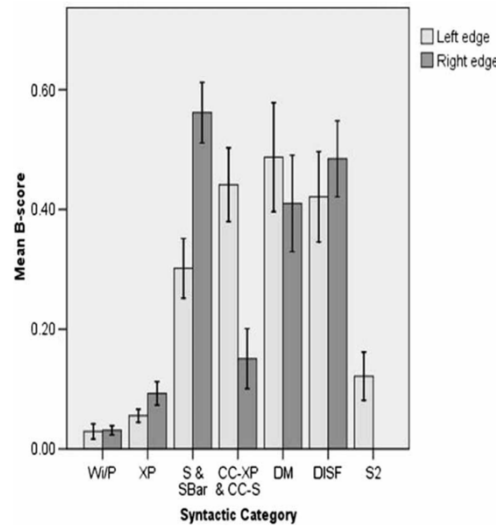


Figure 2. Mean b-score per category at the left and right edge (Cole et al. 2010:1161) (N.B. XP covers NP, VP, ADJP, ADVP, and PP, and S2 codes subordinate clause)

As for prominence, Table 2 shows that prominences are not predicted by Max F0, duration, or intensity in Japanese. Note here that the perception of prosody in pitch language requires some caution. In pitch language, pitches gradually lower as the utterance continues, so listeners have to normalize F0 when they perceive it. Pierrehumbert (1979) claims, based on perception experiments, that when two stressed syllables sound equal in pitch, the second one is lower in F0 due to declination. Figure 3 (right), for example, illustrates two focal prominences on *chocho* ‘major key’ and *tancho* ‘minor key’ marked by the participants of our RPT experiments. We can see that the F0 peak of the second prominent word *tancho* is lower than that of the first prominent word *chocho*, while they sound equally prominent. Pierrehumbert (1979) claims that F0 is normalized when we perceive it. Table 2 shows that range F0 is a better predictor of prominence than Max F0 in Japanese. Compare Figure 3 (left) with Figure 3 (right), for example; in the former, no prominence is marked, while in the latter, with expanded pitch ranges, three prominences are marked by our participants.

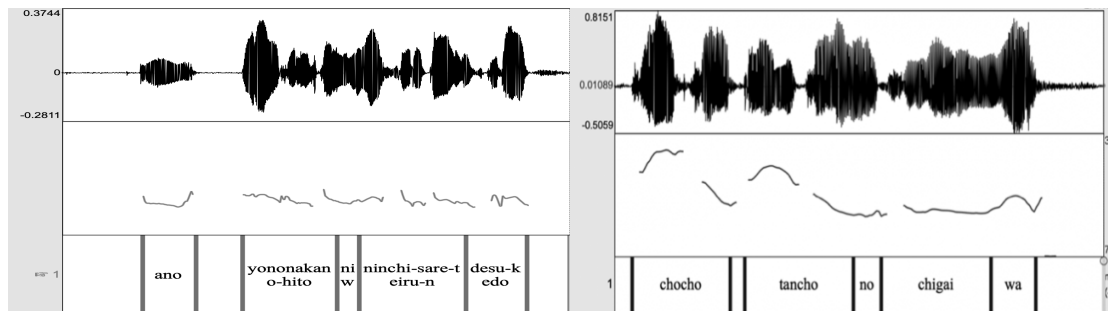


Figure 3. LEFT: Pitch movement without prominence markings in an IP (extracted from CSJ file #A01f0055); RIGHT: Pitch movement with three prominences marked in an IP (extracted from CSJ file #S00f0082)

Thus, our answer to RQ 1, ‘Is F0 a prominence cue in Japanese?’, is ‘no’.

RQ 2 ‘Does prominence reset a phrase boundary, as claimed in the Japanese focus literature?’ is hard to answer unless we know what a prosodic phrase is. Japanese literature has a long history of studies on lexical accents and prosodic phrasing (cf. McCawley 1968; Poser 1984; P&B 1988; Selkirk & Tateishi (S&T) 1988; Kubozono 1988, 2007; Ito & Mester 2012, among many others), but the prosody above the ϕ -phrase level (cf. (2)) is understudied.

(2) Prosodic Hierarchy (Féry 2017:36)

v	utterance	(corresponds roughly to a paragraph or more)
ι -phrase	intonation phrase	(corresponds roughly to a clause)
ϕ -phrase	prosodic phrase	(corresponds roughly to a syntactic phrase)
ω -word	prosodic word	(corresponds roughly to a grammatical word)
F	Foot	(metrical unit)
σ	syllable	(strings of segments)
μ	Mora	(unit of syllable weight)

Japanese prosody is complex and is formed compositionally by the lexical pitch accents, phrasal tones, and boundary tones (Féry 2017: 248). Unlike stressed languages like English, Japanese assigns its lexical accents by bi-tonal H*L. Japanese accent is lexically determined, and its placement is often unpredictable. Tokyo Japanese is a moraic language in which accent is realized on a mora. Figure 4 shows the pitch contours of the declarative sentence *Naoya-mo oyoida* ‘Naoya also swam’ from Venditti et al. (2008:488). It is a good example to show that the same utterance is assigned a variety of pitch movements in Japanese.

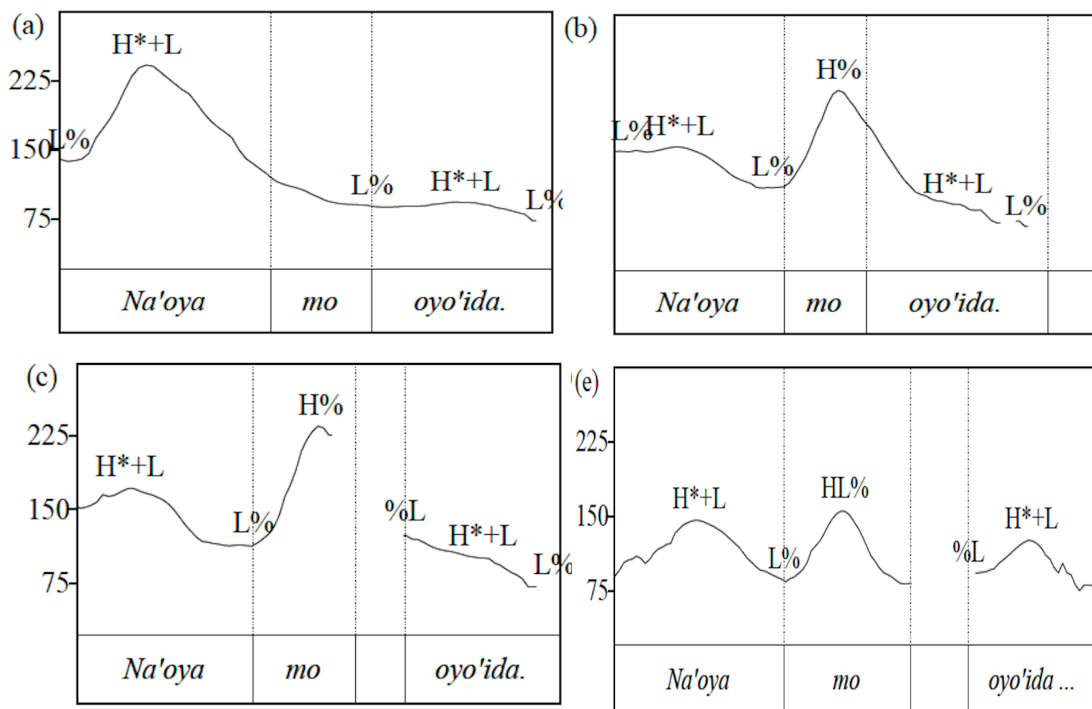


Figure 4. Pitch movements of ‘*Naoya-mo oyoida*’ ‘Naoya also swam’ (Venditti et al. 2008: 488)

McCawley (1968) posits a level of the Minor Phrase (MinP), the Major Phrase (MP), and the utterance for Japanese prosody. McCawley (1968) and Poser (1984) claim that the MP is the domain of ‘catathesis’, i.e., a lowering of F₀-contour induced by the H*L pitch accent. P&B 1988 assume the Accentual Phrase (AP), the Intermediate Phrase, and the utterance. P&B (1988:16) define the AP as a phrase bearing at most one pitch accent whose periphery is marked with an %H at the beginning and an L% at the end. (X)J_ToBI (cf. Venditti 1995; Maekawa 2011, among others) posits the levels of the AP, the Intonational Phrase (IP), and the utterance. Venditti et al. (2008) claim that L%, H%, LH%, and LHL% are phrase-final tones and that they are assigned at the AP in their framework. Intonational tones are also confusing in Japanese literature; Poser (1984) claims that the intonational H% is a final tone inserted at the MP, but P&B 1988 assume that the intonational H% is attached at the utterance level. It is not clear yet at what phrase these tones are aligned. I&M 2012 claim that there is no need to distinguish a MinP from an MP and assume ϕ - and ι -phrases that are recursive. Their framework does not assume a phrase that is unique to Japanese, like AP or MP but fits in the general prosodic hierarchy (2) in the prosodic literature. It is generally accepted in Japanese literature that lexical tones play some part in a larger organization of phrases or intonation patterns, but no standard view for phrasal patterns above the word level is established (cf. P&B 1988:9). To avoid irrelevant confusion in terminology, we will assume the universal prosodic hierarchy (2) in our study (for further discussion, see Mizuguchi & Tateishi (M&T) in press).

Back to our RQ2, P&M (1988) claim that focus appears at the leftmost position of an ι -phrase (i.e. the intermediate phrase in their term) and resets a prosodic phrase. In their framework, a boundary is inserted after the focus. Their theory predicts every ι -phrase has a focus at its left-most position. But Shinya 1999 and Kubozono 2007 empirically refuted P&B’s reset theory on the sentence level. Our RPT experiments found that the correlation between boundary and prominence perception is very weak ($r=0.12$ on Pearson’s Correlation), and also show that focal prominence appears in phrase-initial, phrase-mid, and phrase-final positions in spontaneous speech (cf. Table 1). We found that some phrases do not mark focal prominence at all (cf. Figure 3 (left)). We consider these as serious problems to the reset theory, and our answer to RQ2 is ‘no’.

3.2. NEW FINDINGS. In the course of our RPT experiments, we have two new findings. One is that not only content words (73% - 84%) but also function morphemes (16% - 27%) are perceived as prominent in Japanese spontaneous speech (cf. Table 1). The RPT studies on stress languages report that prominence is perceived on content words only (cf. Baumann & Winter 2018 for German; Cole et al. 2019 for English, French, and Spanish; Bishop et al. 2020 for English). Japanese is an agglutinative language, and Japanese literature has long observed that function morphemes are also subject to highlighting (cf. Oishi 1959; Kawakami 1963; Kuroda 1965; Kuno 1973; Taniguchi & Maruyama 2001; Hara 2006; Tomioka 2009, among others). (3) is an example from Tomioka (2009), where prominence is marked either on subject *KEN* or contrastive particle *WA* in (3A).

- (3) Q: Dare-ga ukat-ta-no?
 Who-GA pass-PAST-Q
 ‘Who passed?’
 A: KEN-wa/Ken-WA ukat-ta.
 Ken-TOP pass-PAST
 ‘(At least) Ken passed.’ (Tomioka 2009: 119)

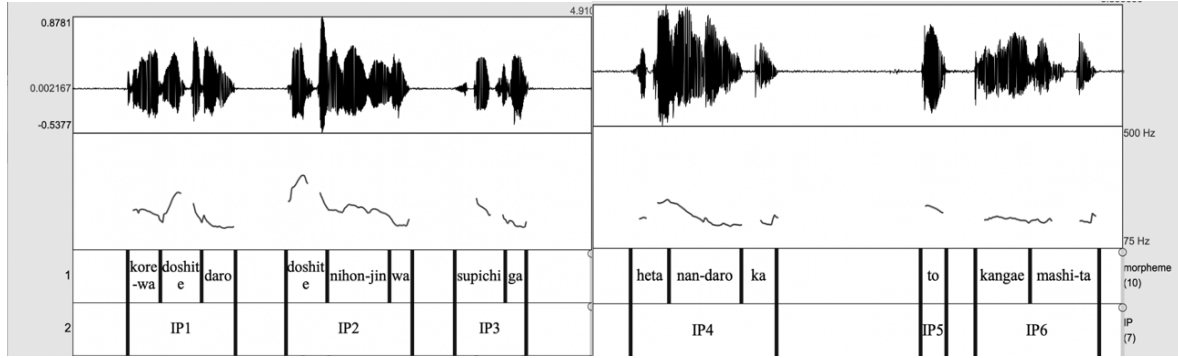


Figure 6. Pitch movements across multiple IPs, extracted from CSJ File #S00f0208 (Text: kore(this)-wa(TOP) doshite(why)-daro(AUX), doshite(why) nihon-jin(Japanese)-wa(TOP) supichi(speech)-ga(NOM), heta(poor)-nan(COPULA)-daro(AUX)-ka(Q), to(COMP), kangae(think)-mashi(COPULA)-ta(PAST))

token	morpheme	b(oundary)-score			p(rominence)-score		
		Tokyo	Osaka	NK	Tokyo	Osaka	NK
1	kore	0	0	0	0.14	0.11	0
2	wa	0	0	0	0.09	0.11	0
3	doshite	0	0	0	0.17	0.11	0.36
4	daro	0.71	0.78	0.82	0.14	0.15	0.09
5	doshite	0.06	0.07	0.09	0.71	0.48	0.55
6	nihon	0	0	0	0.29	0.19	0
7	jin	0	0	0	0.26	0.19	0
8	wa	0.57	0.7	0.64	0.23	0.15	0
9	supichi	0.03	0.04	0	0.31	0.37	0.27
10	ga	0.51	0.67	0.73	0.17	0.22	0
11	heta	0	0.04	0	0.4	0.44	0.27
12	nan	0	0	0	0.23	0.19	0.18
13	daro	0.03	0	0	0.26	0.19	0.18
14	ka	0.66	0.82	0.82	0.23	0.19	0.09
15	to	0.17	0.37	0.18	0.17	0.26	0.36
16	kangae	0	0.04	0	0.06	0.07	0
17	mashi	0	0	0	0	0	0
18	ta	0.17	0.11	0.36	0.06	0	0

Table 3. Coding of boundaries and prominences of Figure 6

Besides the difference in boundary marking, Table 1 also shows that prominence marking differs more than boundary marking among the listeners of the three dialects. TJ listeners perceive more prominences than OJ and NK listeners. It is tough to answer why, but Table 3 may give us a clue, as it shows that TJ listeners perceive the whole syntactic category highlighted, while OJ and NK listeners perceive that only the head is prominent given the same highlighting maneuver. If our observation of the perception bias of TJ listeners is correct, this will lead us to reconsider the framework of Alternative Semantics we discussed above in (4), where only the head of FoC is aligned with a prosodic prominence. We need lots for further consideration.

4. Conclusion and implication. Japanese, a pitch language, employs a different perception strategy from stress languages. We have observed that there are two ways to mark prominence in Japanese: local expansion of pitch range and boundary tone. The former is realized locally to content words, and the latter is assigned to function morphemes at the boundary mid and final positions. Cross-linguistic perception variations have been observed in the literature (cf. Cole et al. 2019, among others). Our study suggests a cross-dialectal perception bias. Japanese literature has long reported a cross-dialectal difference in production. If the difference in production will affect perception, it will open a new perspective on human cognition. Future study is expected.

References

- Baumann, Stephan & Bodo Winter. 2018. What makes a word prominent?: Predicting untrained German listeners' perceptual judgment. *Journal of Phonetics* 70. 20–38. <https://doi.org/10.1016/j.wocn.2018.05.004>.
- Cole, Jennifer, Yun-S. Mo, & S-D Beak. 2010. The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes* 25. 1141–1177. <https://doi.org/10.1515/labphone.2010.022>.
- Cole, Jennifer, Hose I. Hualde, Caroline L. Smith, Christopher Eager, Timothy Mahrt, & Ricaodo Napoleão de Souza. 2019. Sound, structure, and meaning: The bases of prominence ratings in English, French, and Spanish. *Journal of Phonetics* 75. 113–147. <https://doi.org/10.1016/j.wocn.2019.05.002>.
- Elordieta, Gorka, & Elizabeth Selkirk. 2022. Unaccentedness and the formation of prosodic structure in Lekeitio Basque. In Kubozono et al. 2022, 374–419.
- Féry, Caroline. 2017. *Intonation and prosodic structure*. Cambridge: Cambridge University Press.
- Fukui, Naoki. 1995. *Theory of projection in syntax*. Tokyo: Kurosio Publishers.
- Igarashi, Yosuke. 2014. Typology of intonational phrasing in Japanese dialects. In Jun, S-A. (ed.) *Prosodic Typology II*. Oxford: Oxford University Press. 464–491.
- Ishihara, Shin'ichiro. 2016. Japanese downstep revisited. *Natural Language & Linguistic Theory* 34. 1389–1443. <https://doi.org/10.1007/s11049-015-9322-8>.
- Ito, Junko & Armin Mester. 2012. Recursive prosodic phrasing in Japanese. In Borowsky, Toni, Shigeto Kawahara, Takahiro Shinya, & Mariko Sugahara (eds.). *Prosody matters*, 280–303. London: Equinox.
- Kori, Shiro. 1989. Kyocho-to intoneshon [Emphasis and intonation]. In Miyoko Sugito (ed.), *Nihongo-no Onsei Onin* [Japanese Phonetics and Phonology], 316–342. Tokyo: Meiji Shoin.
- Kratzer, Angelika & Elizabeth Selkirk. 2020. Deconstructing information structure. *Glossa* 5(1): 113. <https://doi.org/10.5334/gjgl.968>.

- Kubozono, Haruo. 2007. Focus and intonation in Japanese: Does focus trigger pitch reset? *Workshop on Prosody Syntax, and Information Structure (WPSI) 2*. 1–27.
- Kubozono, Haruo, Junko Ito & Armin Mester (eds.). 2022. *Prosody and prosodic interfaces*. Oxford: Oxford University Press.
- Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.
- Lee, Seunghum J. & Elizabeth Selkirk. 2022. Xitsonga tone: The syntax-phonology interface. In Kubozono et al. 2022, 337–73.
- Mahrt, Tim. 2016. LMEDS: Language makeup and experimental design software. <https://github.com/timmahrt/LMEDS>.
- McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.
- Mizuguchi, Shinobu & Koichi Tateishi. In press. *Prominence in a pitch language: The production and perception of Japanese*. Lanham, MD: Lexington Books.
- National Institute of the Japanese Language and Linguistics (ed.). 2004. Corpus of Spontaneous Japanese. <https://clrd.ninjal.ac.jp/csj/en/index.html>.
- Pierrehumbert, Janet. 1979. The perception of fundamental frequency declination. *Journal of Acoustic Society of America* 66(2). 363–369. <https://doi.org/10.1121/1.383670>.
- Pierrehumbert, Janet & Mary Beckman. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.
- Poser, Bill. 1984. *The phonetics and phonology of tone and intonation in Japanese*. Cambridge, MA: MIT dissertation.
- Rooth, Mats. 1986. *Association with focus*. Amherst, MA University of Massachusetts dissertation.
- Rooth, Mats. 2016. Alternative semantics. In Caroline Féry & Shin'ichiro Ishihara. (eds.), *The Oxford handbook of information structure*, 260–288. Oxford: Oxford University Press.
- Selkirk, Elizabeth & Koichi Tateishi. 1988. Constraints on minor phrase formation in Japanese. *Chicago Linguistics Society (CLS)* 24. 316–336.
- Shinya, Takahiro. 1999. Eigo to Nihongo ni okeru fokasu ni yoru daunsuteppu no soshi to cho'on-undo no chogo [The blocking of downstep by focus and articulatory overlap in English and Japanese]. *Proceedings of Sophia University Linguistic Society* 14. 35–51.
- Tomioka, Satoshi. 2009. Contrastive topics operate on speech acts. In Malte Zimmermann & Caroline Féry (eds.). *Information structure: Theoretical, typological, and experimental perspectives*, 115–138 Oxford: Oxford University Press.
- Venditti, Jennifer, Kikuo Maekawa & Mary Beckman. 2008. Prominence marking in the Japanese intonation system. In Shigeru Miyagawa & Mamoru Saito (eds.). *The Oxford handbook of Japanese linguistics*, 456–512. Oxford: Oxford University Press.