Assessing tonal specifications through simulation and classification: The case of post-wh accent in Japanese

BACKGROUND: Given a phonetic trajectory, be it articulatory or acoustic, deciphering whether a segment has a phonetic target in that trajectory, or is entirely targetless, is not an easy task (e.g. [1, 2]). [3] developed a general computational toolkit to address this issue, which allows us to assess the presence of a phonetic target on a token-by-token basis. [3] analyzed tongue dorsum trajectories, addressing whether devoiced [u] in Japanese retains its articulatory target; they found that some tokens retain a full target, whereas some tokens are entirely targetless. The current paper applies this computational toolkit to intonational patterns.

The empirical focus of the current study is realization of lexical LHL tonal accent following wh-elements in Japanese. [4] argues that lexical accent after wh-elements in Japanese is "eradicated," and that this domain of eradication continues up to the question particle that licenses that wh-element. This claim led [5] to propose a general theory of wh-movement in natural languages. Japanese allows wh-elements to stay in-situ because wh-elements and their licenser can be prosodically grouped together; since English does not have this prosodic means, it has to overtly move its wh-elements. Other studies however cast doubt on the claim that lexical accent after wh-elements is entirely deleted (e.g. [6, 7]). [6] shows that degrees of rise of lexical accent are in fact *reduced* after wh-elements, but the averaged F0 contours seem to show that there are some visible rises on the lexical items at issue.

METHOD: The current study thus reanalyzes a subset of the data obtained by [6]. We compare two sentence structures in (1) and (2).

(1) Control sentences: $Word_1$ $Word_{2[-wh]}$ $Word_3$ $Word_4$ Verb
(2) Test sentences: $Word_1$ $Word_{2[+wh]}$ $Word_3$ $Word_4$ Verb

The sentences in (1) serve as the control sentences in which the lexical accent of both $Word_3$ and $Word_4$ is realized (i.e. full target). What we are interested in is whether $Word_3$ and $Word_4$ in (2) maintain some traces of lexical accent or whether the accent is completely eradicated. There were six types of sentences for both (1) and (2); nine native speakers of Tokyo Japanese repeated those sentences, together with other sentences, twice. The intonational contours of $Word_3$ and $Word_4$, delimited by their L tones, were extracted using YAAPT. These contours were decomposed into a set of DCT coefficients. In addition to the control sentences, we simulated linear trajectories between the two L tones, and injected the same degree of variability as those observed in the sentences in (2)—these contours simulate how targetless intonational contours (i.e., tone eradication) would be realized given naturalistic variability. Finally, a Bayesian classifier was trained, and for each tonal contour for the sentences in (2), it assigned a posterior probability of belonging to the linear interpolation category.

RESULTS: Figure 1 shows the posterior probability of linear interpolation for $Word_3$ and $Word_4$ For $Word_3$, many speakers (Speakers 1, 2, 4, 5, 7, 8, 9) show at least some tokens that have a high posterior probability of linear interpolation. These tokens support the view expressed by [4]. However, a Speakers 6, 7, 8, and 9, show a large number of tokens that are better classified as belonging to the full target category; these tokens show no trace of reduction. We finally observe those tokens whose posterior probabilities are in the middle range (Speakers 2, 4, 5, 7); these tokens are phonetically reduced. For most speakers (all except 3), within-speaker variability is evident. For $Word_4$, most speakers (all but Speakers 1 and 6) show tokens of high linear interpolation probability. Speakers 1, 3, 5, 6, and 8, also produced tokens with full tonal targets and there are many tokens as well that are phonetically reduced. Again, just like $Word_3$, we observe both inter- and intra- speaker variability.

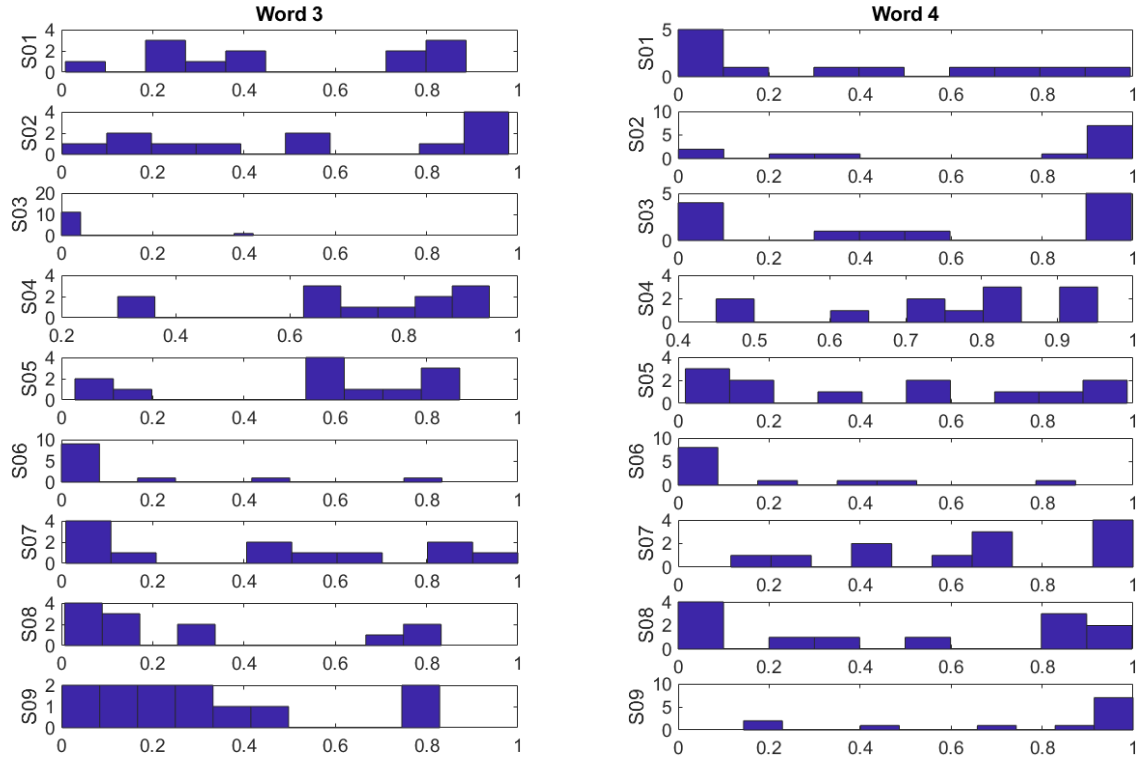

Figure 1: Posterior probabilities of linear interpolation. Left = Word$_3$; Right = Word$_4$.

DISCUSSION: Our analysis shows that some speakers show some tokens that are best characterized as eradicated; at the same time, however, no speakers consistently show eradication. These results imply that [4]'s observation was correct at some level of analyses, but the current results pose an interesting challenge to [5]. If eradication is what allows Japanese wh-elements to stay in-situ, how come those tokens without eradication show no wh-movement? How come Speaker 6, who almost always showed high probability of full target, does not move wh-elements?

Next, we discuss some advantages of our analytical toolkit. First the toolkit allows us to assess the presence of intonational targets on a token-by-token basis. Most previous studies analyze averaged contours, but analyzing only averaged contours can be misleading. Take the case of Word$_4$ for Speakers 3 and 4, for example. In our analysis, Speaker 4 shows reduction for all tokens; Speaker 3 on the other hand shows a bimodal distribution of full targets and eradication. If we were only looking at averages, we would have erroneously concluded that both speakers show reduction. This highlights the importance of analyzing each token separately. Second, [7] addressed the question of whether "deaccented" phrases and unaccented phrases are different or not, by fitting linear regression lines (see also [2]), and found that the regression lines are different between deaccented phrases and unaccented phrases. One distinct advantage of our approach is that it does not have to assume linearity, as the first step of our analysis decomposes trajectories into a sum of cosine waves. Token-by-token analysis offers great promise for the study of intonational variation.

REFERENCES: [1] Cohn, A. 1993. Nasalisation in English: phonology or phonetics. *Phonology* 10. 43–81. [2] Pierrehumbert, J. & M. Beckman. 1988. Japanese tone structure. MIT Press. [3] Shaw, J. & S. Kawahara. 2018. Assessing surface phonological specification through simulation and classification of phonetic trajectories. *Phonology* 35. 481–522. [4] Deguchi, M. & Kitagawa, Y. 2002. Prosody and wh-question. *Proceedings of NELS* 32: 73-92. [5] Richard, R. 2010. Uttering Trees. MIT Press. [6] Ishihara, S. 2010. Focus prosody in Tokyo Japanese wh-questions with lexically unaccented wh-phrases. *Proceedings of ICPhS* 946-949. [7] Maekawa, K. 1994. Is there 'dephrasing' of the accentual phrase in Japanese? *Ohio State University Working Papers in Linguistics* 44. 146-156.