Do Japanese speakers always prosodically group wh-elements and their licenser? Implications for Richards' (2010) theory of wh-movement

BACKGROUND: Languages differ with respect to whether wh-phrases move overtly or not: English wh-phrases move overtly; Japanese wh-phrases stay in situ. This cross-linguistic syntactic variation is argued by Richards (2010) to follow from a language universal—all languages attempt to create a prosodic structure for wh questions in which the wh-phrase and corresponding complementizer are separated by as few prosodic boundaries as possible. This is accomplished in in English via wh-movement. Inspired by a body of work on Japanese intonational patterns, Richards proposes that Japanese has a prosodic means to group the wh-phrase and its complementizer, and hence doesn't need to resort to overt wh-movement. In Japanese, a Minor Phrase contains at most one accented lexical item, and is signaled by a phrase-initial %LH rise and a H^*+L accentual fall (Pierehumbert & Beckman 1988). Deguchi & Kitagawa (2002) argue that these tonal events associated with Minor Phrases are eradicated after wh-elements up to the complementizer that licenses the wh-elements, effectively grouping the wh-phrase and complementizer within a single Minor Phrase. Richards' theory of wh-movement assumes that Japanese resorts to this sort of prosodic means in one way or another (see below for different alternatives); we feel that it is important to test this assumption with actual utterances.

THE CURRENT STUDY: This paper reexamines wh-phrase conditioned tone eradication in Japanese. We adopt a token-by-token analysis of pitch contours, using the computational toolkit recently developed by Shaw & Kawahara (2018). Our analysis shows that almost all speakers produce some sentences that behave as predicted by Richards' theory, as well as some that don't. We conclude that the prosodic properties of Japanese that allow their wh-elements to stay in situ must be more abstract than currently assumed.

METHOD: The current study reanalyzes a subset of the data obtained by Ishihara (2011). The corpus features carefully controlled pairs consisting of a declarative sentence and wh-question counterpart. All sentences consist of five accented words . For Wh-questions, the second word is the wh-phrase. Schemata of the item pairs are given below:

(1) Declarative sentence (control): Word_1 $\text{Word}_{2[\text{-wh}]}$ Word_3 Word_4 Verb
(2) Wh-question sentence (test): Word_1 $\text{Word}_{2[\text{+wh}]}$ Word_3 Word_4 Verb

The sentences in (1) serve as the control sentences. Both Word_3 and Word_4 are Minor Phrases; as such, the phrasal tones (%LH) and lexical accent (H^*+L) of both Word_3 and Word_4 are realized (i.e. full target). We are interested in whether Word_3 and Word_4 in (2) also retain these %LH^*+L tones, or whether these tones are completely eradicated, as implied by Deguchi & Kitagawa (2002). 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 each. The intonational contours of Word_3 and Word_4, delimited by %L and +L, were extracted using YAAPT (Zahorian & Hu 2008). 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 eradication category.

RESULTS: Figure 1 shows the posterior probability of eradication 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 eradication. These tokens support the view expressed by Deguchi & Kitagawa (2002), and instantiate tonal patterns assumed by Richards (2010). However, 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 also evident. For $Word_4$, most speakers (all but Speakers 1 and 6) show tokens of high eradication 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.

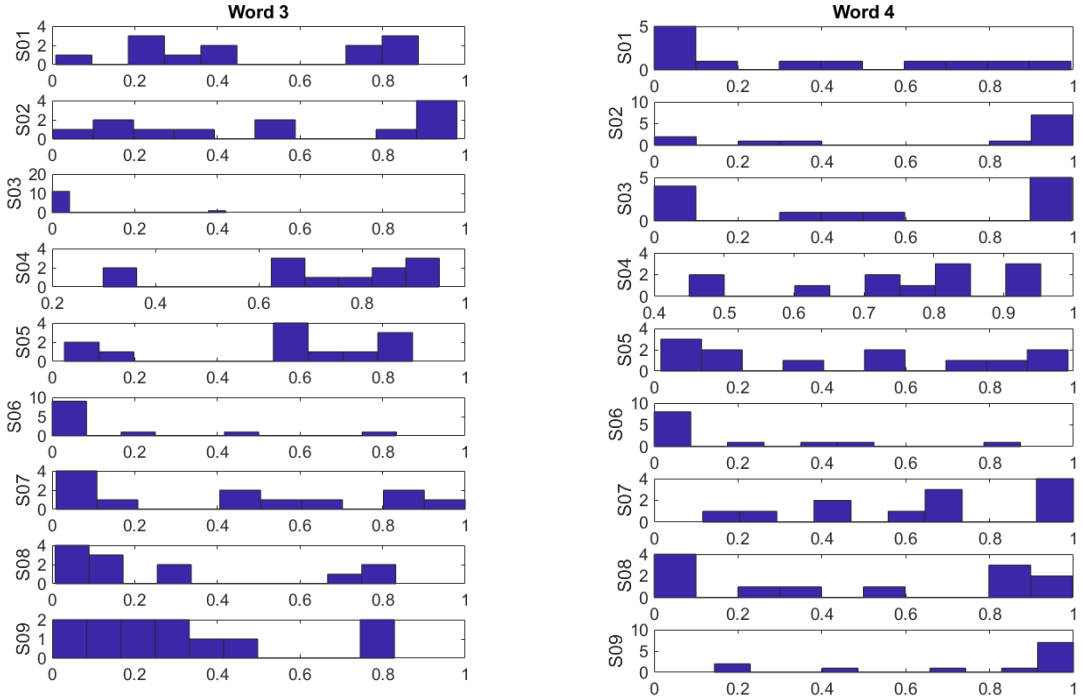

Figure 1: Posterior probabilities of tone eradication. Left = $Word_3$; Right = $Word_4$.

DISCUSSION: Our analysis shows that some speakers do show some tokens in which %LH*+L tones are eradicated (i.e. linear interpolation between %L and L); at the same time, however, no speakers consistently show eradication. These results pose an interesting challenge to Richards' (2010) theory. 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?

One may argue that the licensing domain is higher in the prosodic hierarchy—e.g., a Major Phrase (or a higher recursive Minor Phrase, as actually assumed by Richards), which can be signaled by reduction (a.k.a. downstep) of lexical H*+L accents (Pirrehumbert & Beckman 1988). However, our results show that some tokens do not even show reduction, and hence there is no prosodic evidence for grouping the wh-elements and their licensors. Overall, our results provide robust quantitative support from a reasonable sample of naturally produced utterances that tone eradication, as assumed by Richards (2010), does indeed happen in Japanese. However, the process is variable, which implies that what is crucial is not tone eradication per se but rather the *availability* of the prosodic grouping of wh-phrases and their licensors; i.e., it is not the case that a prosodic grouping pattern is checked for each syntactic derivation but instead that it is a more abstract prosodic feature of Japanese that allows wh-phrases to stay in situ.