What does pitch range tell us about speech and speakers?
– A sociophonetic study of pitch range variation in spoken Japanese –

Céleste Guillemot¹, Shinichiro Sano²
¹ICU, ²Keio University

Synopsis: The phonological organization of a language restricts the way in which pitch change can be realized lexically and/or intonationally (Ward and Hirschberg 1985, Pierrehumbert and Hirschberg 1990, Ladd 1996). Japanese is a language with lexical pitch accent, but no intonational pitch accent reported. In this study, based on the analysis of the Corpus of Spontaneous Japanese, we investigated the domain of pitch range in spoken Japanese. Specifically, we examined the effects of extra-linguistic factors on f0 range variations as well as how pitch range affects listeners’ perception. Findings suggest that language-specific restrictions on pitch range can be overridden by the influence of extra-linguistic factors such as gender and register, and influence the listeners’ perception of the speech characteristics.

Background: Previous studies have shown that a phonological grammar that demands a particular tonal classification affects the variability in pitch range; namely, in tonal languages (e.g., Chinese) smaller variations are observed compared to non-tonal languages (e.g., English) (Grieser and Kuhl 1988). The same observation was also made for Japanese compared to Western languages (Fernald et al. 1989). Indeed, contrary to English where accent placement is determined by an interplay of various linguistic and extra-linguistic factors, pitch accent in Japanese is a property of a given word. Lexical pitch accent is involved in a variety of minimal pairs, suggesting its crucial role for lexical distinction in Japanese (e.g., ame HL ‘rain’ vs. ame LH ‘candy’). A logical consequence of this should be a restriction of the domain of possible variations in pitch (Venditti 2005), as it might interfere with the meaning, resulting in a neutralization of contrastive items.

Previous research has shown that gender does not play a crucial role in predicting pitch range variations; that is, male and female speakers exhibit similar patterns (Henton 1989). However, it appears that speakers adjust pitch range to different degrees of formality (e.g., more informal => wider pitch range in Switzerland French; Sertling-Miller 2009) or to speech register (e.g., read speech > spontaneous speech in English; Gut 2009). Moreover, studies focusing on infant-directed speech in Japanese shed light on the fact that even in languages where f0 is used as a cue for lexical distinction, in specific pragmatic situations speakers can make use of pitch range modulation. Grieser and Kuhl (1988) for Mandarin Chinese and Igarashi et al. (2013) for Japanese find significant pitch expansion for interactions with infants when compared to those with adults.

Previous studies have investigated variations in pitch mainly from an experimental and qualitative point of view. Furthermore, they deal with pitch range with a focus on specific pragmatic contexts (e.g., infant-directed speech). There is a need for more studies on the interaction between pitch and intonation in a spontaneous (adult) speech context. Thus, this study explores the variations of pitch range in spontaneous speech using corpus data. Specifically, we examined how extra-linguistic factors affect its variations, and identified connections between the patterns of pitch range variation and information about speech and speakers.

Method: Data was retrieved from the Corpus of Spontaneous Japanese (Maekawa 2003, henceforth CSJ) a large-scale database of spontaneous speech. We extracted data from the CSJ core, which comprises 201 speech samples (about 162,000 seconds, 111 male, 90 female speakers, age range: 25-69 y.o.) from four different registers: Academic Presentation Speech (A, N=70), Simulated Public Speaking (S, N=107), Dialogue (D, N=18), and Reading (R, N=6). The CSJ also comes with TextGrid files and a variety of annotations regarding speaker attributes such as gender, age, and other indexes regarding perceived characteristics of speech. We analyzed a total of 60,045 tokens. F0 measurements were conducted using a Praat (Boersma and Weenink 2018) script based on the CSJ TextGrid files for each utterance which was defined as Maekawa’s (2003) transcription units: utterances separated by “longer-than-200 ms” pauses. The data was normalized by converting the raw f0 measurements in semitones (Ward 1970) for pitch range analysis. A linear mixed-effects model (Tukey’s method for post-hoc test) was fit to test the statistical significance of distributions using R (R Core team 2017).
Results and discussion: [Gender] The analysis of the data indicated that Japanese speakers have an average pitch range of 19.565 semitones. Contrary to the results observed for English speakers in Henton (1989), we found a significant effect of gender on pitch range ($t = 3.154, p < 0.01$). Male speakers exhibited a wider pitch range (20.715 st) than female speakers (18.201 st). The result is in contrast not only with Henton’s (1989) observation, but also with our assumption that it would show the opposite pattern (e.g., in the case of English, male speakers are more likely to be identified as “gay-like” (i.e., feminine) when they show a wider pitch range, Levon 2007).

[Register] The analysis of the overall results for register indicate the following pattern in pitch range: $D$ (24.276 st) $>$ $A$ (19.294 st), $S$ (19.378 st) $>$ $R$ (16.344 st). D speech register was significantly higher ($z = 22.682, 22.726, p < 0.01$) than A and S between which no difference was observed ($z = 2.245$, n.s.), and R was the lowest ($z = 5.503, 3.181, p < 0.01$). These results are consistent with the properties of each of the given speech registers. Higher pitch range is observed in D when compared to A and S, which corresponds to the affective dimension of D where two participants are interacting intimately, while A and S consist of public speaking. On the other hand, the read speech in R does not involve any kind of interaction, and therefore exhibits the narrowest pitch range, contrary to Gut’s (2009) observation. As shown in [1], an interaction was also observed between register and gender. Pitch range expansion was observed in male speakers (blue) for D and there was no significant variation between A, R and S, while female speakers (red) were more sensitive to the affective dimensions of register (expansion in D, compression in R). The different behavior of men and women when it comes to the use of pitch variations in different registers can be imputed to the propensity of women to be more sensitive to affective dimensions (presence or absence of audience, of interaction); while on the other hand, men are more likely to be affected by referential ones (stylistic level, formality).

[Style] A significant effect of speech style (degree of formality perceived by listeners, from 1 the less, to 4 the most formal) on pitch range patterns was also observed ($t = 2.112, p < 0.05$). As presented in [2], speeches that were perceived as more formal by listeners also exhibited a narrower pitch range, and vice versa.

[Other perceptual factors] Based on the analysis of the perceptual rating of CSJ speech samples by Japanese listeners, the following factors were shown to be active as predictors of pitch range patterns: spontaneity, speed, fluency, confidence and intelligibility. First, pitch range was consistent with spontaneity [3]: the wider the pitch range, the more spontaneous the speech was perceived ($t = 3.472, p < 0.01$). Pitch range expansion was also associated by listeners with speed [4], as listeners tended to judge a speech as faster when the pitch range was wide ($t = 2.926, p < 0.01$) than when it was narrow. The results also showed an effect of pitch range on the perceived characteristics of the speech/speaker. Namely, speeches with a narrow pitch range tended to be perceived as less fluent ($t = -6.635, p < 0.01$) and less confident ($t = 4.192, p < 0.01$) when compared to those with a wider one. On the other hand, a significant influence of pitch range on intelligibility was also observed ($t = -3.409, p < 0.01$) as a narrower pitch range was judged more intelligible by Japanese listeners.

Conclusion: This study confirmed that language-specific restrictions on pitch range can be overridden by the influence of extra-linguistic factors. In addition to the variations observed in gender, Japanese speakers make use of pitch expansion to adjust to specific pragmatic situations. Moreover, variations in pitch range affect listeners’ perception of the speech and the speakers. The results also suggest that there is a specified domain of pitch range in Japanese, and if the value deviates from the norm the
speech/speaker tends to be identified as less natural (e.g., fluency, confidence, intelligibility), confirming that languages with contrastive pitch accent are subject to less pitch range variations: Variation in Japanese is allowed, but to a restricted degree.

References