

## **Revisiting f0 range production in Japanese-English simultaneous bilinguals**

Calbert Graham

Phonetics Lab, Univ of Cambridge

This study reports an experiment in which 10 Japanese-American English simultaneous bilinguals (5 males, 5 females; all undergrads at the University of California, Berkeley) were recorded performing comparable reading tasks in their two languages. The study builds on a relatively new approach to measuring f0 range - proposed by Patterson 2000 and operationalised in Mennen et al. 2012 - that computes its high and low points from actual tonal targets in the intonational phonology. Also, unlike in most previous studies where f0 range is traditionally treated as a one-dimensional measure, f0 range in both languages was measured along two quasi-independent dimensions: level and span. The results reveal that Japanese was realised at a significantly higher level and with a wider range of frequencies (span) than English. This finding provides new insights into the relation between intonational structure and f0 range in two typologically different prosodic systems.

**Keywords:** Pitch range, f0, Japanese, English, bilinguals, level, span, intonation

### **1. Introduction**

Japanese has been described as a pitch-accent language (Akinaga 1966; McCawley 1968; Goldsmith 1974, Haraguchi 1977, among others). In pitch accent languages pitch<sup>1</sup> serves as the main cue to signal lexical and phrasal distinctions. This contrasts with stress-accent languages like English in at least two important ways:

1. In English pitch accents are primarily prominence lending (i.e. their main function is to make accent-bearing units more intonationally prominent than others to signal various discourse functions).
2. Whereas, on the one hand, accent in Japanese is manifested solely by pitch modulation with no significant use of other material, on the other, in addition to f0 English uses a combination of other acoustic material including duration and amplitude (Beckman, 1986).

An important question for cross-language analysis is whether and in what way these basic differences in the intonational organisation of the two languages interact with the phonetic realisation of speaking fundamental frequency range. Previous research comparing the realisation of pitch range in the two languages has yielded inconclusive findings with some claiming that speaking f0 range in Japanese is typically realised higher than in English due to its phonemic use of pitch (e.g. Yamazawa & Hollien (hereafter Y&H, 1992) on the one hand, and others who argue that such differences are only observable in Japanese women's use of sharp pitch range distinctions to their male counterparts as a gender-defining sociophonetic behaviour (e.g. Ohara, 1999), on the other.

Two of the of the earliest works, to the best of my knowledge, to have investigated cross-language differences in pitch range involving Japanese are Hanley et al. (1966) and Hanley & Snidecor (1967). The Hanley studies compared acoustic data for Japanese and other languages that have been traditionally described as belonging to typologically different classes (Spanish, English; and Tagalog which was

---

<sup>1</sup> I am not making the technical distinction between fundamental frequency as an objective notion of mathematical periodicity and pitch as a more subjective notion of perceived sine wave frequency. The terms will be used interchangeably.

analysed only in the 1967 study). Hanley and his colleagues compared the median fundamental frequencies and the standard deviations in data elicited in a read passage and a spontaneous speech task. For both speech samples, they found a statistically significant main effect of language on the median f<sub>0</sub>s, but not on the standard deviations. Based on the findings of the first study they ranked Japanese as the highest on a pitch continuum from high to low, followed by Spanish and American English, respectively. However, the second study failed to replicate these earlier findings and the ranking appeared far less conclusive than earlier reported. More precisely, they found no statistically significant contrasts between the four languages, only strong tendencies of English and Tagalog to have lower pitch ranges than Japanese and Spanish. If the f<sub>0</sub> excursion from a high to a low tonal target associated with an accented mora in Japanese was inherently greater than that associated with the pitch accent movement in stress-accent languages like English or Spanish then Japanese would, logically, be expected to manifest higher f<sub>0</sub> ranges than those language, but this has not been sufficiently established from these studies, and it is not even entirely clear whether these measures of speaking f<sub>0</sub> range were adequate to determine this either way.

Loveday (1981) focussed on the sociophonetic influences of pitch use in Japanese by examining the pitch correlates of politeness formulae produced by both male and female Japanese and English speaking informants. The study showed that there were significant pitch-related contrasts in the expression of politeness between the two language groups. He reported that the Japanese females adopted a significantly higher pitch range that differentiated them acoustically from their male counterparts in contrast to the English males and females who appeared less differentiated in this way. Based on this finding, Loveday (1981:1) concludes that there is some contrast in the “sociosemiotic function” assigned to pitch in the two speech communities whereby high pitch is reserved for the enactment of female roles in Japanese society, as opposed to in America where high pitch is adopted by both genders to express politeness.

Van Bezooijen (1995) conducted a comparative study on the sociocultural aspects of f<sub>0</sub> range differences between Japanese and Dutch women and found that Japanese women in the study used higher f<sub>0</sub> ranges than their Dutch counterparts, somewhat similar to what Loveday (1981) found for Japanese and American women.

Y&H, 1992 compared f<sub>0</sub> range realisations of Japanese females to that of Caucasian American females (both groups being bilingual speakers of the two target languages) and found that the Japanese women in the study exhibited significantly higher f<sub>0</sub> profiles in their English productions than the Caucasian Americans. They argued that these differences are likely to have resulted from fundamental prosodic differences between the two languages whereby, as mentioned earlier, f<sub>0</sub> serves a phonemic function in Japanese, as opposed to in English where it serves primarily an intonational function in marking discourse prominence. However, it is perhaps regrettable that this study only focussed on Japanese females, making it difficult to rule out the possibility that these differences may, at least in part, be attributable to the sociophonetic influences on f<sub>0</sub> range reported in Loveday (1981) and van Bezooijen (1995), among others. Furthermore, it is also not adequately established whether all the bilinguals in the study were of comparable competency in both of their languages, which could conceivably be another source of confound. In fact, Ohara (1999) did not find the sort of differences reported in Y&H (1992) in Japanese male speakers and argued instead that observed contrasts in f<sub>0</sub> range between Japanese and English

cannot be linked to prosodic differences between the two languages but rather to the sociocultural influences. Finally, it cannot be entirely ruled out that the findings of Y&H (1992) may be attributable, at least in part, to presumed physiological differences between Japanese and Caucasian Americans due to the inverse correlation between the size of a speaker's vocal tract and the formant frequencies it typically realises (see Ohala 1983 for a further discussion of the effect of organ size in pitch production).

Yet, despite criticisms of Y&H (1992), there is ample perceptual evidence that listeners are indeed able to tell languages apart based solely on their prosody (Ohala & Gilbert, 1981). Ohala & Gilbert sought to establish whether Japanese, English and Cantonese native speakers can tell these three languages apart when given stimuli with segmental material removed, and only prosodic cues ( $f_0$ , amplitude and timing characteristics) remaining. The results of this experiment confirm that listeners can indeed identify languages based solely on prosodic cues, particularly their own language. Other studies have corroborated this finding (Maidment, 1983), and have even shown that at just a few days after birth infants are able to differentiate between languages in this way (Mehler et al. 1988) – see Komatsu (2007) for a detailed review of these and other similar experiments. As traditional measures of  $f_0$  range have generally not taken serious account of systematic differences in the prosodic structure of languages in the computation of  $f_0$  range, it is conceivable that with proper quantification we may yet be able to unravel the operations of the underlying structural properties that give rise to these perceived prosodic differences between languages.

For completeness, many other factors have been reported to influence the realisation of  $f_0$  range, including segmental composition (see Gussenhoven 2004 for an extensive discussion).  $F_0$  range has also been found to differ between languages, depending on speech materials (Keating & Kuo, 2012), with various categorical effects of sentence type (e.g. Prieto 2004 who found that Spanish declarative statements tend to have lower first  $H^*$  peaks compared to interrogatives and exclamatory sentences). These differences appear to be language specific and suggest that languages may vary widely in how they implement prosodic cues that differentiate different sentence types. This is a particularly important factor to consider in measuring  $f_0$  range, and it is perhaps not without consequence that previous studies comparing Japanese and English have not attempted to tease out its potential effects on  $f_0$  range manifestation.

In building a framework for measuring pitch range, Ladd (1996) argues that pitch range can vary along two quasi-independent dimensions: (1) level, which refers to how high or low a speaker's pitch is; and (2) span, which refers to how narrow or wide the range of frequencies is realised. Taking Ladd's approach as a starting point, Patterson (2000) carried out further investigations into how to best quantify these two related but distinct aspects of pitch range in a way that optimises perceptual validity. Such an approach may shed new light on the apparent perception-production discrepancy between the findings of Ohala & Gilbert (1981) and Ohara (1992).

Patterson (2000) compared traditional measures of pitch range to his new prosodic approach. Traditionally measures of mean  $f_0$  and median  $f_0$  have been used for level (Mennen et al., 2008) and long-term distributional measures (LTDs), such as maximum minus minimum  $f_0$  (Cosmides 1983), four standard deviations around the mean (Jassem 1971), the difference between the 95<sup>th</sup> and 5<sup>th</sup> percentile (Horii 1975), among others have been used for span (Patterson, 2000). Patterson (2000) confirms that measures of pitch range based on long-term distributional properties of  $f_0$  are

indeed problematic (e.g. octave errors towards the end of an utterance would prevent accurate measurements being taken) and instead proposes that (Patterson 2000:36) “an alternative to measuring span and level in terms of long term f0 distribution is to see them as fundamentally linked to *tonal targets* found in speech.” This proposal is a logical extension of the idea that f0 targets are the phonetic manifestations of underlying high and low points associated with pitch accents and boundary tone movements, as represented in the autosegmental metrical approach (Pierrehumbert, 1984; Beckman & Pierrehumbert, 1986; Ladd, 1996; Silverman et al., 1992, among others).

More recently, Mennen and her colleagues (2012) operationalised Patterson’s claim in a methodological study on the linguistic dimensions of pitch range differences between German and English. The research focussed on female speakers of Southern Standard British English (SSBE) and Northern Standard German (NSG) in order to test an anecdotal observation that speakers of SSBE have a wider pitch range than German speakers. The results of the study revealed that the linguistic measures of pitch range outperformed the long-term distributional measures in uncovering evidence in support of these hitherto unproven contrasts between SSBE and NSG.

On the basis of the foregoing review, I propose that an approach to measuring pitch range based on underlying tonal targets in an utterance fits intuitively with perceptual judgments of f0 structure in speech and, accordingly, could conceivably be better able than long term distributional measures to reveal f0 range differences arising from structural differences in the prosodic systems of Japanese and English, if indeed such contrasts exist. These new considerations provide sufficient justification to warrant revisiting the question of whether pitch-accent languages are typically realised with higher f0 range profiles than stress accent languages. To that end, this study proposes to tease apart the sociophonetic or cultural influences that have been reported to be associated with the enactment of female roles from prosodic influences that are likely to apply irrespective of gender. Using linguistic measures of f0 range that take intonational differences and sentence-type information into account I ask:

1. Do simultaneous Japanese-English bilinguals realise f0 range differently depending on which language they are speaking?
2. If so, what aspect(s) of f0 range (i.e. level/span) vary between the two languages?

## 2. Method

### 2.1. Participants

Participants in the study were 10 simultaneous<sup>2</sup> bilingual adult speakers of Tokyo Japanese (hereafter Japanese) and American (Californian) English. All 10 participants were undergraduate students at the University of California Berkeley, but have lived both in Japan and the United States at various stages of their life. There were five males and five females - All between 19 and 25 years of age. All participants were judged by monolingual speakers of their respective languages to be indistinguishable from other native speakers (i.e. having no trace of a foreign accent in either of their

---

<sup>2</sup> These participants grew up speaking both Japanese and English and were judged by independent raters to speak English with no noticeable trace of a foreign accent

two languages). These judges listened to a wide range of recorded speech samples for 20 participants who identified themselves as balanced (simultaneous) bilinguals, rating them in terms of how foreign accented they perceived them to be on a scale of 1 to 3, where 1 is not at all (native), 2 is advanced and 3 means basic. Only those who were placed in the native group for their two languages by at least two of three raters were used in the study. Further, all participants completed a questionnaire on their linguistic background, and only those who indicated that they did not speak a third language at an advanced level were included in the study reported here. None of the participants reported any hearing or speaking difficulties. All participants were compensated for participation as appropriate.

## 2.2. Materials

The datasets consisted of declarative statements and three types of questions, adapted from a set of the materials used in the IVIE project (Grabe, Post & Nolan, 2000; 2001). The target sentences were translated into Japanese, with a few minor adaptations to enable Japanese-English cross language analysis (e.g. the need to carefully select only accented words that would normally receive a pitch accent in Tokyo Japanese). However, on the whole the segmental composition of target words was generally comparable between the two languages (e.g. as far as possible, only fully voiced target words were used in order to avoid f<sub>0</sub> discontinuity associated with voiceless segments). Short sentences corresponding typically to a single intonation phrase were chosen to control for differences in pitch accent realisation across intonational phrase boundaries. Below is a sample of the target utterances that were used to elicit the subjects' intonation realisation in various sentence patterns<sup>3</sup> (see appendix 1 for the complete list of test items):

- (a) 6 declarative statements (e.g. *We remembered Lil*; [リルの事を思い出した])
- (b) 6 declarative questions (i.e. without morphosyntactic markers). (e.g. *You remembered Lil?* [リルの事を思い出した?])
- (c) 5 Wh-questions (e.g. *Why will you be in Ealing?* [なんでイーリングにいる?])
- (d) 5 alternative questions with the conjunction 'or' (e.g. "Did you say red or bed?" [レッドとベッドのどっちを言ったの?])

## 2.3. Recording procedure

Materials were presented in PowerPoint (with half of the participants in each group having the Japanese test items first followed by the English ones, and was reversed for the other half to avoid any possible task presentation effects – such as general declination or task fatigue). Recordings were made in a recording booth in the Phonology Lab of the University of California, Berkeley using a Zoom H2 recorder with a sampling frequency of 44.1 KHz and a 16 bit resolution. The participants were asked to carefully read each sentence at their normal pace and to repeat if they made an error.

## 2.4. Analysis

The study adapted the research methodology of Mennen et al (2008), itself based on

Patterson (2000), in the quantification of pitch range using linguistic measures. This approach distinguishes between prominent and non-prominent peaks and valleys and makes a further distinction between initial and non-initial peaks in line with Patterson's findings that they patterned differently. Linguistic measurements were made for the two aspects of pitch range described earlier: level and span.

The following phonetic landmarks were annotated in Praat by two independent labellers; inter-annotator agreement according to Cohen's kappa was 0.78. A kappa greater than .07 is generally considered satisfactory.

- (1) Utterance onset f0 – this is defined as the f0 of the first voice frame, marked as “O”
- (2) The highest f0 in each H-star accent was marked as iH\* (initial peak) and H\*2 (non-initial peaks) depending on its position within the utterance. The troughs after all H\*s, where present, were marked as L for low.
- (3) The final low tone (marked as FL in statements and non-rising Wh-questions).  
FL is also the f0 landmark just before the rise in declarative questions

The labelling was done according to TOBI annotation guidelines. Only falling pitch accents with clear f0 peaks (H\*) & troughs (L) were labelled; accents realised as plateaus were excluded from these analyses as they lack visible peaks and troughs to take measurements. Where voices end in a creak, measurements were taken immediately before. All values were initially extracted in Hertz using a Praat script and converted to a semitone scale. See the Figure 1 below for a sample annotation

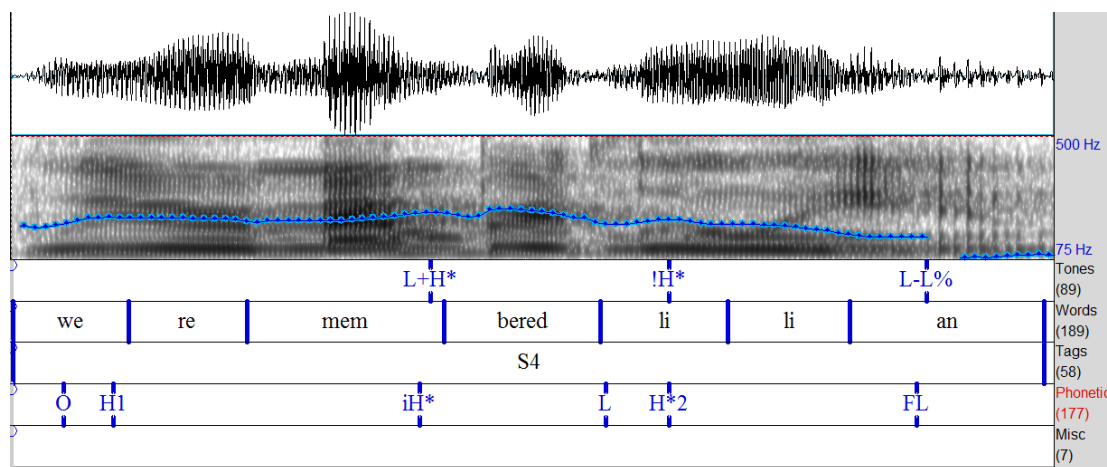


Fig. 1. Praat textgrid showing a sample of measurements point in a declarative statement produced by one of the participants.

## 2.5. Linguistic measures of pitch range

From these data measurements, the following measures for span and level are derived:

### 2.5.1. Measures of pitch level

1. Mean utterance onset f0 ('O')
2. Mean of initial iH\* peaks
3. Mean of non-initial H\* peaks
4. Mean of post-accentual low (L) tones
5. Mean of phrase final FL tones

In order to derive a comparable measure of pitch level in the four sentence types in the two languages I initially used 1, 2, 4 and 5 and excluded 3 (i.e. non-accentual peaks), as they were generally absent from declarative questions and Wh-questions. However, they are included in a separate analysis on overall peak difference (i.e. the difference between the first and second H\* peaks).

### 2.5.2. Measures of pitch span

There were three measures for span:

1. iH\* minus L
2. iH\* minus FL for the four utterance types
3. Peak difference (iH\* - H\*)

### 2.5.3. Statistical analysis

A Shapiro-Wilk test determined that the distribution of the data did not vary significantly from normality and therefore parametric ANOVAs were used for each measure of pitch range in separate analyses. For level, the data for males and females are processed separately to control for gender effects on f0 range. However, they are merged for span as there was no statistically significant effect of gender on this measure.

## 3. RESULTS

### 3.1. Linguistic level

#### 3.1.1. Females

For pitch level a repeated measures ANOVA with language (English, Japanese) as the independent variable and utterance type (4 levels: statements, Wh-questions, declarative questions and alternative questions) and pitch level measures (4 levels: sentence onset f0, first H\*peak f0, post accentual low and phrase final f0) as the dependent variable was conducted. The ANOVA revealed a significant effect for language ( $F(1, 5) = 19.10$ ,  $P < 0.05$ , partial eta squared .83); and a significant interaction between language and measures, ( $F(3, 12) = 33.12$ ,  $p < 0.001$ , partial eta squared .89), and a three way interaction between language, level measures and utterance type ( $F(3, 12) = 11.04$ ,  $p < 0.05$ ). On the significant effect of language, overall level in Japanese was determined to be higher than English by an average of 26 Hz. Post hoc tests further determined that Japanese was realised significantly higher than English on two measures: iH\* and onset f0 for all sentence types. Furthermore, declarative questions in English were realised significantly lower than statements and Wh-questions. Overall results are shown in Fig. 2 followed by a breakdown by sentence type in Table 1, below:

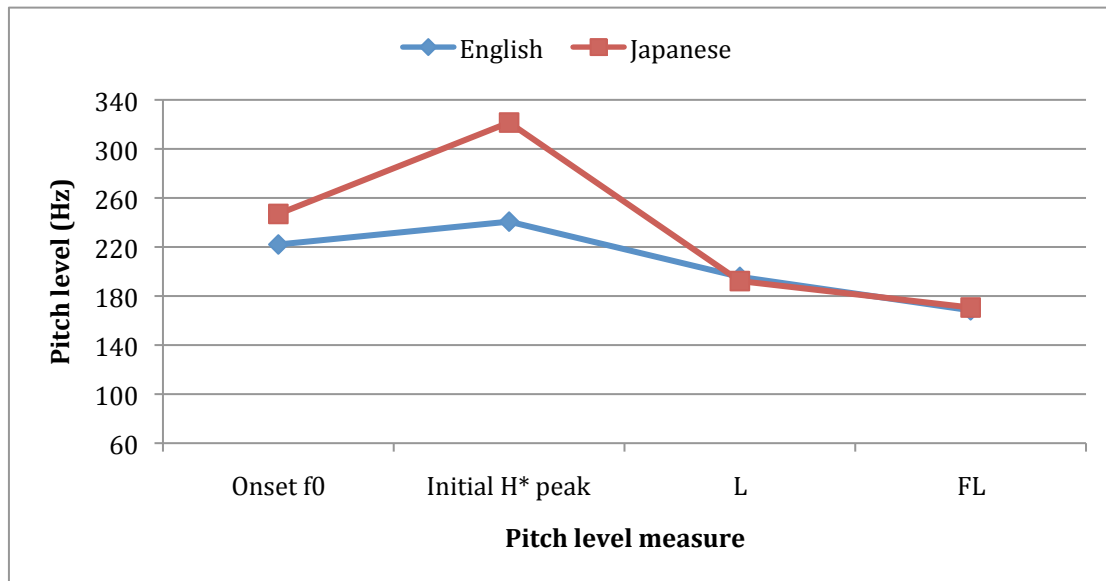


Fig. 2. Overall f0 level distribution in Japanese and English sentences produced by Japanese-English bilingual females. Onset f0 and initial H\* peaks (iH\*) are realised significantly higher in Japanese than in English.

Table 1. F0 level in Japanese and English by sentence type

F0 measure	level	Sentence type	English	Japanese	Sig. (p<.05)
<i>Onset f0</i>		S	225	248	*
		DQ	222	244	*
		WhQ	219	247	*
		AQ	222	249	*
<i>iH*</i>		S	245	317	*
		DQ	190	311	*
		WhQ	280	334	*
		AQ	248	324	*
<i>L</i>		S	190	191	n.s.
		DQ	194	194	n.s.
		WhQ	207	198	n.s.
		AQ	191	187	n.s.
<i>FL</i>		S	164	165	n.s.
		DQ	159	176	n.s.
		WhQ	174	173	n.s.
		AQ	176	169	n.s.

Table 1 shows the breakdown of f0 level realisations by language & sentence type (S – statements; DQ – declarative questions; WhQ – Wh-questions; AQ – alternative questions) in Japanese-English bilingual females. Stars (\*) represent significant contrasts; ‘n.s.’ stands for ‘not significant’.

### 3.1.2. Males

The results were generally similar to those of the females as follows. The ANOVA confirmed a significant effect for language ( $F(1, 4) = 17.73, P < 0.05$ , partial eta squared .82); and a significant interaction between language and level measures, ( $F(3,$



12) =10.81,  $p < 0.01$ , partial eta squared .65), and a three-way interaction between language, sentence type and measures ( $F(9, 36) = 3.9$ ,  $p < 0.01$ ). Overall, Japanese was realised higher than English by an average 11Hz by the male bilinguals. Post hoc tests revealed that similar to the results for the females, the male bilinguals realised Japanese significantly higher than English in terms of onset  $f_0$  and initial  $H^*$  (or  $iH^*$ ). Also, onset  $f_0$  and  $iH^*$  in declarative questions were significantly lower than in statements and Wh-questions. Overall results are shown in Fig. 3, followed by a breakdown by sentence type in Table 2:

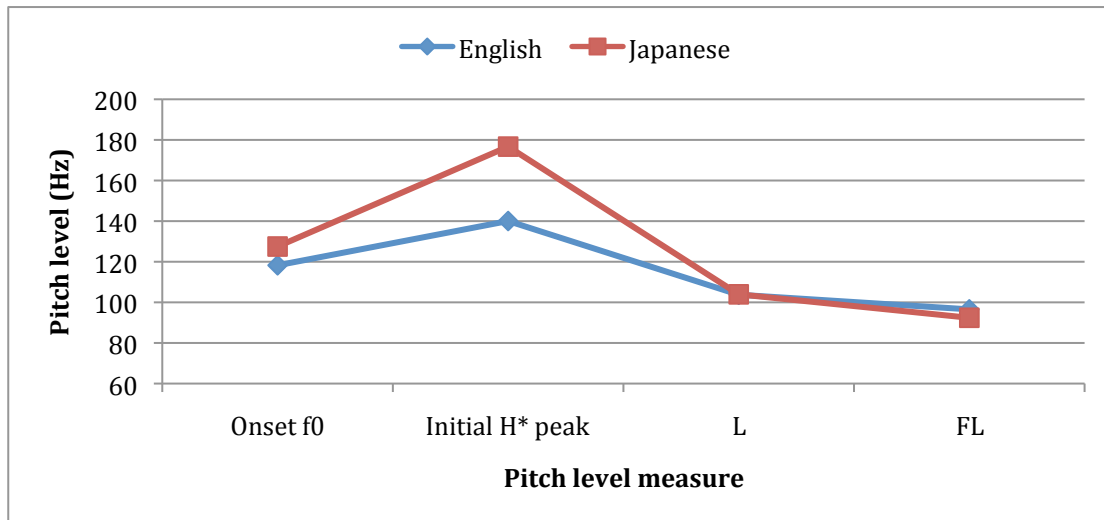


Fig. 3. Overall  $f_0$  level distribution in Japanese and English sentences produced by Japanese-English bilingual males. Similar to the findings in the females, in the males, onset  $f_0$  and initial  $H^*$  peaks ( $iH^*$ ) are realised significantly higher in Japanese than in English.

Table 2.  $f_0$  level in Japanese and English by sentence type

<b><math>f_0</math> measure</b>	<b>level</b>	<b>Sentence type</b>	<b>English</b>	<b>Japanese</b>	<b>Sig. [<math>p &lt; .05</math>]</b>
<i>Onset <math>f_0</math></i>		S	119	126	*
		DQ	120	125	*
		WhQ	115	130	*
		AQ	118	129	*
<i><math>iH^*</math></i>		S	135	167	*
		DQ	124	177	*
		WhQ	156	183	*
		AQ	146	179	*
<i>L</i>		S	103	101	n.s.
		DQ	100	109	n.s.
		WhQ	107	103	n.s.
		AQ	105	102	n.s.
<i>FL</i>		S	94	90	n.s.
		DQ	110	94	n.s.
		WhQ	94	95	n.s.
		AQ	92	91	n.s.

Table 2 shows the breakdown of  $f_0$  level realisations by language & sentence type (S – statements; DQ – declarative questions; WhQ – Wh-questions; AQ – alternative questions) in Japanese-English bilingual males. Stars (\*) represent significant contrasts and n.s. stands for ‘not significant’.

### 3.2. SPAN

As earlier indicated, the results for male and females are reported together as the ANOVA determined that there was no significant effect of gender on span. There was an overall significant effect of language ( $F(1, 8) = 88.9, p < 0.001$ , partial eta squared .92); language and sentence type ( $F(3, 24) = 5.95, p < .01$ , partial eta squared .43) and a three way interaction between language, sentence type and measures ( $F(3, 24) = 4.89, p < .01$ , partial eta squared, .38). Overall, span in Japanese was realised 4.8 semitones wider than in English. Post hoc tests revealed that for Japanese there was no effect of span on sentence type. For English, span was comparable across sentence types, apart from declarative questions which had a significantly lower span compared to the other three sentence types, all at the 5% level of significance. These results are shown in figures 10 -14 below:

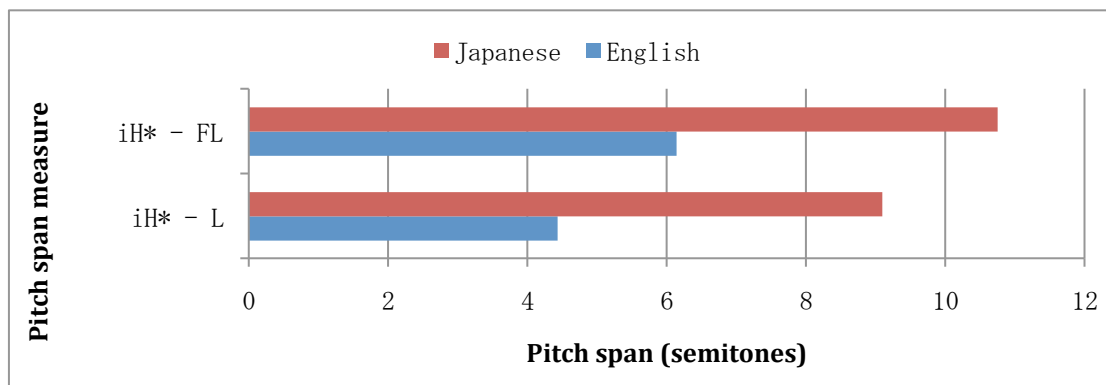


Fig. 4. Overall span averaged across all sentences of each of the two languages produced by all Japanese-English bilinguals in the study.

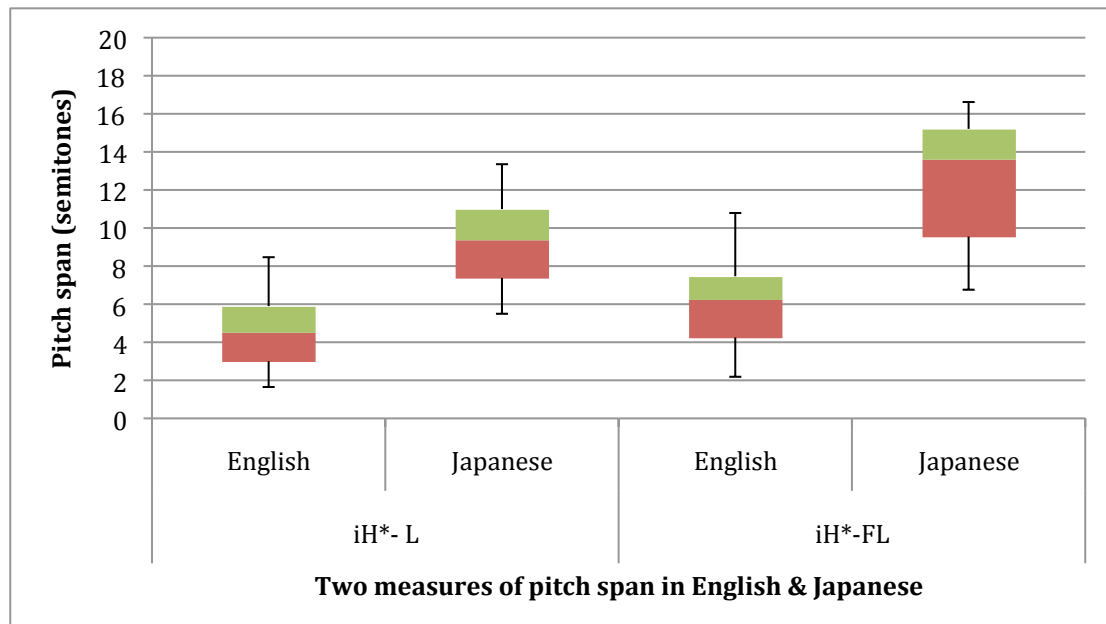


Fig. 5. Overall F0 span in Japanese and English sentences produced by all Japanese-English bilinguals in the study. Japanese was realised significantly wider than English on both measures.

**Table 3. F0 span in Japanese and English by sentence type**

F0 measure	span	Sentence type	English	Japanese	Sig. (p<.05)
<i>iH* - L</i>		S	4.45	8.58	*
		DQ	2.64	8.96	*
		WhQ	5.71	9.39	*
		AQ	4.92	9.43	*
<i>iH* -FL</i>		S	6.90	10.92	*
		DQ	2.49	10.30	*
		WhQ	8.40	11.33	*
		AQ	6.75	11.44	*

Table 3 shows pitch span distribution by sentence type (S – statements; DQ – declarative questions; WhQ – Wh-questions; AQ – alternative questions) in Japanese-English bilingual males and females. Japanese was realised significantly wider than English in all sentence types on both measures.

### 3.3. Peak difference

#### 3.3.1. Females

This analysis compares the level and span of all first and second H\* peaks in the dataset in both languages. Second H\* peaks were generally missing from English declarative questions and Wh-questions.

A paired sample t-test determined that *level* is also significantly higher in Japanese than in English if defined as non-initial accent peaks ( $t(5)=2.87, p<.05$ ). The difference between first and second peaks was significantly larger in Japanese, suggesting a greater degree of declination: ( $t(5) = 8.15, p<.05$ ). The combined f0 distribution across all four sentence types is presented in figure 15 below:

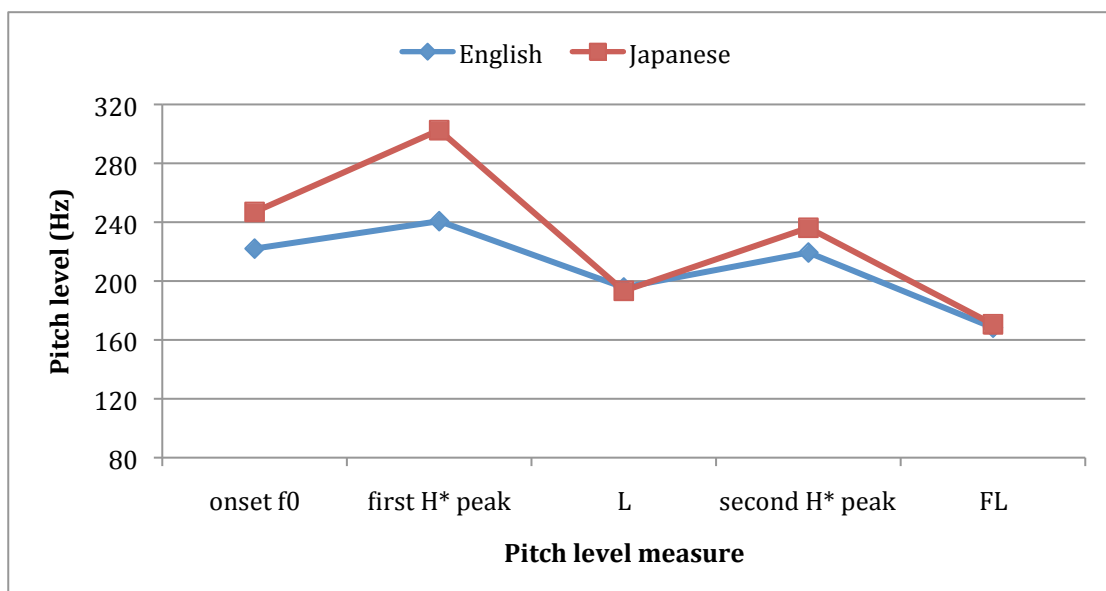


Fig. 6. shows overall f0 at different pitch points produced by the bilingual females. Japanese was significantly higher than English at onset f0, first H\*peak and second H\* peak points; both languages were comparable at L and FL.

### 3.3.2. Males

The results for the males were comparable to the females and confirm that the level of non-initial peaks was significantly higher in Japanese than in English ( $t(5) = 4.91$ ,  $p < .01$ ), and that Japanese was realised with a wider span between successive peaks than English: ( $t(5) = 2.77$ , all at  $p < .05$ ). The combined  $f_0$  distribution across all four sentence types is shown in figure 16 below:

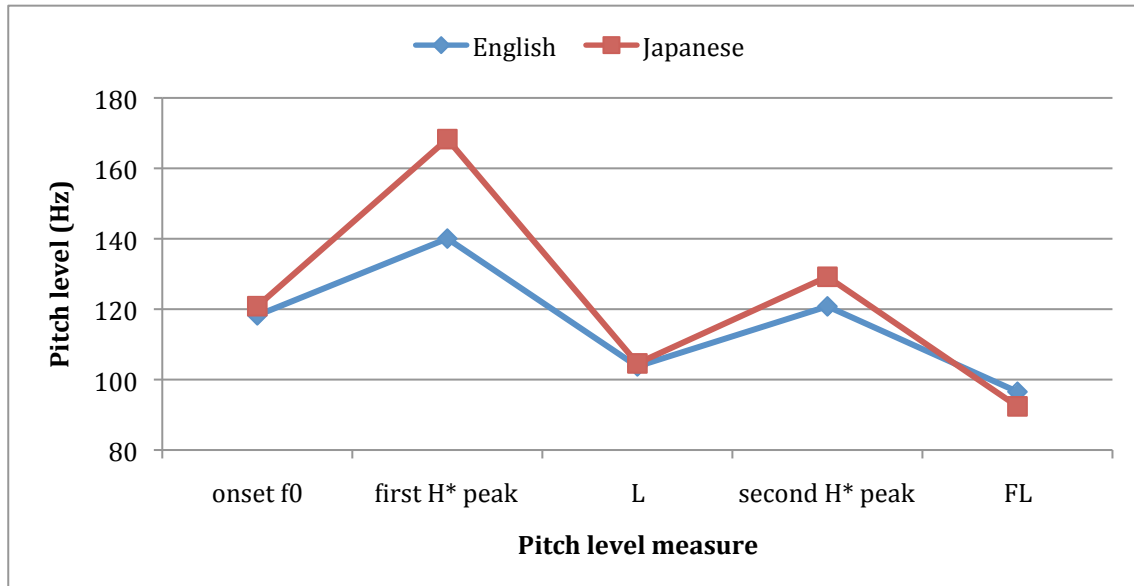


Fig. 7. shows overall  $f_0$  at different pitch point measures produced by the bilingual males. Japanese was significantly higher than English on onset  $f_0$ , first H\* peak and second H\* peak; they were comparable on F and FL.

## 4. Discussion

This study examined whether the  $f_0$  range of simultaneous Japanese-English bilinguals vary depending on which of their two languages they are speaking. I addressed the hypothesis that pitch range in Japanese as a language that uses pitch accents to mark lexical distinctions is likely to be realised at a higher level (and perhaps with a wider span) than in English which does not systematically use pitch to make such phonemic contrasts. The overall results confirm this hypothesis that  $f_0$  range is realised differently between the two languages in both dimensions (i.e. pitch level in Japanese is higher and the range of frequencies is wider), and revealed that these differences were not gender-specific, and thus unlikely to be due merely to the sociophonetic factors discussed earlier. Instead, I argue that they can be attributed to inherent differences in the prosodic systems of the two languages. It is remarkable that the bilinguals were realising high tonal targets at a higher level and low targets at a lower level in Japanese, which resulted in an expanded pitch span. These results appear to accord with Ohala & Gilbert's (1981) perceptual finding that languages (including Japanese and English) can be distinguished based solely on their prosody, and lend further support to Patterson's (2000) suggestion, operationalised in Mennen (2012), that linking pitch range to tonal targets in speech is more likely to uncover linguistically and perceptually valid results in pitch range analysis. Pitch accent distribution is highly constrained in Japanese, which in turn imposes important constraints on spoken-word recognition (Cutler and Otake, 1999). Perhaps this may partly determine the kinds of strategies that the bilinguals in this study used that

resulted in the differentiation between the  $f_0$  ranges of their two languages. These strategies include realising higher utterance onset levels, higher accent peaks and greater declination in Japanese (e.g. the difference between initial and non-initial peaks was larger in Japanese than in English). Given that pitch accents in Japanese create lexical distinctions depending on their location in words, and can contrast syntagmatically with unaccented words, it is unsurprising that there is an apparent greater constraint in Japanese than in English to make the high and low elements of the single pitch accent movement more perceptually salient. Although utterance onset does not necessarily coincide with any pitch accent movement in either language, it might be argued that the significantly higher utterance onset in Japanese allows the transition to the  $H^*$  of the pitch accent to be less steep and therefore able to be realised more quickly and with less  $f_0$  fluctuation.

## **5. Conclusion**

This study investigated the realisation of pitch range in simultaneous Japanese-English bilinguals in order to test an unresolved issue in the literature on whether Japanese which uses pitch accents to mark lexical distinctions will be realised at a higher pitch level or with a wider range of frequencies than a stress accent language like English. Overall, the results of the study reveal that Japanese is spoken by balanced bilinguals at a higher level and within a wider span than English irrespective of sentence type, and confirm that analysis of  $f_0$  range can be enhanced by an approach that takes both intonational phonology and phonetic implementation into account, whilst controlling for sociophonetic and non-linguistic sources of variation. These findings offer support for Yamazawa & Hollien's claim (1992), but against Ohara (1999), and provide new insights into the relation between  $f_0$  range and intonation structure in two prosodically different languages.

## **6. Acknowledgements**

I wish to thank the following persons who have given their support at various stages of this research:

Brechtje Post and Bob Ladd for very insightful discussions on intonation theory and pitch range.

Keith Johnson, John Ohala, Larry Hyman, Jonah Katz and members of the UC Berkeley Phorum who attended the initial presentation of the results presented in this paper. Their feedback has been invaluable. Many thanks to the Berkeley students who willingly participated in this study, and of course to Keith and Ron Sprouse who have provided tremendous support during my semester away at Berkeley. Thank you all.

## **References**

- Akinaga, K. 1966. *Kyootsuugo no akusento. Nihongo hatuon akusento ziten*, ed. by Nihon Hoosoo Kyookai, 45-90. Tokyo: Nihon Hoosoo.
- Beckman, M. 1986. *Stress and non-stress accent*. Dordrecht: Foris.
- Brown, A., and Docherty, G. J. (1995). "Phonetic variation in dysarthric speech as a function of sampling task," *Eur. J. Disord. Commun.* 30(1), 17-35.
- Chen, S. H. (2005). "The effects of tones on speaking frequency and intensity ranges in Mandarin and Min dialects," *J. Acoust. Soc. Am.* 117(5), 3225-3230.
- Deutsch, D., Henthorn, T., and Dolson, M. (2004). "Speech patterns heard early in life influence later perception of the tritone paradox," *Music Percept.* 21, 357-372.
- Goldsmith, J. 1974. An autosegmental typology of tone: And how Japanese fits in. *Papers from the fifth annual meeting, North Eastern Linguistic Society*, 172-182.

- Grabe, E., Post, B., Nolan, F., and Farrar, K. (2000). Pitch accent realisation in four varieties of British English. *Journal of Phonetics* 28, 161-185.
- Hanley, T.D., Snidecor, J.C., and Ringel, R. (1966). "Some acoustic differences among languages," *Phonetica* 14, 97-107.
- Hanley, .D. and Snidecor, J.C. (1967). "Some acoustic similarities among languages," *Phonetica* 17, 141-148.
- Haraguchi, S. 1977. *The Tone Pattern of Japanese: An auto segmental theory of tonology*. Tokyo: Kaitakusha.
- Keating, P. & Kuo, G. (2012). "Comparison of speaking fundamental frequency in English and Mandarin", *J. Acoust. Soc. Am.* 132, 1050-1060.
- Komatsu, M. (2007). *Reviewing Human Language Identification*. In Müller, C. (Ed.), *Lecture notes in Computer Science* vol. 4441, Springer-Verlag, Berlin.
- Ladd, R. (1996). *Intonational Phonology*. Cambridge University Press.
- Ladd, D. R., Faulkner, D., Faulkner, H., & Schepman, A. (1999). Constant "segmental anchoring" of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America*, 106, 1543-1554.
- Loveday. (1981). "Pitch, politeness and sexual role: An exploratory investigation into the pitch correlates of English and Japanese politeness formulae," *Lang Speech* 24, 71-89.
- McCawley, J. D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.
- Maidment, J. A. (1983). "Language recognition and prosody: further evidence," *Speech, hearing and language: Work in progress*, University College London 1, 133-141.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J., and Amiel-Tison, C. (1988). "A precursor of language acquisition in young infants," *Cognition* 29, 143-178.
- Mennen, I., Schaeffler, F., and Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English & German. *Journal of the Acoustical Society of America*, 131 (3), 2249-2260.
- Ohala, J. J., and Gilbert, J. B. (1979). "Listeners' ability to identify languages by their prosody," in *Problèmes de Prosodie*, edited by P. Léon and M. Rossi (Didier, Ottawa), pp. 123-131.
- Ohara, Y. (1992). "Gender-dependent pitch levels: A comparative study in Japanese and English," *Locating power: Proceedings of the second Berkeley women and language conference*, edited by K. Hall, M. Bucholtz, and B. Moonwoman (Berkeley: Berkeley Women and Language Group, Berkeley), pp. 469-477.
- Patterson, D. (2002). "A linguistic approach to pitch modeling." Doctoral dissertation, Edinburgh University, UK.
- Prieto, P. (2004). "The Search for Phonological Targets in the Tonal Space: Evidence from Five Sentence-Types in Peninsular Spanish". *Laboratory Approaches to Spanish Phonology*, ed. by Timothy Face, pp. 29-59. Mouton de Gruyter: The Hague.
- Van Bezooijen, R. (1995). "Sociocultural aspects of pitch differences between Japanese and Dutch women. *Lang Speech* Jul-Sep; 38-3, 253-65.
- Yamazawa, H., and Hollien, H. (1992). "Speaking fundamental frequency pattern of Japanese women," *Phonetica* 49, 128-140.

## Appendix 1

### English test items

#### 1. Declarative Statements

We remembered Lil.

We remembered lilies.

We remembered Lillian.

We will lose Bill.

We will lose Billy.

We will lose billions.

#### 2. Declarative Questions

You remembered Lil?

You remembered lilies?

You remembered Lillian?  
You will lose Bill?  
You will lose Billy?  
You will lose billions?

3. Wh-questions

When will you be in Ealing?  
Where is the manual?  
Why are you in Ealing?  
Why are we in a limousine?  
Why is he on the bed?

4. Alternative questions

Is his name Miller or Mailer?  
Do you live in Ealing or Reading?  
Did you say mellow or yellow?  
Did he say red or bed?  
Are you growing limes or lemons?

*Japanese test items*

1. Statements

リルの事を思い出した  
'riru no koto wo omoidashita'  
リリズの事を思い出した  
'ririzu no koto wo omoidashita'  
リリアンの事を思い出した  
'ririan no koto wo omoidashita'  
ビルを失うことになる  
'biru wo ushinau koto ni naru'  
ビリーを失うことになる  
'birii wo ushinau koto ni naru'  
「ビリアン」を失うことになる  
'birian (billy-ann) wo ushinau koto ni naru'

2. Declarative questions

ビルを失うことになる?  
'biru wo ushinau koto ni naru'  
ビリーを失うことになる?  
'birii wo ushinau koto ni naru?'  
「ビリアン」を失うことになる?  
'birian wo ushinau koto ni naru?'  
リルの事を思い出した?  
'riru no koto wo omoidashita'  
リリズの事を思い出した?  
'ririzu no koto wo omoidashita'

リリアンを失うことになる？  
'ririan wo ushinau koto ni naru?'

### 3. Wh-questions

いつイーリングにいる？

'itsu iiringu ni iru?'

マニュアルはどこにある？

'manuaru wa doko ni aru'

なんでイーリングにいる？

'nande iiringu ni iru?'

なんでリムジンの中にいる？

'nande rimujin no naka ni iru?'

なんで彼はベッドの上にいる？

'nande kare wa beddo no ue ni iru?'

### 4. Alternative questions

彼の名前はミラーと、メイラーのどっちなの？

'kare no namae wa miraa to mairaa no dochi na no?'

イーリングとレディングのどっちに住んでいるの？

'iiringu to reddingu no dochi ni sunde iru no?'

メローとイエローのどっちを言ったの？

'meroo to ieroo no dochi wo itta (iutta) no?'

レッドとベッドのどっちを言ったの？

'reddo to beddo no dochi wo itta (iutta) no?'

ライムとレモンのどっちを育てるの？

'raimu to remon no dochi wo sodatere no?'