9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

Bilingualism, pitch range and social factors: preliminary results from sequential Japanese-English bilinguals

Elisa Passoni[1], Adib Mehrabi[1], Erez Levon[1], Esther de Leeuw[1]

[1]Queen Mary University of London, United Kingdom
{e.passoni, a.mehrabi, e.levon, e.deleeuw }@qmul.ac.uk

Abstract

Previous research shows that pitch range varies across languages and dialects as a result of different linguistic structures as well as of extra-linguistic factors which influence prosody. Mastering the pitch range of a second language (L2) has been reported to be a particularly challenging task, due to the multiple functions of prosody and the various domains along which it varies, from linguistic to socio-cultural. We designed a reading task aiming to explore the effect of bilingualism on socially constrained attributes of pitch range in female and male Japanese native speakers, who have acquired English as an L2 after having fully acquired their first language (L1, i.e. Japanese). Specifically, we looked at the effect of formality on the pitch range of both female and male speakers, while they were addressing both female and male recipients. Initial results from a small pilot indicate that, irrespective of whether female or male, surprisingly, Japanese-English bilinguals had a lower mean F0 in Japanese than in English; and a wider span in Japanese than in English. Moreover, the Japanese-English bilingual female displayed more pitch variation in the different formality settings than did the Japanese-English bilingual male.

Index Terms: pitch range, prosody, L2 acquisition, bilingualism, Japanese, gender.

1. Introduction

Do bilinguals vary their pitch range according to the language they speak? Does such variation have social significance? These are the questions we explored in this preliminary study. Mastering the pronunciation of an L2 is considered extremely difficult, yet mastering the prosodic system of an L2 is perhaps an even more challenging task due to the complexity of prosody [1]. Previous research reports that prosody is affected by social factors [see 2; 3, among others], and it could be argued that social variation in prosody is very challenging for an L2 speaker to acquire. One motivation of the present work is to further understand to what extent this social variation is acquired, particularly with regard to addressing people of both sexes in different social contexts with varying degrees of formality.

Following Ladd [4], here pitch will be considered as a manifestation of Fundamental Frequency Range (FFR, also called pitch range), which varies across two quasi-independent dimensions: (1) pitch level and (2) pitch span. Level is the height of a speaker's F0, and span is the range of F0s used by an individual when speaking. As noted by Mennen and colleagues [5], there is no consensus on which is the best approach to quantify these two dimensions. In this case, we followed a method operationalized by Ordin & Mennen [3].

Regarding terminology, following the guidelines of the American Psychological Association [6], *sex* will be used to refer to the "individual biological status as female, male or intersexual" [6:2], whereas *gender* will be used to refer to "the attitudes, feelings and behaviours that a given culture associates with a person's biological sex" [6:2]. The term *bilingual* will be used to describe people who use two or more languages in their everyday lives [14], in line with similar research examining differences between individuals who speak more than one language [12; 13].

Existing research has reported that pitch varies greatly among individuals [3]. Such variation can be i) linguistic in nature or ii) linked to extra-linguistic factors. With regard to the population investigated here, previous research examining pitch in Japanese-English bilinguals does not provide a clear picture of the reasons for the differences reported between the two languages. Some researchers have attributed differences in pitch to linguistic factors [7; 8] and others to socio-cultural constraints linked to 'being a female' in the Japanese society [9; 10; 11]. In particular, Graham [7] reported that Japanese-American English simultaneous bilinguals showed a higher pitch level and wider span in Japanese compared to English, irrespective of whether they are female or male, in a reading task. On the other hand, Ohara [10] reported that only female Japanese-American English bilinguals showed differences in pitch range when reading Japanese compared to English. Moreover, Ohara reported the same language effect on female Japanese-English bilinguals in a free speech production task, i.e. in a voicemail task [11]. Interestingly, in [11], Ohara also found an effect of formality of the addressee (i.e. leaving the voicemail to a professor *versus* leaving it to a friend) only on the pitch range of her female participants. Specifically, the formal addressee (i.e. the professor) consistently elicited a higher pitch level that the informal addressee (i.e. the friend) and this was only valid when the female participants were speaking Japanese.

The present project investigated whether and how sequential Japanese-British English individuals manipulate their pitch in the case of a reading task in which formality and sex of the addressee were manipulated. In particular, we decided to investigate Japanese-English sequential bilinguals, i.e. individuals who acquired English L2 after the acquisition of Japanese [15]. The research questions we set out to answer were the following:

- Will the language spoken affect pitch range in female and male Japanese-English sequential bilinguals differently?
- Will formality and/or sex of the addressee affect pitch range of female and male Japanese-English sequential bilinguals differently?

10.21437/SpeechProsody.2018-78

2. Methods

2.1. Participants

Four participants, all residents of London (UK) at the time of the data collection, took part in the pilot study: 1 female Standard Southern British English (SBBE) functional [12] monolingual; 1 male SBBE functional monolingual; 1 female sequential Japanese-English bilingual and 1 male sequential Japanese-English bilingual. These individuals represented respectively the female control group, male control group, female experimental group, and male experimental group, for which data collection is ongoing. Prior to data collection, participants were asked to fill in a language background questionnaire adapted from LEAPQ [16] and MSI-Goldsmiths [17], enquiring about general information which could be a source of potential variation and confounds in this study. Some relevant data gathered from the background questionnaire are presented in Table 1.

Table 1: *Background information*

Participant (number, sex & nationality)	**Age**	**Years of English learning**	**Years of residency in English speaking countries**
1M (UK)	32	from birth	from birth (UK)
2F (UK)	24	from birth	from birth (UK)
3F (JP)	31	18	1 (UK)
4M (JP)	35	21	5 (USA) + 4 (UK)

2.2. Stimuli

Sentences from [7] were used for this project as they were considered to be favourable for pitch analysis, i.e. they contain a high amount of fully voiced segments to avoid F0 discontinuity associated with voiceless segments. Moreover, the sentences were short enough to typically correspond to a single intonational phrase (IP), enabling to control for declination differences. The study in [7] used 20 sentences, comprising four different types of sentence, namely:

- 5 Wh-questions (e.g. When will you be in Ealing? / なんでいいリングーにいる？)
- 5 declarative questions (e.g. You remembered Lil? /リルのことを思い出した？)
- 5 alternative questions (e.g. Did you say red or bed? / レッドとベッドのどっちを行ったの？)
- 5 declarative statements (e.g. We remembered Lil. / リルのことを思い出したリルのこと出した。)

A novel component was added to the elicitation procedure, namely addressees' photographs, to account for the effect of formality of the situation reported by Ohara [11]. Three types of addressee were chosen: (1) highly formal-looking addressees, i.e. members of the Royal family; (2) formal-looking addressees, i.e. business people (3) informal-looking addressees, i.e. students in school uniforms. In each condition, addressees of both sexes, i.e. females and males, were presented to the participants to investigate if the sex of the addressee had an effect on participants' pitch production. The sentences were judged to be informal in style by monolinguals of both languages, hence if there was a variation in pitch, this would have been due to the different levels of formality of the addressees and/or sex of the addressee.

The reading task presented in this paper is part of a longer experiment comprising 3 tasks and 3 short questionnaires: Task 1 was a perception task and Task 3 was a freer production task. One questionnaire (Opening questionnaire) preceded the tasks and two questionnaires (Gender questionnaire and Ending questionnaire) followed the tasks. Bilingual participants were asked to repeat the experiment twice in each of their languages, with a 30-minute break in between the two sections of the experiment to account for language modes [18]. The testing environment was in English, and the data collection always started in English: both factors are counterbalanced in the ongoing data collection. Only Task 2 will be presented in this paper.

2.3. Procedure

The study took place in a soundproof IAC booth in the Queen Mary University of London Phonetics Lab. The recording chain was a Neumann TML 103 large diaphragm condenser microphone (cardioid polar pattern) and an RME Fireface UC audio interface (microphone preamp and analogue to digital converter). All audio was recorded on an iMac at a sample rate of 44.1 kHz, 16-bit. Participants were comfortably seated at a computer and presented with a basic interface created in PsychoPy [19] that showed the stimuli along with the photograph of the addressee. Participants were instructed by an animated character created in Adobe Character Animator [20] to address each sentence to the person they saw in the picture. The instructor, "Blobby", was a gender-neutral animation (see Figure 1) which gave written directions to the participants via speech bubbles to avoid any potential influence in terms of (1) phonetic imitation [21; 22; 23] and/or (2) gender interaction [24].

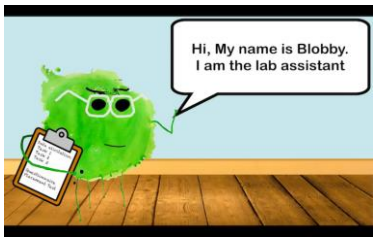

Figure 1: *Gender-neutral animation character, English version.*

For the workflow, each stimulus was presented only once, always with the same addressee. The order of the presentation of each combination of stimulus+addressee was randomized. Participants were instructed to read out loud without changing the content of the sentence in any way and were able to set their

own pace by pressing the space bar to move on to the next stimulus+addressee.

2.4. Analysis

Following the procedure from Ordin & Mennen [3], FFR measurements were made for the two aspects of pitch range described previously: level and span. The recordings were manually segmented in Praat [25]. Three measures were taken for level, namely F0 maximum, F0 minimum and mean F0, all reported in Hertz. F0 min and F0 max were extracted manually in Praat [25] for each utterance (i.e. 20 per task language), whereas mean F0 values were calculated using the Praat internal auto-correlation algorithm with 0.01 time step over each sentence. Pitch floor and ceiling were set at 75Hz and 300Hz for male participants and 100Hz and 500Hz for female participants, following Praat recommendations [25]. Pitch span, expressed in semitones (ST), was computed over each sentence, following [26], as:

$$\text{Pitch span} = 39.863 * \log(F0Max/F0Min) \quad (1)$$

3. Results

Figures 2 and 3 show the relationship between the language spoken by each participant and the formality of the addressee, for mean F0 and span separately. In all graphs, the English monolingual male is on the far left, followed by the English monolingual female; thereafter the Japanese-English bilingual female and male, firstly in English and then Japanese.

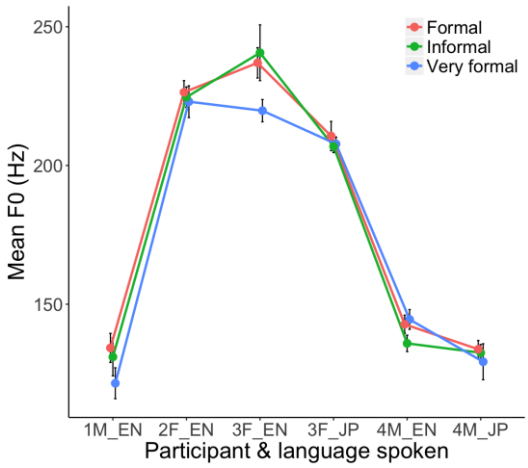

Figure 2: *Mean F0 (Hz) for each participant according to formality of the addressee. Error bars indicate standard error of the mean.*

With regard to mean F0, it is apparent in Figure 2 that there was little variation according to formality of the addressee in the English monolinguals. However, the Japanese-English female bilingual had a lower mean F0 in the very formal setting, whilst the male did not. Surprisingly, the mean F0 was lower in Japanese than English irrespective of the whether the participant was female or male.

The results were analysed with R [27] using a two-way Anova to compare the main effects of task language and formality, and the interaction of task language and formality on mean F0 and span, respectively. In all models, n=120, 20 tokens for each monolingual and 40 tokens for each bilingual were entered due to low participant numbers and to test the models for the future main data collection. For mean F0, the results showed only a significant effect of the task language [F (2, 5) = 261.5910, p<0.001]: Japanese evidenced a significantly lower mean F0.

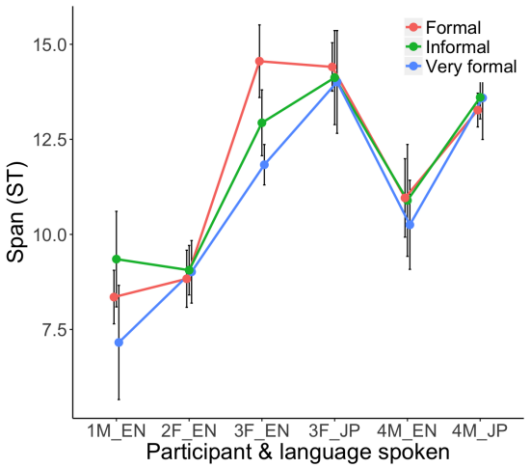

Figure 3: *Span (ST) for each participant according to formality of the addressee. Error bars indicate standard error of the mean.*

With regard to pitch span, it is apparent in Figure 3 that the monolingual male had a wider pitch span in the informal setting, which narrowed in the formal setting, and was narrowest in the very formal setting, similar to the monolingual female, although she displayed less span variation across different settings. Overall, one can see an effect of language spoken in the bilinguals: the span was wider in Japanese than in English; however, in the formal setting, the Japanese female's pitch span remained stable in Japanese and English. For span, the Anova yielded significant results only for task language [F (2, 5) = 18.4967, p<0.001].

Figures 4 and 5 show the relationship between the language spoken by each participant and the sex of the addressee, i.e. female *vs.* male addressee, for mean F0 and span separately.

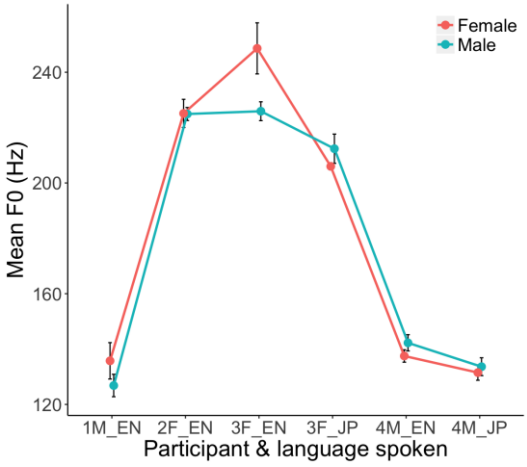

Figure 4: *Mean F0 (Hz) for each participant according to sex of addressee. Error bars indicate standard error of the mean.*

Similar to the previous plots, the overall effect of language spoken can be seen in Figure 4. Mean F0 of the English controls seemed to show a sex of the addressee-related pattern; when the addressee was female, the speakers' F0 was slightly higher than when the addressee was male. The Japanese female speaker replicated this pattern in her English. However, in Japanese, her F0 was slightly higher when addressing a male; this was also the case for the Japanese-English bilingual male in both Japanese and English.

Again, we used a two-way Anova to investigate the main effects of task language and sex of the addressee, and the interaction of task language and sex of the addressee on mean F0 and span, respectively. For mean F0, the Anova yielded a significant effect of task language [$F(1, 5) = 294.15$, $p<0.001$], and a significant interaction between task language and sex of the addressee [$F(1, 5) = 3.5158$, $p<0.01$].

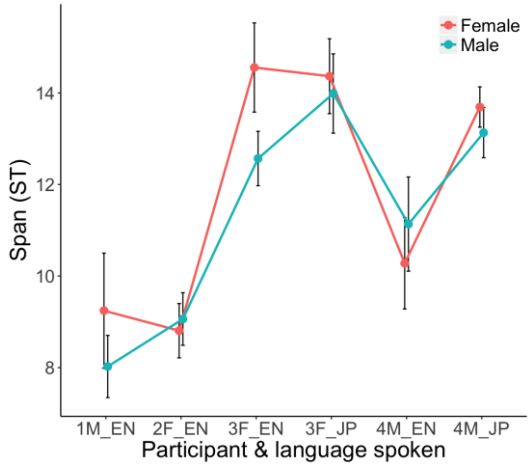

Figure 5: *Span (ST) for each participant according to sex of addressee. Error bars indicate standard error of the mean.*

With regard to span (Figure 5), overall one can see an effect of language spoken. In most cases span appeared to increase when addressing females, apart from the English female and the Japanese male in his English. Moreover, the span of the Japanese-English female increased considerably when addressing Japanese males compared to English males, but it remained roughly the same across languages when speech was addressed to females. The results of the Anova yielded a significant effect of task language [$F0(1, 5) = 19.125$, $p<0.001$].

The relationship between language spoken and type of sentence will not be presented here due to space limitations.

4. Conclusions

Although these are only preliminary results stemming from a very small pool of participants, we summarise some noteworthy observations here. Interestingly and unexpectedly, our results showed that Japanese speakers had a higher mean F0 when they spoke English than Japanese, irrespective of whether they are female or male. This is different from what has been reported in the literature, with one study reporting a higher mean F0 in Japanese for bilinguals of both sexes [7] and others reporting that only female Japanese speakers have a higher mean F0 in Japanese than English [9, 10, 11]. Span results are in accordance with Graham [7] for Japanese-English sequential bilinguals, who reported that they have a wider span in Japanese than English, irrespective of whether they are female or male. Due to the limited sample size, at present we are not able to draw any conclusions on the discrepancy between our results and the literature.

However, an interesting remark from the female bilingual during the debriefing could tentatively be used to explain her mean F0 results [i.e. English mean F0 = 235Hz; Japanese mean F0 = 208Hz]. When informed about the aim of the project, the participant stated that she "hates that female way of speaking in Japanese". This might be reflected in her Japanese results: her mean F0 in Japanese is not only lower than her mean F0 in English but also far from the peaks of 400Hz reported in the literature [9]. In her English, though, her mean F0 increases considerably, surpassing the mean F0 of the female control [mean F0 = 225Hz]. It might be that in English she applies the "female way of speaking" more so than in Japanese. Looking at Table 1, we can see that this speaker had been studying English for 18 years, but had been living in an English-speaking country only for 1 year at the time of testing. Hence, it might be that when speaking English, she is still unable to consciously repress the female prosodic feature which is characteristic of her native speech.

This speculation does not hold for the result of the male speaker, for whom a slightly different explanation might exist: Loveday [9] reports that Japanese male speakers, differently than English speakers, use a low mean F0 which in English implies 'being cool…rude' and even 'unpleasant' [9:83]. This participant had lived in the UK for 4 years, and previously in the USA for 5 years at the time of testing, so he potentially might have adjusted his mean F0 in English in order to not be perceived as 'cool and unpleasant' by native English speakers.

As per the effect of formality of the addressee and sex of the addressee on the pitch range of the participants, from this pilot it appears that sex of the addressee had a greater effect on pitch range than formality. Interestingly, the Japanese-English female participant seemed to clearly distinguish her pitch range, i.e. a higher F0 and wider span, in English when addressing females *vs.* males. Such a pattern did not hold in her Japanese, where a higher mean F0 was used to address males, similar to the results of the male Japanese-English bilingual. Both bilinguals showed a wider span when addressing females compared to males in Japanese; interestingly the span of the Japanese female participant while addressing males increased considerably in her Japanese compared to her English. Formality of the addressee showed a less stable pattern. This could be due to the fact that the addressees are shown in pictures, so potentially sex is more salient than formality.

These initial findings motivate further exploration of pitch range in bilingual speakers as an effect of societal factors in order to confirm whether the variation reported was due to language spoken and social factors, rather than speaker-specific characteristics.

5. Acknowledgements

We wish to thank the participants that kindly gave their time to carry out this pilot; Kimbal Bumstead for his help in creating the animation; Tomoko Taguchi for translating the experiment into Japanese and answering all of our questions about this fascinating language; and Hiroko Mori for answering all of our questions regarding formal and informal Japanese language. This work was funded by the ESRC.

6. References

[1] Mennen, I., & de Leeuw, E. (2014). Beyond Segments. *Studies in Second Language Acquisition*, *36*(02), 183–194.

[2] Deutsch, D., Le, J., Shen, J., & Henthorn, T. (2009). The pitch levels of female speech in two Chinese villages. *The Journal of the Acoustical Society of America*, *125*(5), EL208-EL213.

[3] Ordin, M., & Mennen, I. (2017). Cross-Linguistic Differences in Bilinguals' Fundamental Frequency Ranges. *Journal of Speech Language and Hearing Research*, *60*(6), 1493.

[4] Ladd, D. R. (2008). *Intonational Phonology* (2nd ed.). Cambridge: Cambridge University Press.

[5] Mennen, I., Schaeffler, F., & Docherty, G. (2008). A methodological study into the linguistic dimensions of pitch range differences between German and English. In *Proceedings of the 4th Conference on Speech Prosody, Campinas*.

[6] American Psychological Association. (2012). Guidelines for psychological practice with lesbian, gay, and bisexual clients. *American Psychologist*, *67*(1), 10–42.

[7] Graham, C. (2015). Fundamental Frequency Range in Japanese and English: The Case of Simultaneous Bilinguals. *Phonetica*, *71*(4), 271–295.

[8] Yamazawa, H., & Hollien, H. (1992). Speaking Fundamental Frequency Patterns of Japanese Women. *Phonetica*, *49*(2), 128–140.

[9] Loveday, L. (1981). Pitch, politeness and sexual role: An exploratory investigation into the pitch correlates of English and Japanese politeness formulae. *Language and Speech*, *24*(1), 71–89.

[10] Ohara, Y. (1992). Gender-dependent pitch levels: A comparative study in Japanese and English. In K. Hall, M. Bucholtz, & B. Moonwomon (Eds.), *Locating Power: Proceedings of the Second Berkeley Women and Language Conference, April 4 and 5, 1992*. Berkeley Women and Language Group, University of California, Berkeley.

[11] Ohara, Y. (1999). Performing gender through voice pitch: A cross-cultural analysis of Japanese and American English. In U. Pasero & F. Braun (Eds.), *Wahrnehmung und Herstellung von Geschlecht: Perceiving and Performing Gender* (pp. 105–116).

[12] De Leeuw, E., & Bogulski, C. A. (2016). Frequent L2 language use enhances executive control in bilinguals. *Bilingualism: Language and Cognition*, *19*(05), 907–913.

[13] Bogulski, C. A., Rakoczy, M., Goodman, M., & Bialystok, E. (2015). Executive control in fluent and lapsed bilinguals. *Bilingualism: Language and Cognition*, *18*(03), 561–567.

[14] Grosjean, F. (2010). *Bilingual: life and reality*. Cambridge, Mass: Harvard University Press.

[15] Li Wei (Ed.). (2009). *The bilingualism reader* (2. ed). London: Routledge.

[16] Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research: JSLHR*, *50*(4), 940–967.

[17] Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLOS ONE*, *9*(2), e89642.

[18] Grosjean, F. (2001). *Life with two languages: an introduction to bilingualism* (Nachdr.). Cambridge, Mass.: Harvard Univ. Press.

[19] Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13.

[20] The Adobe. (2017). Adobe Character Animator CC (Version Beta). Adobe Systems incorporated.

[21] Adank, P., Stewart, A. J., Connell, L., & Wood, J. (2013). Accent imitation positively affects language attitudes. *Frontiers in Psychology*, *4*.

[22] Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, *40*(1), 177–189.

[23] Giles, H., Coupland, J., & Coupland, N. (Eds.). (1991). Contexts of accommodation: developments in applied sociolinguistics. Cambridge [England] ; New York : Paris: Cambridge University Press ; Editions de la Maison des science de l'homme.

[24] Biemans, M. (1998). The Effect of Biological Gender (Sex) and Social Gender (Gender Identity) on Three Pitch Measures. *Linguistics in the Netherlands*, *15*, 41–52.

[25] Boersma, P., & Weenink, D. (2016). Praat: doing phonetics by computer. (Version 6.0.26). Retrieved from http://www.praat.org/

[26] Hewlett, N., & Beck, J. M. (2010). Phonation. In *An introduction to the science of phonetics*. London [u.a.]: Routledge.

[27] R Development Core Team (2008). R: A language and environment for statistical computing. Foundation for Statistical Computing, Vienna, Austria.