

Est.
1841

YORK
ST JOHN
UNIVERSITY

Muradás-Taylor, Becky ORCID

logoORCID: <https://orcid.org/0000-0001-7275-6016> (2017)

Japanese pitch accent in an English/Nupe/Hausa trilingual. In: Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017. ISMBS, pp. 199-212

Downloaded from: <https://ray.yorks.ac.uk/id/eprint/2727/>

The version presented here may differ from the published version or version of record. If you intend to cite from the work you are advised to consult the publisher's version:

<http://ismbs.eu/data/documents/ISMBS-2017-Proceedings.pdf>

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. [Institutional Repository Policy Statement](#)

RaY

Research at the University of York St John

For more information please contact RaY at ray@yorks.ac.uk

Japanese pitch accent in an English/Nupe/Hausa trilingual

Becky Muradás-Taylor
b.muradas-taylor@yorks.j.ac.uk

York St John University

Abstract. Pitch accent, which distinguishes words in Standard Japanese, is difficult for speakers of English to acquire. This is likely to be because pitch in English does not have lexical function. However, alternative explanations could be insufficient Standard Japanese input or lack of explicit instruction. This paper reports on an English/Nupe/Hausa trilingual learner of Japanese who uses Standard Japanese pitch accent accurately, in spite of no residence in Japan or explicit instruction on pitch accent. Nupe and Hausa are tonal i.e. have lexical pitch. The aim of the paper is to report on the accuracy and stability of the participant's pitch accent; to consider how their language background has aided this acquisition, and to discuss implications for monolingual English speaking learners of Japanese. The data consists of a three minute audio recording of a presentation given in Japanese. The participant produced 90% of words with accurate Standard Japanese pitch accent and 93% of repeated words with accurate stable pitch accent. The participant's successful acquisition of pitch accent is argued to be because of the presence of lexical pitch in Nupe and Hausa. Since they achieved this despite no explicit instruction, stay in Japan, or a native-speaker tutor, the difficulty monolingual English speakers have acquiring pitch accent cannot be easily dismissed due to lack of explicit instruction or input. This finding provides support for the argument that English speakers' difficulty acquiring Japanese pitch accent is due to pitch not being lexical in English. However, other bilinguals (English + tone language) are needed to strengthen this claim.

Keywords: L2 acquisition, L1 influence, input, instruction, suprasegmental, stress, tone

Introduction

This is a case study of a person with three L1s – English and the Nigerian languages Nupe and Hausa – who started studying Japanese as an additional language in the UK at the age of 30. The study investigates the participant's production of pitch accent in their spoken Japanese, focussing on its accuracy (i.e. adherence to Standard Japanese norms) and stability (i.e. the extent to which repeated words have the same accent type). The participant produces pitch accent that is both accurate and stable. This is of interest because previous research on the acquisition of pitch accent by monolingual speakers of English learning Japanese showed that pitch accent was neither accurate nor stable (Taylor, 2012). Since the trilingual participant has acquired Standard Japanese pitch accent despite not receiving explicit instruction on pitch accent nor living in Japan, the difficulty that monolingual English speakers have acquiring pitch accent cannot be attributed to insufficient input or lack of explicit instruction. Instead the difference between the trilingual participant and the monolingual English speakers in Taylor (2012) is argued to be due to the trilingual participant's two other L1s, Nupe and Hausa, which are both tonal, i.e. have lexical pitch.

Pitch in Japanese, Nupe, Hausa and English

Pitch in Japanese is used to distinguish words. For example in Standard Japanese *hashi* 'chopsticks' has initial accent and *hashi* 'bridge' has final accent. What distinguishes different accent types is the presence or absence of an accent (realised phonetically as a sharp fall in pitch) and its position (Vance, 2008). Initially-accented words have an accent on the first syllable; finally-accented words have an accent on the final syllable; unaccented words have no accent i.e. no sharp pitch fall. Words longer than two syllables can have an accent on the second, third, fourth, etc. syllable: for simplicity, they are often referred to as having 'medial' accent.

Pitch in Nupe and Hausa is also lexical. Unlike Japanese pitch accent, where the presence or absence of an accent and its position distinguish the accent types of words, Nupe and Hausa are tonal, with

syllables specified for tone. Nupe has five tones: high, mid, low, rise, fall; Hausa has three: high, low, fall (Yip, 2002).

Pitch in English, however, does not have lexical function. Instead, stressed syllables have high, low, falling or rising pitch (Pierrehumbert, 1980). The pitch shape differences depend on whether the utterance is, for example, a statement, question, or conveying surprise.

That pitch in English does not have lexical function is central to this study, but may go against the intuitions of users of English. Even in academic writing, pitch is sometimes described as an acoustic correlate of English stress. Fry (1958) is often quoted as evidence of this: for minimal pairs such as the noun *object* and the verb *object*, listeners used pitch to distinguish these words. However, calling pitch a correlate of stress is a ‘common misunderstanding’ (Beckman & Edwards, 1994:13) caused by a confound between lexical stress and utterance-level pitch (Sluijter & van Heuven, 1996a). Stressed syllables are potential ‘docking sites’ for utterance-level pitch (Sluijter and van Heuven, 1996b:2471); although listeners may use pitch to distinguish two words in citation form, it does not follow that pitch is a correlate of stress. Nor is pitch a reliable cue to stress: words that are not focussed do not have a pitch movement (Sluijter & van Heuven, 1996a, 1996b) and, for focussed words, the shape of the pitch movement depends on whether the utterance is a question, statement, conveying surprise, etc.

The following example illustrates the difference between pitch in Japanese (lexical function) and English (post-lexical function). If you take the word *happy* and produce it with rising intonation, you get the word *happy* as a question. However, if you take the initially-accented word *hashi* ‘chopsticks’ and produce it with rising pitch, you get a different word: finally accented *hashi* ‘bridge’. Since pitch has lexical function in Japanese, change the pitch of *hashi* and you change the word meaning. Since pitch does not have lexical function in English, change the pitch of *happy* (to falling pitch, rising pitch, falling-rising pitch, rising-falling pitch, etc.) and you will still have the same word.

Perception of Japanese pitch accent by L1 English speakers

The ability of L1 English-speaking L2 learners of Japanese to perceive Japanese pitch accent is task-dependent. In an AX discrimination task, where listeners judge whether two words have the same or different accent type, English-speaking learners perform similarly to L1 Japanese listeners (Hirano-Cook, 2011). English-speaking learners also perform similarly to L1 Japanese listeners on an ABX task, where listeners judge whether the word X has the same accent type as A or B (Sakamoto, 2011). On an identification task, however, where listeners identify which accent type a word has (e.g., initial, medial, final, unaccented), L1 Japanese speakers out-perform English-speaking learners (Hirano-Cook, 2011; Sakamoto, 2011). Similarly, unlike L1 Japanese speakers, English-speaking learners are unable to judge whether the accent types of words are correct (i.e. adhere to Standard Japanese norms; Shibata & Hurtig, 2008). Hirata (2015:736) concludes from findings such as these that L1 English-speaking L2 learners of Japanese can perceive pitch but do not categorise it like L1 Japanese speakers.

Interestingly, the correctness judgement task is difficult for all English-speaking learners, even advanced ones (Shibata & Hurtig, 2008). However, individual variation is observed on the identification task, with some learners performing well (Hirano-Cook, 2011). This may imply that English speaking learners of Japanese do not encode accent type into lexical representations in long-term memory, even if they learn to distinguish and identify them.

Production of Japanese pitch accent by L1 English speakers

Yamada (1994) reports on a learner who says the word *yappari* ‘obviously’ with three different accent types in one conversation. Such accent type instability is likely to be a result of pitch not having lexical function in English.

Taylor (2012) investigated pitch accent acquisition by English-speaking learners of Japanese, focussing on both accuracy and instability. This will be described in detail here, since it is the trilingual participant’s different behaviour to that of the participants in Taylor (2012) that is of interest in this study. The participants were L1 speakers of Standard Southern British English: a less-experienced group (n=13) who had studied Japanese for one or two years (mean 250 hours), and a

more-experienced group (n=8) who had completed four years of a Japanese degree including a year in Japan. The learners read aloud 180 Japanese words in three contexts:

- (i) In isolation e.g., *ame* 'rain'
- (ii) Before a function word e.g., *ame da* 'it's rain'
- (iii) Before a content word e.g., *ame ga furu* 'rain falls'

Japanese phoneticians identified the accent type of each word that the participants produced. The accuracy (i.e. percentage match with Standard Japanese) was 43% for the less-experienced group (lowest participant 32%, highest participant 52%, SD=6) and also 43% for the more-experienced group (lowest participant 36%, highest 48%, SD=5). The learners' accent types showed considerable instability: only 12% of the words produced by the less-experienced group, and 13% of the words produced by the more-experienced group had accent types that were produced accurately and stably across the three contexts.

The question that arises out of this research is why Japanese pitch accent is so difficult to acquire for English speakers. The most likely explanation is a linguistic one: because pitch is not lexical in English, L1 English-speakers do not encode pitch in their lexical representations when learning Japanese. This is supported by the perception findings described above: even if English-speaking learners learn to identify Standard Japanese accent types, they cannot say whether the accent types of words are correct. This implies that the learners' lexical knowledge does not contain information about pitch, even if accent types can be identified.

If pitch is not lexical in English, why do learners of Japanese not learn to encode it lexically on exposure to their L2? This could be because the functional load of pitch accent in Standard Japanese is low (Kitahara, 2001): pitch is rarely needed to disambiguate minimal pairs. Another possible reason is the considerable dialectal variation in Japanese, with words' accent types and even the accentual system varying between dialects (Kubozono, 2012). Together these could mean that L2 pitch accent is not necessary for learners and, therefore, not learned.

However, there could be other explanations. Perhaps the participants in Taylor (2012) had received insufficient Standard Japanese input (see e.g., Flege, 2009)? Or perhaps they had received insufficient explicit instruction (see e.g., Thomson and Derwing, 2015)? The second of these is particularly likely, since commonly used textbooks in the UK introduce pitch accent at the beginning, but do not refer to or practice it in later chapters (Shport, 2008).

Aims

The aims are: (i) to measure the accuracy and stability of the participant's pitch accent; (ii) to consider how the participant's language background has aided this acquisition; and (iii) to discuss implications for monolingual L1 English speaking learners such as those in Taylor (2012).

Method

Participant

The participant was raised in Nigeria and the UK. They (with 'they' being used as a singular gender neutral pronoun) were born in Nigeria, lived in the UK from age 1 to age 5, and Nigeria from age 6 to age 23. They were raised trilingually: English and the Nigerian languages Nupe and Hausa were all used in the home from birth. They were educated in English both in the UK and in Nigeria, and studied Hausa at school from age 9.

The participant started a beginner-level Japanese class at a UK university at the age of 30. The course was delivered by a British English tutor for two hours a week for 12 weeks. They then took the intermediate-level course with the same tutor which was three hours a week for a further 12 weeks. They did not receive any explicit pitch accent instruction on these courses. At the time that the data was collected, they had never been to Japan.

The key facts regarding the participant's language background are firstly that, as well as English, they have two further L1s, Nupe and Hausa, which are tonal and therefore have lexical pitch; and secondly, that the participant had not received explicit instruction on pitch accent and had not been to Japan.

However, the participant's language background is actually more complex: as well as their three L1s (English/Nupe/Hausa), they also learned Arabic from the age of 3 for Quranic recitation, and Hindi from the age of 6 through interaction with Hindi-speaking neighbours. They also spent time in Malaysia and India between the ages of 23 and 27. Additionally, before entering the beginner-level Japanese class, they had a long-standing interest in Japanese poetry, role-playing games and anime.

Data collection and analysis

The data was a three-minute audio recording of a presentation on a topic chosen by the participant. The presentation was part of an assessment; informed consent for it to be used for research was obtained subsequently.

The data was analysed by comparing the accent type that the learner used for each word against the Standard Japanese norm as described in the Shin Meikai accent dictionary (Kindaichi & Akinaga, 2001). This was used to give a percentage of words with accurate accent types. In addition, where words appeared more than once in the recording, their accent types were coded as the same or different, and a percentage of words with stable accent types calculated.

Unlike in Taylor (2012), where the participants read aloud a list of words and phrases, the first stage of analysis in this study involved deciding what to count as a word. The accent types of a 'content word + function word' unit were judged where possible, even if there was a pause before the function word. The second stage involved deciding whether to exclude any words due to their accents being deleted or compressed. In Standard Japanese, two accented words can be grouped together in one 'accentual' phrase, with the second accent deleted, or in one 'intonational' phrase, in which case the second accent is retained but compressed or 'downstepped' (Venditti, 2005:175). An attempt was made to exclude such contexts systematically but did not prove possible. Instead, words were included in the analysis, if it was possible to identify their accent type aurally.

In total, the accent types of 113 words were identified (tokens, not types, as the stability of the accent types was also of interest). Fifteen words appeared more than once.

Results

The participant produced 102 out of 113 words (i.e. 90%) with accurate Standard Japanese accent types. Figure 1 compares the accuracy of the trilingual participant (TRI) to the 21 monolingual English learners of Japanese in Taylor (2012), where LE is the less-experienced group (n=13) and ME is the more-experienced group (n=8).

From Figure 1, it is clear that the learners in Taylor (2012) all have low accuracy, with little variation between learners, and the trilingual participant has high accuracy.

In addition, of the 15 words that were uttered more than once during the recording, 14 (93%) had the same accent type on both occasions. This contrasts with the learners in Taylor (2012), where even the more-experienced learners only produced 13% with stable accurate accent types.

Although the focus of this paper is the accuracy and stability of the participant's pitch accent, it is worth pointing out that the eleven words whose accent type was not Standard varied in part of speech (noun, verb, etc.), syllable number and structure, Standard Japanese accent type, and produced accent type. No pattern (e.g., initially-accented words consistently being produced with final accent, for example) was observed.

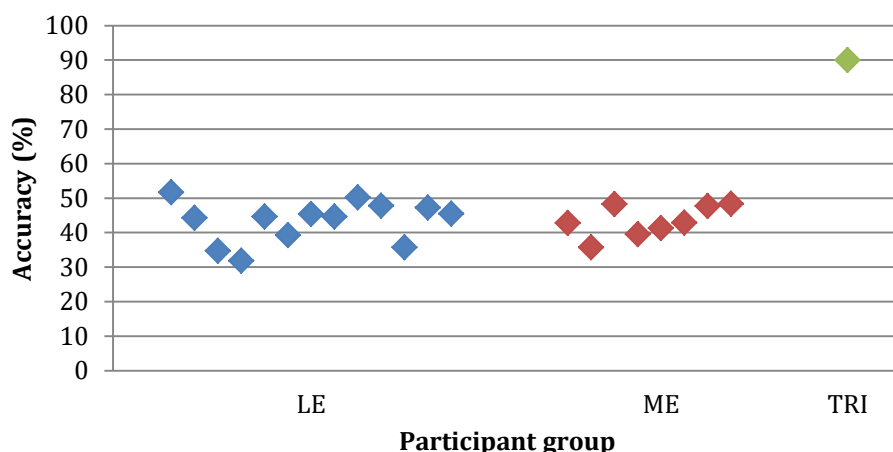


Figure 1. Accent type accuracy of the trilingual participant compared to the Taylor (2012) participants

Discussion

The participant's pitch accent was accurate and stable. 90% of words produced had accurate Standard Japanese accent types, and 93% of repeated words had accurate stable accent types. These findings are in striking contrast to the English monolingual speakers in Taylor (2012), where even the more-experienced group of learners only produced 43% of words with accurate Standard Japanese accent types and 13% with stable Standard Japanese accent types.

It was suggested earlier that the monolingual English speakers' difficulty acquiring Japanese pitch accent could be attributed to insufficient Standard Japanese input or lack of explicit instruction. However, the trilingual participant had only received 60 hours of Japanese language instruction over 24 weeks, had had a British English tutor, had had no explicit instruction on pitch accent, and had not lived in Japan. The participants in Taylor (2012) had received more Standard Japanese input (mean 250 hours for the less-experienced group; 970 hours for the more-experienced group) during their year in Japan. And neither the trilingual participant nor the Taylor (2012) participants had received explicit instruction on pitch accent. The difference between the trilingual participant and the monolingual English participants cannot therefore be attributed to these factors.

The most likely explanation for the difference between the trilingual participant and those in Taylor (2012) is the trilingual participant's two additional L1s with lexical pitch. These must have enabled the trilingual participant to encode Standard Japanese accent type into the lexical representation. Monolingual English speaking learners, however, must not encode pitch in the lexical representation, even if they are able to distinguish different accent types and identify them correctly.

Note that the trilingual participant produces Japanese pitch accent with high accuracy and stability, despite the fact that Japanese has pitch accent, with one or no syllables in a word specified as having a pitch fall, whereas Nupe and Hausa have tone with syllables specified as high, mid, low, rise, fall for Nupe, and high, low, fall for Hausa. Further research could investigate whether this difference has any effect on the participant's L2 Japanese.

A limitation of the current study is that the participant appears to be a particularly good language learner, who picked up Hindi from neighbours, for example. They (singular) also had considerable exposure to Japanese culture before starting formal study, casting doubt on the claim that their Standard Japanese input is limited. However, many of the monolingual English speakers in Taylor (2012) also expressed a keen interest in Japanese culture, and yet the highest accuracy of any learner was only 52%. This returns us to the original argument that is a linguistic factor (the presence or absence of a tone language L1) not amount of Standard Japanese input that explains the difference between the trilingual participant and the monolingual ones.

Conclusion

This paper reports on the pitch accent production of a learner of Japanese who has the tone languages Nupe and Hausa as L1s, as well as English. This participant's accurate and stable Standard Japanese pitch accent contrasts with monolingual L1 English-speaking learners of Japanese, whose pitch accent is neither accurate nor stable (Taylor, 2012). Since the trilingual participant had received no explicit instruction on pitch accent, less Japanese instruction than the Taylor (2012) participants, and had not lived in Japan, the monolingual participants' difficulty acquiring Standard Japanese pitch accent cannot be dismissed as due to lack of explicit instruction or insufficient Standard Japanese input.

Instead, the difference between the participants is argued to be due to the presence or absence of an L1 with lexical pitch. Since pitch in English has post-lexical function, monolingual English speaking learners of Japanese are not expected to encode pitch accent into long term memory. In contrast, the trilingual participant's tonal L1s Nupe and Hausa must enable the encoding of Japanese accent types into long term memory, resulting in stable and accurate Standard Japanese pitch accent production.

In order to strengthen this claim, other bilinguals (English + tone language) are needed. Further research could include acoustic analysis, as an aural impression of the current participant is that their accent types are not just phonologically accurate, but similar to Standard Japanese in phonetic realization, too.

References

- Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (ed.), *Phonological structure and phonetic form: Papers in laboratory phonology III* (pp. 7-33).
- Flege, J. (2009). Give input a chance! In T. Piske & M. Young-Scholten (eds), *Input Matters in SLA* (pp. 175-190).
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech*, 1(2), 126-151.
- Hirano-Cook, E. (2011). *Japanese pitch accent acquisition by learners of Japanese: Effects of training on Japanese accent instruction, perception, and production*. PhD Thesis, University of Kansas.
- Hirata, Y. (2015). L2 Phonetics and phonology. In H. Kubozono (ed.), *Handbook of Japanese phonetics and phonology* (pp. 719-762). Berlin: De Gruyter Mouton.
- Kindaichi, H., & Akinaga, K. (eds.) (2001). *Shin Meikai nihongo akusento jiten [New Meikai Japanese accent dictionary]*. Tokyo: Sanseido.
- Kitahara, M. (2001). *Category structure and function of pitch accent in Tokyo Japanese*. PhD Thesis, Indiana University.
- Kubozono, H. (2012). Varieties of pitch accent systems in Japanese. *Lingua*, 122(13), 1395-1414.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation* (PhD Thesis, Massachusetts Institute of Technology).
- Sakamoto, E. (2011). *Investigation of factors behind foreign accent in the L2 acquisition of Japanese lexical pitch accent by adult English speakers*. PhD Thesis, University of Edinburgh.
- Shibata, T., & Hurtig, R. R. (2008). Prosody acquisition by Japanese learners. In Z. Han & E. S. Park (eds.), *Understanding second language process* (pp. 176-204). Clevedon, UK: Multilingual Matters.
- Shport, I. A. (2008). Acquisition of Japanese pitch accent by American learners. In P. Heinrich & Y. Sugita (eds.), *Japanese as foreign language in the age of globalization* (pp. 165-187). München: Iudicium Verlag.
- Sluijter, A. M. C., & van Heuven, V. J. (1996a). Acoustic correlates of linguistic stress and accent in Dutch and American English. *Proceedings of the Fourth International Conference on Spoken Language Processing*, 2, 630-633.
- Sluijter, A. M. C., & van Heuven, V. J. (1996b). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100(4), 2471-2485.
- Taylor, R. L. (2012). *Eigo washa ni yoru nihongo no go akusento no shuutoku [The acquisition of Japanese lexical accent by English speakers]*. PhD Thesis, Nagoya University.
- Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36, 326-344.
- Yamada, N. (1994). Nihongo akusento shuutoku no ichidankai—gaikokujin gakushuusha no baai— [A stage in acquiring Japanese accent: the case of foreign learners]. *Journal of Japanese Language Teaching*, 83, 108-120.
- Vance, T. (2008). *The sounds of Japanese*. Cambridge, UK: Cambridge University Press.

- Venditti, J. J. (2005). The J_TOBI model of Japanese intonation. In S. Jun (ed.), *Prosodic typology and transcription: A unified approach* (pp. 172-200). Oxford, UK: Oxford University Press.
- Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.