NATIVENESS AND INTELLIGIBILITY OF JAPANESE ACCENTED ENGLISH CONSONANTS BY FRENCH LISTENERS

Rubén Pérez-Ramón[1], Mariko Kondo[1], Sylvain Detey[1], Lionel Fontan[2], Maelle Amand[3], Takeki Kamiyama[4]

[1]Waseda University, [2]Archean Labs, [3]CeReS Université de Limoges, [4]TransCrit Université Paris 8 Vincennes - Saint-Denis
rperez@aoni.waseda.jp, mkondo@waseda.jp, detey@waseda.jp, fontan.lionel@gmail.com, maelle.amand@unilim.fr, takeki.kamiyama@univ-paris8.fr

ABSTRACT

The presence of a Foreign Accent (FA) in the communicative process may interfere with the correct transmission of a message. This interference can arise from various traits, such as rhythm, intonation and speech rate. This can affect perception by both L1 and L2 listeners. Therefore, in this study we focused on the segments of Japanese FA in English as perceived by French listeners. Seven consonants ([l, ɹ, t^h, k^h, θ, f, v]) were extracted from the beginning of CVC English words. Using sound manipulation techniques, they were transformed into their most expected Japanese realisation ([ɾ, ɾ, t, k, s, ɸ, b] respectively) to generate a new set of words in which only the first segment was accented. The French cohort showed clear differences in perceived intelligibility of rhotic consonants and the [f/ɸ] contrast compared with previously analysed groups of Japanese and American listeners.

Keywords: perception, foreign accent, intelligibility, consonants

1. INTRODUCTION

There are many factors involved in second language (L2) phonological acquisition, such as native-likeness, fluency, intelligibility and comprehensibility. Acquiring "nativelike" pronunciation used to be considered the key issue in teaching or learning a foreign language, which meant that having a foreign accent was considered to be a failure of L2 acquisition. However, now the focus of L2 teaching and learning has shifted to intelligibility and comprehensibility [1]. So, while speech with a strong foreign accent may not always be easy to follow, studies such as [2] have shown that simply having a foreign accent does not necessarily hinder intelligibility or comprehensibility. In addition, since English has become a lingua franca, there are many varieties of English throughout the world, and the number of L2 English users has surpassed first language (L1) English users [3, 4]. This means that the majority of English users now have a foreign accent.

In terms of L2 perception, some studies have found that segmental accuracy is more important than suprasegmental accuracy in relation to correct word recognition, i.e., intelligibility [5, 6, 7], while other studies have shown that suprasegmental accuracy has a greater effect on native speakers' judgement of L2 speakers' fluency and nativeness of speech [8, 9].

Most assessments of learner pronunciation are done by native speakers of the target language. However, it is also important to consider how speakers of other languages perceive L2 pronunciation. Earlier studies [10, 11] compared assessments of Japanese accented English by assessors of 10 different L1s. It was found that assessors' L1 had a significant influence on their fluency assessment scores. However, more studies are needed to better understand how listeners' L1 affects perception of foreign accent.

Therefore, in this study, we examine the assessments of nativeness and intelligibility of Japanese-accented English – especially on the consonantal level – by native speakers of French and compare it with a prior assessment by native speakers of American English and Japanese.

2. JAPANESE AND FRENCH CONSONANTAL SYSTEMS

In the study, we focused on the following English consonants, [l, ɹ, t^h, k^h, θ, f, v], in the word initial position of a monosyllabic consonant-vowel-consonant (CVC) word. These segments were chosen because they are the most troublesome segments for native Japanese speakers in terms of their English production and perception [12, 13, 14, 15]. The important considerations regarding Japanese are as follows. First, Japanese has only one liquid consonant, /r/, and there is no phonemic distinction between /r/ and /l/. The most common realisation of /r/ is an alveolar tap [ɾ] [16, 17], although many other variants including [l] also occur [18, 19]. Japanese does not have the labio-dental fricatives /v/ and /f/. The most common substitutions are [b] for /v/ and bilabial fricative [ɸ] for /f/. Furthermore, Japanese lacks the voiceless inter-dental fricative /θ/ which is usually substituted with [s]. Finally, while the

plosives /t/ and /k/ are common consonants in Japanese, they are produced with less aspiration in all positions including the word-initial position that is the focus of this study.

Next, let us consider the French consonant inventory, and the main potential interferences with English [20] and Japanese [21, 22] consonants described in the literature. It is relatively similar to the English consonant system, but there are differences which may affect the evaluation of Japanese accented English by French speakers. For example, like Japanese, /θ/ is absent in French, and the most common substitute used by French speakers is [s], just like Japanese speakers [23, 12, 13]. While the other consonants examined in this study are found in French, the phonetic realisations can be quite different from English. For example, the French plosives /t/ and /k/ are unaspirated in word initial position, unlike their English equivalents, in word-initial position [24, 13]. In addition, while French has the voiceless labio-dental fricative /f/, the bilabial fricative [ɸ], which is a common Japanese substitution for /f/, does not occur in French. The French /v/ is similar to English /v/. With regard to lateral consonants, /l/ is always clear in present-day French ([l]). However, French /r/ is either a uvular fricative [ʁ] or slightly fronted velar [ɣ] in prevocalic position in most accent varieties in France, both of which are different from the English approximant [ɹ] which can be perceived by French speakers as [w] [25]. Therefore, such phonemic and phonetic differences between English, Japanese and French may influence the French speakers' judgement on Japanese accented English.

3. METHODOLOGY

3.1 Tokens

Seven target English consonants were selected according to their known impact on intelligibility when accented by Japanese speakers. The rationale for this selection is detailed in [26]. For each of these seven segments, three words were selected based on their structure and frequency, i.e., all of them had a consonant-vowel-consonant structure and ranked among the 1000 most frequent words in English according to the n-gram analysis of the Google corpus (Table 1).

A bilingual speaker was recruited based on her nativeness in both American English and Japanese, as judged by native speakers of each language [27]. In addition to the recorded English words, the speaker also recorded a set of non-words with a Japanese accent, one for each English word. The accented segment of each non-word was placed in the same (or similar) phonetic context to its equivalent English word. For example, for the English word *fish* [fɪʃ], the Japanese non-word *fishi* [ɸiɕi] was recorded, so [f] and [ɸ] would be in a similar context [_/i/]. Afterwards, by applying the splicing technique [28], a new token was generated in which only the initial consonant was pronounced with a Japanese accent, while the rest of the word remained with an unaltered English pronunciation (in the case of the example given above, this new word would be [ɸɪʃ]). Finally, the gradation technique [29] was applied, and, for each pair of generated tokens, a 7-step continuum was generated in which the accent of the target consonant varied gradually from fully accented to fully native. The final number of tokens was, therefore, 147 (7 consonants × 3 words × 7 steps).

Native segment	Foreign segment	English word	Japanese non-word
l	ɾ	leave	rifi
		league	rigi
		long	rogo
ɹ	ɾ	run	rana
		reach	richi
		roof	rufu
t^h	t	tag	taga
		teach	tichi
		took	tuku
k^h	k	keep	kipi
		come	kama
		call	koro
θ	s	thick	shiki
		thought	soto
		thumb	sama
f	ɸ	feet	fiti
		fish	fishi
		food	fudu
v	b	vet	bete
		van	bana
		ville	biri

Table 1: Target segments in English, Japanese segment for the accented version generation and selected English words for the experiment. Note that phonotactics in Japanese do not allow for a combination of [s]+[i], so the accented version of *thick* is closer to [ɕik].

3.2 Participants

Four cohorts were recruited for the experiment. A group of native American English participants were recruited using the Amazon Mechanical Turk tool (n=14). All of them were native speakers of English,

born in the United States and reported no contact with native speakers of any other language or from any other linguistic backgrounds. Two cohorts of Japanese participants, all members of the student population at Waseda University (Tokyo, Japan), also took part in the experiment: one group had high proficiency in English (n=25) and the other had low proficiency in English (n=26). These two groups were analysed separately because proficiency in a language has been demonstrated to have a significant effect on the perception of a foreign accent [29], and the Japanese cohort was the only one with enough heterogeneity to create two separate groups. Proficiency was self-reported and standardised according to the Common European Framework [30]. Finally, a group of French listeners was recruited from a body of university students in France (n=7). None of the non-native participants were bilingual in any language and had not lived for long periods of time in foreign countries. All subjects were paid for their participation.

3.3 Tasks

All participants underwent two tasks, one for intelligibility and one for nativeness, which were part of a bigger experiment.

In the intelligibility task, the participants were asked to listen to tokens and then type the word they heard on a screen. Once written, the screen instructed them to press a button with the text "Next", and the following token played automatically, leading to a self-paced task. They were asked to type always real English words.

The nativeness task consisted of choosing between two buttons in which the labels "Native" and "Foreign" could be read. Participants were asked to press the appropriate button based on their perception of every token. In this task, the word appeared written on the screen in order to avoid the Ganong effect [31], i.e., so listeners can decide whether a token that sounds like [bæn] is a good exemplar or not of the word *van* without ascribing it to the word *ban*. Therefore, this task was conducted last, so that any learning effect would not affect the intelligibility task.

For both tasks, stimuli were pseudo-randomised to ensure that no two tokens from the same target consonant were presented consecutively. The experiment was designed using the platform for online experimenting Gorilla [32], and participants completed all the tasks in their own house with their own equipment, following two requirements: to complete the tasks in a quiet setting and to use headphones.

4. RESULTS

Results for the Native American English (N) and the Japanese cohorts (HP for High Proficiency and LP for Low Proficiency) have been reported in [26]. In this study, we will focus on the information collected from the French cohort (F) and compare it with the other groups.

4.1 Preprocessing and statistical analysis

Before proceeding with the analysis of the results, answers deemed too fast (< 0.3s) or too slow (> 5s) were removed from the dataset. All statistical analysis was carried out with R [33]. General linear mixed models were calculated using the `lme4` package [34], and significant effects were measured using the *Anova* function of the `car` package [35]. Post-hoc pairwise comparisons were conducted using the `emmeans` package [36].

4.2 Nativeness results

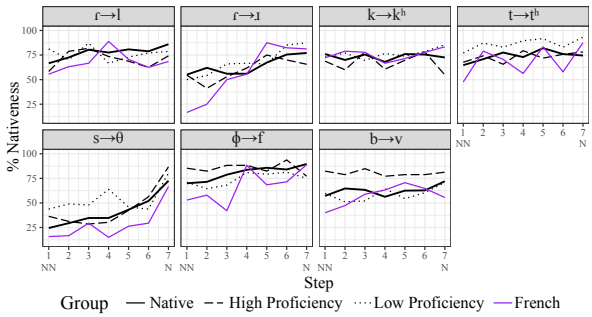

Figure 1: Results for the nativeness task for the four groups. Results are presented from the most accented end of the continuum (NN) to the most native-like token (N). Error bars are omitted for clarity.

A visual inspection of the results reveals some difference of the French cohort with the other groups. In [ɸ→f] and [ɾ→ɹ], a somewhat lower nativeness was reported, meaning that the accented realisation of these tokens was more harshly judged by the French cohort than the other groups. Additionally, in [s→θ], the French listeners seemed to perceive the change in category later than the other groups.

A general linear mixed model with CONSONANT and COHORT as fixed factors and PARTICIPANT as a random factor revealed a significant interaction between CONSONANT and COHORT ($p < .001$) as well as a significant effect of CONSONANT ($p < .001$). Averaged across steps, the French group showed a significant difference with the HP group for the [b→v] continuum ($p < .05$) and with the LP group for the [t→tʰ] and [s→θ] continua ($p < .05$). Marginal differences were found with the LP group for the [ɸ→f] continuum. Individual analysis carried out for

each continuum with p-values corrected for multiple comparisons revealed a significant difference with the HP group for the first 3 steps of the [ɸ→f] continuum ($p < .05$) and for step 4 of the [s→θ] and [t→t^h] continua ($p < .05$).

4.3 Intelligibility results

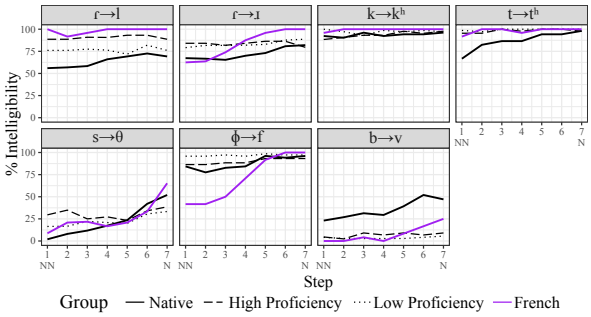

Figure 2: Results for the intelligibility task for the four groups. Results are presented from the most accented end of the continuum (NN) to the most native-like token (N). Error bars are omitted for clarity.

The replacement of [f] with [ɸ] seems to have affected the French listeners significantly more than any of the other three groups. There is also a clearly different pattern of the [ɾ→ɹ] continuum, which follows a trend similar to [ɸ→f], specifically, a sharp change in perceived category at the central part of the continuum (i.e., categorical perception).

As with the nativeness task, the results from the intelligibility task were used to generate a similar general linear mixed model. The model showed a significant interaction between sound and cohort ($p < .001$) and a significant effect of sound ($p < .001$). The analysis averaged across steps revealed significant differences of the French group with the native, high proficiency ($p < .05$) and low proficiency ($p < .001$) groups for the [ɸ→f] continuum. Similarly, for [ɾ→l], a significant difference was found between the French and the other three groups (N, LP: $p < .001$; HP: $p < .05$). The analysis of individual sounds only showed significant differences for the [ɸ→f] continuum, with the native cohort at the first step ($p < .05$) and with the other two cohorts at steps 1, 2 and 3 (HP: $p < .05$; LP: $p < .01$).

5. DISCUSSION AND CONCLUSION

The main goal of this study was to compare the perception of Japanese accented English consonant segments by French listeners with that of a native cohort (American English) and a matched non-native cohort (Japanese). For the latter, results from two sub-groups were collected, attending to their proficiency in English.

Previous research [26] analysed the impact of a foreign accent on perception by both native speakers of the target language and the accent of the speaker. In [37], the concept of matched interlanguage speech intelligibility benefit was coined to refer to the fact that listeners whose L1 is the same as the speaker report a less extreme drop in intelligibility. Moreover, the authors also suggested a mismatched interlanguage speech intelligibility benefit in the sense that listeners with an L1 that differs both from the target language and the accent of the speaker could experience a similar "benefit".

However, this mismatched interlanguage benefit in intelligibility (which is the case of the French cohort analysed in this study) has been disputed [38] or, at least, not found for other language pairs. In our study, we have found that this advantage in intelligibility is not only dependent on the language of the speaker, but also on the phonotactic strategy that leads to a certain transformation when speaking with a foreign accent.

For French listeners, the changes in [ɸ→f] and [ɾ→ɹ] involve unfamiliar (Japanese) units, which may explain the noticeable difference in the intelligibility patterns. In these cases, the French cohort perceived a more abrupt change in the category along the continuum, suggesting that a new segment conveys a bigger loss in intelligibility as the listener is not able to assign the proper category.

On the other hand, [s→θ] and [t→t^h] are changes that are also expected among French listeners (though the latter is seen more often in low proficiency speakers with little metaphonological awareness). Our results confirm that, when this is the case (i.e., when the transformation of accentedness is similar in two languages), intelligibility follows a similar pattern.

In conclusion, our study complements the classical theories of the mismatched interlanguage speech intelligibility benefit by providing detailed information about individual segments. By applying the segmental FA generation technique and the gradation technique, we can conclude that patterns of perception are not dependent on the L1 of the listener, but are probably related to the phonological transformations that arise in their mind at the time of speaking a non-native language.

7. REFERENCES

[1] Levis, J. 2020. Revisiting the intelligibility and nativeness principles. *J. of Second Language Pronunciation*, *6*(3), 310-328.

[2] Munro, M. J. & Derwing, T. M. (2020). Foreign accent, comprehensibility and intelligibility, redux. *J. Second Language Pronunciation*, *6*(3), 283-309.

[4] Crystal, D. 2003. *English as a Global Language*. CUP.

[5] Ethnologue Languages of the World, 2022. https://www.ethnologue.com/
[6] Fayer, J. M. & Krasinski, E. 1987. Native and nonnative judgments of intelligibility and irritation. *Language Learning, 37*, 313-326.
[7] Koster, C. J. & Koet, T. 1993. The evaluation of accent in the English of Dutchmen. *Language Learning,* 43, 69-92.
[8] Rognoni, L. & Busà, M. G. 2014. Testing the effects of segmental and suprasegmental phonetic cues in foreign accent rating: An experiment using prosody transplantation. *Proc. the International Symposium on the Acquisition of Second Language Speech, Concordia Working Papers in Applied Linguistics*, Vol. 5. 547-560.
[9] Anderson-Hsieh, J., Johnson, R. & Koehler, K. 1992. The relationship between native speaker judgments of non-native pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42(4), 529-555.
[10] Gut, U. 2003. Non-native speech rhythm in German. *Proc. 15th ICPhS,* Barcelona, 2437-2440.
[11] Kondo, M. & Konishi, T. 2017. Tsuujiru Eigo no tameno hatsuon kyouiku (What we need to teach for better communication in English)," *Conference Handbook of the 35th Conference of the English Linguistic Society of Japan*, Sendai, Japan, 232-237.
[12] Kondo, M., Fontan, L., Le Coz, M., Konishi, T. & Detey, S. 2020 "Phonetic fluency of Japanese learners of English: automatic vs native and non-native assessment", *Proc. Speech Prosody 2020,* 784-788.
[13] Yavaş, M. 2011. *Applied English Phonology (2nd ed.)*, Wiley-Blackwell.
[14] Collins, B. & Mees, I. M. 2013. *Practical Phonetics and Phonology - A resource book for Students* (*3rd ed.*). London: Routledge.
[15] Rogerson-Revell, P. 2011. *English Phonology and Pronunciation Teaching*. Continuum International Publishing Group.
[16] Kondo, M., Tsubaki, H. & Sagisaka, Y. "Segmental Variation of Japanese Speakers' English: Analysis of "the North Wind and the Sun" in AESOP Corpus", J. of Phonetic Society of Japan, 19(1), 3-17.
[17] Saitô, Y. 2006. *Nihongo Onseigaku Nyûmon* (*Introduction to Japanese Phonetics*) (2nd ed.). Sanseido.
[18] Vance, T. J. 2008. *The Sounds of Japanese*. CUP.
[19] Ohnishi, M. 1987. Dôteki jinkô kôgai ni yoru nihongo shiin no kôsatsu [A study of Japanese consonants by means of an artificial palate]. Onsei Gakkai Kaihô 186: 19–24.
[20] Arai, T. (1999), A case study of spontaneous speech in Japanese, *Proc.14th ICPhS*, San Francisco. 615-618.
[21] Tennant, J. 2016. Les anglophones. In S. Detey, I. Racine, Y. Kawaguchi & J. Eychenne (Eds.), *La pronunciation du français dans le monde : du natif à l'apprenant*. CLE international, 97-102.
[22] Kamiyama, T., Detey, S. & Kawaguchi, Y. 2016. Les japonophones. In S. Detey, I. Racine, Y. Kawaguchi & J. Eychenne (Eds.), *La pronunciation du français dans le monde : du natif à l'apprenant*. CLE international, 155-162.
[23] Detey, S., Fontan, L., Le Coz, M., Barcat, C. & Kawaguchi, Y. (in press). Identifying segmental substitutions in spontaneous speech of L3-French/L1-Japanese learners: a corpus-based pilot study. *J. of the Phonetic Society of Japan.*
[24] Tranel, B. 1987. *The Sounds of French*: An Introduction, CUP.
[25] Hallé, P., Best, C. T., Levitt, A. 1999. Phonetic vs. Phonological influences on French listeners' perception of American English approximants, J. *of Phonetics*, 27, 281-306.
[26] Kondo, M. & Pérez-Ramón, R, (in press) Perception of Japanese Accented English segments in Words, *J. of Second Language Studies.*
[27] Pérez-Ramón, R. & Kondo, M. 2021."Relationship between reaction time and nativeness rating", Proc. of 35th Annual Meetings of the Phonetic Society of Japan, 95-100.
[28] García Lecumberri, M. L., Barra-Chicote, B., Pérez-Ramón, R., Yamagishi, J. & Cooke, M. 2014. Generating segmental foreign accent. Proc. INTERSPEECH 2014, 1303-1306.
[29] Pérez-Ramón, Rubén and Kondo, Mariko (2021) "Relationship between reaction time and nativeness rating", The proceedings of the 35th Japan Phonetic Society Annual Meeting, 95-100, 25-26 September, 2021.
[30] Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
[31] Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of experimental psychology: Human perception and performance*, *6*(1), 110.
[32] Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior research methods*, *52*(1), 388-407.
[33] R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/
[34] Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.
[35] John Fox and Sanford Weisberg (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage.
[36] Russell V. Lenth (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.4.
[37] Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, *114*(3), 1600-1610.
[38] Stibbard, R. M., & Lee, J. I. (2006). Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis. *The Journal of the Acoustical Society of America*, *120*(1), 433-442.