INTERSPEECH 2019
September 15–19, 2019, Graz, Austria

Place Shift as an Autonomous Process: Evidence from Japanese Listeners

Yuriko Yokoe

Sophia University, Tokyo, Japan
yuriko.r.yokoe@gmail.com

Abstract

A perception experiment with Japanese listeners is conducted to investigate the nature of place shift phenomenon that was previously found with French and English listeners. Hallé et al. [1] showed that unattested consonant sequences /tl, dl/ are perceptually repaired to form grammatically acceptable consonant clusters /kl, gl/ in the listeners' native language.

In this study, a similar experiment with Japanese listeners, whose mother tongue lacks the onset clusters altogether, is conducted. The result explicitly shows that the place shift phenomenon ought not to be interpreted in relation to the top-down phonotactic feedback. Rather, I will argue that both labial and velar shift reflect an autonomous, signal-driven process. As such, language specificity in speech perception must reside in the listeners' cue weighting, rather than encoded linguistic knowledge.

Index Terms: misperception, unattested consonant clusters, language-specificity, cue weighting

1. Introduction

Tuning-in to the native language speech sounds is known to begin in the earliest stages of our ontogenetic development. Early infants are universal language users in that they are able to discern between any combination of human speech sounds, though this capability diminishes during the first year [2, 3, 4, 5]. In addition to prosodic and segmental attunement, phonotactic adjustment also comes into effect during the earliest years. Nine-month-old infants are capable of generalising across phonotactic regularities to enhance speech perception [6, 7]. The development of the sensitivity to phonotactic structures is not straightforward; some patterns are acquired earlier, some later [8]. However, it is a general understanding that it is during the latter half of the first year that they become aware of structural organisation of sound patterning [9].

Scholars as early as Trubetzkoy [10] were foresighted to point out this language specific nature of speech perception, coining the term 'phonological filter'. The influence of the mother tongue is so overwhelming that the incoming acoustic signal can be overridden, jeopardising the transmission of intended sounds. Specifically, as Polivanov [11] points out, this drastic force often overrules the faithful representation of syllable structures, allowing the Japanese speakers to insert a vowel between adjacent consonants — a sequence that is generally prohibited in Japanese. As a result, a French word université /ynivɛʀsite/ becomes /juniberusite/ in the Japanese adaptation [12].

Although there has been a large body of research on cross-linguistic segmental perception ([13, 14], among others), papers focusing on the way they pattern together as a segmental sequence are relatively few in number.

One of the earliest research to address the relationships between such phonotactic patterning and speech perception is Brown & Hildum [15] who found that the listeners' knowledge of sequential probabilities in the language can affect their perceptual identifications. Unattested word-initial consonant clusters are altered so as to form actual words, showing a strong tendency to distort the incoming signal to pursue lexicality.

In their experiment where both C1 and C2 were continua of /b-d/ and /r-l/, respectively, Massaro & Cohen [16] reconfirmed the participants' bias towards hearing attested clusters. The 'illegal' coupling of C1 and C2 (i.e., /dl/) was avoided: when a C2 was fixed at the /l/-endpoint, the C1 at the /d/-endpoint resulted in fewer correct identification, compared to that of the /b/-endpoint. This way, Massaro & Cohen explicitly showed that both C1 and C2 can be perceptually distorted to abide by the grammaticality in the hearers' native language.

All these research point to the fact that sequences that violate one's native phonotactics are perceptually altered to those that do not. It was in this context that Hallé et al. [1] discovered the velar shift phenomenon — when presented with word-initial /tl/ and /dl/, native speakers of French perceptually assimilated the clusters to /kl/ and /gl/, adding another result in line with the notion of 'phonological filter'. They replicated the velar shift in English listeners, too, but not in Hebrew listeners, whose mother tongue permits the coronal-lateral onset clusters [17, 18]. They concluded that the shift reflects the listeners' urge to repair the incoming signal so that it conforms to their native language phonotactics, supporting top-down perceptual feedback.

Inspired by Berent et al. [19, 20, 21] and Davidson [22], Matsumoto [23] and Matsumoto-Yokoe [24] conducted an open-response dictation task with monolingual Japanese and English listeners. The participants were presented with various combinations of Russian onset clusters (e.g., obstruent-obstruent /zb/, obstruent-sonorant /kn/, sonorant-sonorant /ml/, etc.). The result showed that even Japanese listeners, whose native language lacked onset CC sequences altogether, are biased to perceive certain clusters better than others. The lack of uniformity in their responses was interpreted in favour of Berent et al.'s view that humans are innately endowed with the universal knowledge about what is phonologically marked and what is not, thereby enabling the Japanese listeners to somehow discern between the statistically unlearnable contrasts. Among the stimuli, however, there was a pair of consonants which induced a particularly striking perceptual pattern: coronal-laterals. While the other clusters tended to receive vowel epenthesis, C1/C2 deletion, or metathesis, coronal-laterals deviated from the general tendency in that they induced the place shift of the first consonant. Moreover, although the

http://dx.doi.org/10.21437/Interspeech.2019-2302

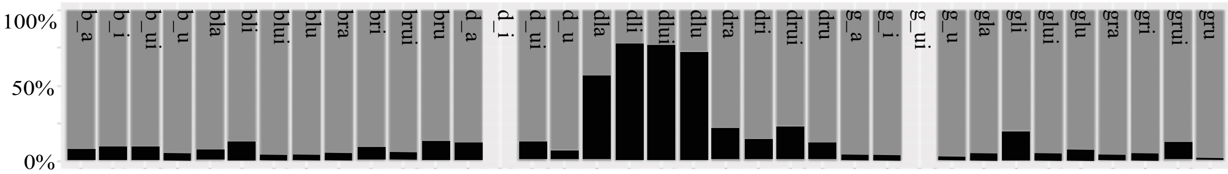

Figure 1: *Proportion of response accuracy to the Russian voiced stimuli by the Japanese listeners.*

proportion of correct identification reached approximately 50 per cent for the labial-lateral and velar-lateral stimuli, coronal-laterals received only 13 per cent (for /tl/) and two per cent (for /dl/) of correct place responses, respectively [24]. This pattern, obviously, could not be accounted for by the universal knowledge about sonority sequencing. In addition, there was even more striking pattern exhibited by the Japanese listeners: they changed the place of the coronal-lateral clusters to labials, not velars (67 per cent labials for /tl/, 51 per cent for /dl/).

Given the results of these preliminary investigations, this paper aims to examine the following questions. First, the reproducibility of the Japanese tendency to exhibit particularly low accuracy towards the coronal-laterals shall be tested. Once it is reconfirmed, the next question is whether the Japanese listeners' labial-oriented repair resurfaces. For convenience hereafter, I shall call this phenomenon the 'labial shift,' as opposed to the already proposed 'velar shift'. Third, in order to make this paper directly comparable to the preceding literature by Hallé et al. [1], I shall incorporate both Hebrew and Russian stimuli. Fourth, I shall attempt to unravel the perceptual mechanism behind these shift patterns.

2. Experiment with Japanese Listeners

A perception experiment with monolingual Japanese listeners was conducted.

2.1. Participants

Monolingual native speakers of Japanese (n=17) participated voluntarily with compensation. None of them had a history of learning Russian/Hebrew languages or residing in countries/regions where Russian/Hebrew is spoken. None reported hearing, speech or language problems.

2.2. Procedure

The experiment was run in a very quiet environment, using a laptop computer with Eprime software (version 2.0.10.356). The selected Russian (n=204) and Hebrew (n=96) tokens were combined to form a block of 300 stimuli. The experiment consisted of a short block with ten practice stimuli and three blocks of 300 stimuli, within each the nonwords were presented in a random order. The listeners were asked to respond as fast as possible by selecting one of the three buttons on their response pad (Cedrus RB-840), which were labelled "p/b", "t/d", or "k/g."

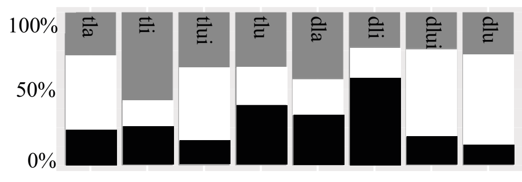

Figure 2(a): *Place identification responses to the Russian coronal-lateral stimuli by the Japanese listeners.*

2.3. Stimulus materials

Three female native speakers of Russian recorded seven repetitions each of 68 nonwords (six C1 conditions (/p, b, t, d, k, g/), three C2 conditions (/l, r/ or none), and four vowel conditions (/a, i, ui, u/)). Sequences /kui, gui/ were omitted from the recording list, for the Russian phonotactics does not allow the continuum of the velar consonants and the central vowel. Sequences /ti, di/ were also omitted in consideration of the prominent palatalisation of the coronal consonants before the high front vowel. The nonwords were embedded in 28 carrier sentences (seven each for four vowel conditions), which were then randomly prompted one by one on a computer screen. The recording was conducted in a soundproof room with a SONY PCM-D50 recorder and a SONY ECM-MS957 microphone, at a sampling rate of 44.1kHz /16 bit. The tokens were then spliced out from the carrier sentences and annotated using Praat for Mac, version 6.0.25 [25].

All the Hebrew stimuli were identical to those presented in Hallé & Best's [18] experiment. I received the sound files from the first author via a link to an online archive on December 11, 2015. There were four C1 conditions (/t, d, k, g/), two C2 conditions (/l, □/), and three vowel conditions (/a, i, u/), resulting in 24 nonwords, which in turn had four tokens each. All 96 tokens were used in the current experiment.

2.4. Result

Figure 1 shows the overall accuracy of the Japanese listeners in identifying the C1 plosive place of the Russian voiced stimuli (grey = correct, black = incorrect). It is evident from this figure that, despite the fact that their mother tongue lacks onset clusters entirely, their performance plummet for the coronal-laterals. Due to limitations of space, I cannot provide details on the voiceless Russian stimuli and the Hebrew stimuli, though the response accuracy pattern was the same: they performed more or less well on the clusters except for coronal-laterals. Figure 2 shows the Japanese listeners' place identification responses to (a) Russian and (b) Hebrew coronal-laterals (white = labial, grey = coronal, black = velar). The disparity between the Russian and Hebrew stimuli is striking: while there is a large amount of 'labial' responses to the Russian stimuli, the trend weakens considerably in the Hebrew nonwords.

Fisher's exact test with Monte Carlo simulation method (with 10,000 trials) was conducted using R version 3.4.3 on a Mac to see if there was any significant difference between the responses to Russian and Hebrew target nonwords. As for

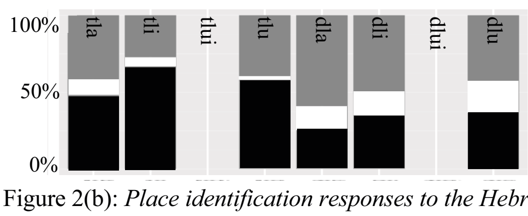

Figure 2(b): *Place identification responses to the Hebrew coronal-lateral stimuli by the Japanese listeners.*

Russian /tla/ and Hebrew /tla/, Japanese participants selected the labial button significantly more often than the velar one for the Russian nonword ($p < 2.3e\text{-}16$). Conversely, velar responses were much more frequent for the Hebrew nonword ($p < 1.8e\text{-}06$). For Russian /tli/ and Hebrew /tli/, labial responses for the former nonword were significantly more often ($p < 0.01$) than for the latter. The number of velar responses was greater for the Hebrew nonword ($p < 5.5e\text{-}15$). As to Russian /tlu/ and Hebrew /tlu/, Japanese listeners exhibited more labial and less velar responses to the former nonword (labial, $p < 1e\text{-}09$; velar, $p < 0.001$).

Comparison between Russian /dla/ and Hebrew /dla/ did not yield significance; still, the tendency for the labial shift in the Russian stimuli is numerically evident. /dli/ was the only nonword to which Japanese listeners gave more velar responses ($p < 6.1e\text{-}06$) to the Russian stimuli — a phenomenon that goes against the overall tendency. Lastly, for Russian /dlu/ and Hebrew /dlu/, there were more labial responses for the former stimulus ($p < 2.1e\text{-}14$). Velar responses were more often seen for the Hebrew stimuli ($p < 1.6e\text{-}07$).

Figure 3 shows the results of Bayesian statistical comparisons between the reaction time (RT) to shifted and non-shifted responses by the Japanese listeners (error bars indicate 95% Bayesian credibility intervals). The shifted responses here include instances where (1) Russian /tl, dl/ were labelled as /pl, bl/, (2) Russian /tl, dl/ were labelled as /kl, gl/, and (3) Hebrew /tl, dl/ were labelled as /kl, gl/. The non-shifted responses include (4) Russian /pl, bl/ correctly identified as /pl, bl/, (5) Russian /kl, gl/ correctly identified as /kl, gl/, and (6) Hebrew /kl, gl/ correctly identified as /kl, gl/. 'DeltaRusLab' refers to the comparisons between (1) and (4). 'DeltaRusVel' refers to that of (2) and (5), and 'DeltaHebVel' (3)(6). Positive posterior distribution indicates that the reaction time to the shifted responses was significantly longer than that of non-shifted responses.

For the Japanese listeners, correct identification of Russian /pl, bl/ requires less time than shifting the place of Russian /tl, dl/ to /pl, bl/. Likewise, correct identification of Russian and Hebrew /kl, gl/ is faster than shifting the coronal place to velar. From these comparisons, it is evident that the place shift requires extra processing time than normal identification responses.

3. Discussion

The current result explicitly shows that it is not language-specific phonotactics that drives the behaviour of the listeners. Replicating Matsumoto-Yokoe [24], Japanese listeners again showed a robust tendency to misperceive only the coronal-lateral clusters, amongst a set of uniformly ill-formed plosive-liquid onsets. This result is eloquent of the fact that the place shift phenomenon does not stem from phonotactic assimilation, negating a view repeatedly supported by Hallé and colleagues, as well as a basic assumption underlying Breen et al.'s [26] EEG experiment.

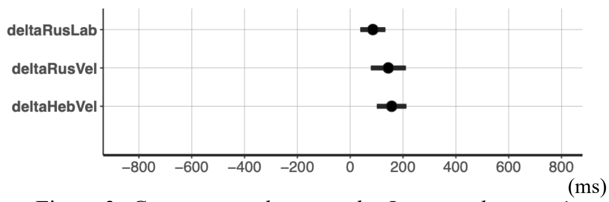

Figure 3: *Comparisons between the Japanese listeners' reaction time to shifted and non-shifted responses.*

In Yokoe [27] a perception experiment with English listeners revealed that, in accordance with the current result, English listeners, too, exhibit the disparate perceptual pattern between Russian and Hebrew stimuli: Russian nonwords induce more labial responses, and Hebrew more velars. Yokoe [27] proposed to interpret these place shift phenomena as stemming from language-specific cue weighting, rather than phonotactics. In other words, the place shift should be understood as an autonomous, signal-driven effect, not as a product of top-down feedback.

Longer RT in the autonomous view is an indicator of ambiguous signals, of which processing takes extra time. As such, one must ask what the definition of an ambiguous signal might be. There must be at least two dimensions that we have to take into account in answering the question. One is the acoustic specifications of the signal itself. These include, among others, the burst properties, segment duration, and formant patterns. The other is the sensitivity of the listeners to those acoustic specifications. Every listener must have his or her own cue-weighting strategy, which has presumably been formed by interacting with his or her native language. A corollary is that it is unacceptable to assume a genuinely robust, or genuinely ambiguous, signal; we must always evaluate the signal robustness/ambiguity in relation to the listeners' perceptual sensitivity.

Figure 4(a) shows mean RT (x axis) and the proportion of place identification (y axis) of Russian coronal-lateral stimuli. The dots represent each stimulus material (token), and a linear model was fit to represent each scatter. Figure 4(b) shows the same plot for Hebrew stimuli.

Readers should note that in these plots, the coronal line represents the accurate identification of coronal, and the other two lines represent errors. This means that the coronal line and the other two lines should exhibit an opposite tilt, if they were to stay consistent. For Russian stimuli, the coronal line for the Japanese listeners starts from approximately 50% of accurate responses at around 800 ms RT, sharply sloping down to reach less than 10% of accuracy at around 1500 ms RT. The labial, but not velar, line is consistent with this move, suggesting that the Japanese listeners wavered in their judgment as to select the coronal response or the labial response. I would interpret these patterns as an evidence that for the Japanese listeners, Russian coronal-lateral stimuli were equivocal between labials and

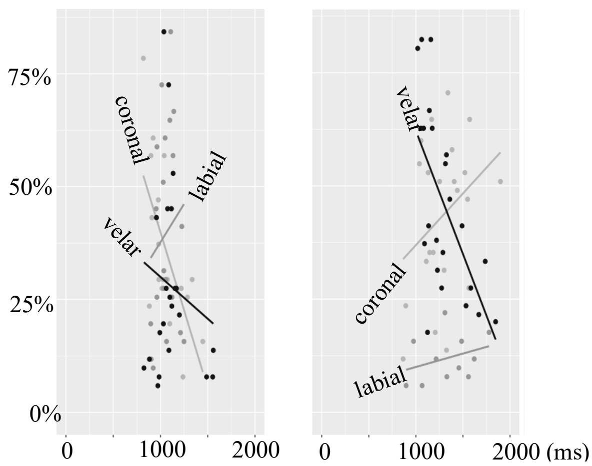

Figure 4(a, left) and 4(b, right): *Correlation between the Japanese listeners' proportion of place identification and RT, to the (a) Russian and (b) Hebrew stimuli.*

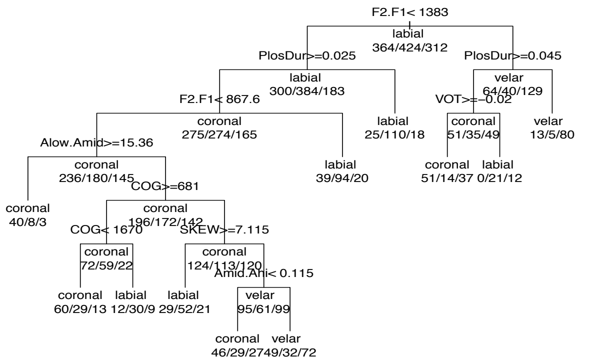

Figure 5(a): *A Decision Tree for the place identification of Russian coronal-lateral stimuli by the Japanese listeners.*

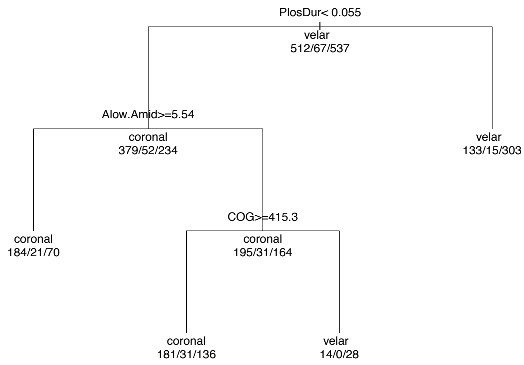

Figure 5(b): *A Decision Tree for the place identification of Hebrew coronal-lateral stimuli by the Japanese listeners.*

coronals, thus resulting in more labial responses than Hebrew stimuli.

The completely opposite tendency can be found in Figure 4(b) for Hebrew stimuli. Here, it is the velar line that is consistent with the coronal line. For the Japanese listeners, Hebrew coronal-initial stimuli were ambiguous between coronals and velars, thus inducing more velar responses than Russian stimuli.

To investigate the Japanese listeners' cue weighting, decision tree analyses were conducted, using R package rpart [28]. The dependent variable was the place identification responses: labial, coronal, or velar. Independent variables were (1) plosive duration, (2) proportion of plosive duration to stimulus duration, (3) voice onset time, (4) centre of gravity (COG) calculated from the spectrum of the first 25ms of the burst, (5) standard deviation (SD) of the spectrum, (6) skewness of the spectrum, (7) kurtosis of the spectrum, (8) Alow-Amid, (9) Amid-Ahi, (10) F2-F1, (11) VowelHL: 'high' includes /i, ui, u/ and 'low' represents /a/, (12) VowelRND: 'rounded' represents /u/ and the rest is 'unrounded.' Of these, Alow, Amid and Ahi represent the mean amplitude of the spectrum, divided into three sub-regions: 0-1,250Hz (low), 1,250-3,000 (mid), and 3,000-5,000 (high). Alow-Amid is retrieved by subtracting the mean amplitude of the mid region from that of the low region, enabling us to have a coarse image of the general shape of the spectrum over the lower half of the bandwidth. Amid-Ahi does the same over the upper half of the bandwidth, and this has been used as one of the criteria to assess acoustic differences between coronal-lateral and velar-lateral clusters in Flemming ([29], though in his study Ahi represented the mean amplitude of the spectrum between 3.5kHz-8kHz, following [30]). F2-F1, following Cooper et al. [31], is calculated by subtracting the frequency of the first formant from that of the second formant, both at the steady state of the lateral approximant, a segment that directly follows the plosive.

Figure 5(a) shows the decision tree for the place identification of Russian coronal-lateral stimuli by Japanese listeners. First, the main branch spreads in terms of F2-F1, where the most of the /i/-final nonwords (F2-F1 difference greater than 1383Hz) are separated from the rest. The /i/-final stimuli are then divided according to plosive duration. If longer than 45ms, they are identified as velars; if shorter, voiceless nonwords are correctly identified as coronals, while voiced ones are labelled labials. Non-/i/-final stimuli are further divided according to plosive duration. If shorter than 25ms, they become labials. Longer ones are inspected again in terms of F2-F1 difference, but this time with the threshold at 867.6Hz. Those that have a larger gap between the formants are identified as labials. The rest of the tree makes use of burst-related criteria, such as Alow-Amid, COG, SKEW, and Amid-Ahi. In a nutshell, Japanese listeners depend primarily on formant-related measures, followed by durational cues, then by burst characteristics.

Figure 5(b) shows the decision tree for the place identification of Hebrew coronal-lateral stimuli by Japanese listeners. Here, the nonwords are first divided according to plosive duration. Longer ones (>55ms) fall into the category of velars, while shorter ones are further split in terms of Alow-Amid, where the nonwords with 'grave' spectra are identified correctly as coronals. The rest, in turn, is distinguished by COG. If a stimulus has its COG below 415.3Hz, it is labelled velar. Those with higher COG are interpreted as coronals. Note that the actual number of the labial responses was not large enough (67, compared to that of coronals (512) and velars (537)) to secure its position in this tree. To summarise, in identifying Hebrew plosive consonants, Japanese listeners first utilised the durational cue, then the burst-related criteria. This interpretation is seemingly inconsistent with the conclusion drawn from the identification of the Russian stimuli by the same population, where the formant characteristics topped the cue weighting. Yet, it has been revealed that the Hebrew stimuli, compared with the Russian ones, lacked formant variety. Specifically, for the Hebrew stimuli, it was impossible to use the formant differences as a measure to discern between /i/-final nonwords from the others. Japanese listeners, who primarily rely on formant characteristics whenever possible, therefore could not utilise this information in identifying Hebrew plosives, and were compelled to begin their decision-making process from the second criterion, the durational aspect.

4. Conclusion

In this study I have shown that even Japanese listeners exhibit particular difficulties in perceiving Hebrew and Russian coronal-laterals, and that they show the 'labial shift.' The place shift phenomenon of the coronal-lateral clusters does not stem from language-specific phonotactic feedback. I have also argued for the non-veridical nature of auditory representation by demonstrating that language-specificity in speech perception must reside within the cue-weighting strategies, rather than the encoded linguistic knowledge.

5. Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP18H05605. This manuscript is based on [27].

6. References

[1] Hallé, P. A., Segui, J., Frauenfelder, U., & Meunier, C. "Processing of illegal consonant clusters: A case of perceptual assimilation?". *Journal of experimental psychology: Human perception and performance*, *24*(2), 592. 1998.

[2] Eimas, P. D. "Developmental aspects of speech perception" In *Perception*. pp. 357-374. Springer, Berlin, Heidelberg. 1978.

[3] Morse, P. A. "Infant speech perception: Origins, processes and alpha centauri". In *Communicative and cognitive abilities: Early behavioral assessment*. University Park Press Baltimore, MD. 1978.

[4] Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. "Developmental aspects of cross-language speech perception" *Child development*, 349-355. 1981.

[5] Werker, J. F., & Tees, R. C. "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life". *Infant behavior and development*, *7*(1), 49-63. 1984.

[6] Chambers, K. E., Onishi, K. H., & Fisher, C. "Infants learn phonotactic regularities from brief auditory experience". *Cognition*, 87(2), B69-B77. 2003.

[7] Friederici, A. D., & Wessels, J. M. "Phonotactic knowledge of word boundaries and its use in infant speech perception". P*erception & psychophysics*, 54(3), 287-295. 1993.

[8] Zamuner, T. S. "Sensitivity to word-final phonotactics in 9-to 16-month-old infants". *Infancy*, 10(1), 77-95. 2006.

[9] Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. "Infants′ sensitivity to phonotactic patterns in the native language". *Journal of Memory and Language,* 33(5), 630-645. 1994.

[10] Trubetzkoy, N. S. *Principles of phonology*. 1939.

[11] Polivanov, E. "La perception des sons d'une langue étrangère", *Travaux du Cercle linguistique de Prague*, *4*, 79-96. 1931.

[12] Shinohara, S. "Emergence of universal grammar in foreign word adaptations". *Constraints in phonological acquisition*, pp.292-320. 2004.

[13] Akahane-Yamada, R., Tohkura, Y. I., Bradlow, A. R., & Pisoni, D. B. "Does training in speech perception modify speech production?". In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP'96* (Vol. 2, pp. 606-609). IEEE. 1996.

[14] Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. "Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese/r/and English/l/and/r/". *Journal of Phonetics*, *32*(2), 233-250. 2004.

[15] Brown, R. W., & Hildum, D. C. "Expectancy and the perception of syllables". *Language*, 32(3), 411-419. 1956.

[16] Massaro, D. W., & Cohen, M. M. "Phonological context in speech perception". *Perception & psychophysics*, 34(4), 338-348. 1983.

[17] Hallé, P. A., Best, C. T., & Bachrach, A. "Perception of /dl/ and /tl/ clusters: A cross-linguistic perceptual study with French and Israeli listeners". In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2893-2896). 2003.

[18] Hallé, P. A., & Best, C. T. "Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters". *The Journal of the Acoustical Society of America*, *121*(5), 2899-2914. 2007.

[19] Berent, I., Lennertz, T., Jun, J., Moreno, M. A., & Smolensky, P. "Language universals in human brains." *Proceedings of the National Academy of Sciences*, 105(14), 5321-5325. 2008.

[20] Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. "What we know about what we have never heard: Evidence from perceptual illusions." *Cognition*, 104(3), 591-630. 2007.

[21] Berent, I., Lennertz, T., Smolensky, P., & Vaknin-Nusbaum, V. "Listeners' knowledge of phonological universals: Evidence from nasal clusters". *Phonology*, 26(1), 75-108. 2009.

[22] Davidson, L. "Phonology, phonetics, or frequency: Influences on the production of non-native sequences". *Journal of Phonetics*, 34(1), 104-137. 2006.

[23] Matsumoto, Y. *Factors in non-native speech perception: typological markedness and linguistic experience*. M.A. thesis. Sophia University. Tokyo. 2012.

[24] Matsumoto-Yokoe, Y. Perception of Plosive-lateral Clusters: A Summary of Experiments. *Phonological Studies*, (17), 43-50. 2014.

[25] Boersma, P., & Weenink, D. Praat: doing phonetics by computer. [Computer program]. Version 6.0.25. Retrieved from http://www.praat.org/. 2017.

[26] Breen, M., Kingston, J., & Sanders, L. D. "Perceptual representations of phonotactically illegal syllables". *Attention, Perception, & Psychophysics*, *75*(1), 101-120. 2013.

[27] Yokoe, Y. *Native and Non-native Speech Perception: the Locus of Language-specificity in the Place Identification of Unattested Sequences.* PhD dissertation, Sophia University, Tokyo. 2019 (to be published).

[28] Therneau, T., Atkinson, B., Ripley, B., & Ripley, M. B. Package 'rpart'. cran.ma.ic.ac.uk/web/packages/rpart/rpart.pdf. 2018.

[29] Flemming, E. "Stop place contrasts before liquids", In *Proceedings of the 16th International Congress on Phonetic Sciences*, pp.233-236. Saarbrücken. 2007.

[30] Suchato, A., & Punyabukkana, P. "Factors in classification of stop consonant place of articulation", In *Ninth European Conference on Speech Communication and Technology*. 2005.

[31] Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., & Gerstman, L. J. "Some experiments on the perception of synthetic speech sounds", *The Journal of the Acoustical Society of America*, 24(6), pp.597-606. 1952.