Journal of Education and Educational Research
ISSN: 2957-9465 | Vol. 10, No. 3, 2024

A Literature Review of Phonetic Transfer

Wenbo Zhao

School of International Studies, Hangzhou Normal University, China

Abstract: Phonetic transfer stands out as a vital part of language transfer in second language acquisition. This paper renders a pithy literature review apropos of current studies concerning the phenomenon of phonetic transfer that is arguably contemplated from three phonetic transfer error pertinent aspects: segment, phonology and prosody for the purpose of gaining new insight for the future research.

Keywords: Phonetic Transfer; Segment; Phoneme; Suprasegmentals.

1. Introduction

The phenomenon of phonetic transfer has long been a pivotal point in the study of second language phonetics. Second language learners display a tendency to be infected with the phonetics of their native language, resulting in a particular foreign brogue when speaking the second language. This enables listeners to effortlessly discern the speaker's nationality. For instance, the term Chinese English often refers to systematic phonetic changes that occur when Chinese language phonetic features permeate the English spoken by Chinese learners. These changes are readily noticeable to native English speakers.

It has been fathomed out in early researches (Weinreich,1953;Brière,1966;Tarone,1980; Broselow,1984) that foreign accents appear to be germane to five common types of phonetic transfer errors: 1) phonetic substitution, such as superseding the Chinese vowel /i/ by the English vowel /ɪ/; 2) non-correspondence of phonemes due to disparities in phonological structure, for example, the arbitrary and erroneous application of the English lateral /l/'s two allophone (clear [l] and dark [ɬ]) to the Chinese /l/; 3) diversified interpretations of phonetic distinctions, such as treating the English tense/lax vowels /i, ɪ/ as long/short vowels /i:, ɪ/; 4) interference in syllable structure and phonotactic patterns, i.e., altering the syllable structure or phonetic combinations of the second language to match those of the native language, such as pronouncing English consonant clusters like /st/ as similar to the Chinese si te; 5) scrambling from prosody, such as using Chinese tones to supplant English intonation.

The aforementioned phonetic transfer errors pertain to three phonetics-related aspects: segments, phonology, and prosody. Segments refer to the specific realization of speech sounds in flow of speech. Phonology reflects the correlations among phonemes. Each language boasts a specific set of phonemes whose configuration and production feature distinctive rules. For example, the English lateral /l/ is pronounced as a dark [ɬ] when it appears at the end of a syllable, while the Chinese counterpart cannot emerge at the end of a syllable and does not have this dark pronunciation. Ergo the phonological rules of different languages and the associated syllable structures are not identical. Prosody refers to the supra-segmental features that are superimposed on different phonetic units such as segments, syllables, words, phrases, and sentences.

In response to the phonetic transfer phenomenon at the segmental level, Flege (1995) proposed the Speech Learning Model, while Best (1995) introduced the Perceptual Assimilation Model. These two models have exerted a profound impact, and a multitude of studies have sprung around them (Jiang & Liu, 2014; Li, 2019). Theoretical models developed for the phonological level include Brown's (1998) Feature-based Theory of Perception and Archibald's (2005) Redeployment Hypothesis. As for the prosodic level, there are models such as So & Best's (2014) Perceptual Assimilation Model for Suprasegmentals and Altmann & Vogel's (2002) Stress Typology Model. These models have been buttressed by experimental data from their respective proponents and numerous opportune studies (Altmann, 2006; Jiang & Liu,2014; So & Best,2014; Holliday, 2015).

2. Current Research

Phonetic transfer errors involve three distinct levels: segmental, phonological, and prosodic, mirroring the differences between the native language and the second language at these three levels. At the segmental level, phonetic transfer alludes to segmental substitution, without considering the sway of phonology and prosody. According to Flege (1987), second language learners, through comparison, categorize second language and native language segments into three categories: identical, similar, and new categories. Segments in the identical category have no differences in production or perception between the second language and the corresponding native language segments. In this case, native language segments in the identical category can replace second language segments. This substitution is positive transfer, as the replacement does not induce any negative effects. More commonly encountered are segments in the similar category, where segments in the native and second languages fall into the same category, even though they can be represented by the same International Phonetic Alphabet symbol. However, differences in production and perception are still perceptible. If native language segments in the similar category took the place of second language segments, it could occasion negative transfer effects, rendering the pronunciation of the second language segments inaccurate. New category refers to second language segments that have no corresponding segments in the native language and are thus completely unfamiliar. These new segments are generally easy to perceive, but due to the strangeness of articulatory mechanisms, second language learners may not be able to enunciate these segments properly. The

phonological level chiefly refers to the acquisition of a certain phonological feature in second language segments and the transfer related to phonemes and allophones. Here, phoneme is an abstract set composed of one or more surface phonological features. Regarding phonological features, errors in the aspects of voicing, length, and tenseness/laxness are quite common. For example, the substitution of the English tense/lax vowels /i, ɪ/ with long/short vowels /iː, ɪ/ illustrates that phonetic transfer occurs not only at the segmental level but also entails the phonological level. In de facto research, the segmental and phonological acquisitions are often intertwined and difficult to separate, so many studies generally do not strictly distinguish between them.

The situation related to phonemes is often engaged with the acquisition of syllables formed by segmental combinations or consonant clusters. Learners are confronted with at least three strands of this issue. Firstly, they must discern between phonemic variants in terms of phonological significance and the different pronunciations of the same phoneme in different contexts at the segmental level. As mentioned earlier, English is destitute of phonemic length contrast, but at the segmental level, English vowels exhibit length contrast and are related to the voicing of the ensuing consonant, such as the greater length of the /i/ before a voiced consonant /d/ compared to the /i/ before a voiceless consonant /t/ by approximately 50% (Major, 2001). In both cases, if second language learners cannot accurately enunciate the different lengths of vowels, even if their articulation is accurate, a foreign accent may still transpire.

In addition, there is the issue of controlling the co-articulation between segments, which pertains to how different articulatory gestures match with each other at various stages. Articulatory gestures refer to the articulatory movements formed by different speech organs and constitute the primary focus of articulatory phonology proposed by Browman & Goldstein (1989). Articulatory phonology theory advocates for directly linking different articulatory gestures and their intrinsic spatiotemporal relationships to speech sounds, supplanting the conventional phonological units with these gestures. Articulatory gestures have both spatial and temporal dimensions, determining how the articulatory organs of adjacent segments coordinate in time and space, and the various stages of articulatory gestures manifest different spatiotemporal coordination patterns across different languages. If learners cannot master these co-articulatory patterns, difficulties in acquisition may appear. For instance, Yang & Chen (2019) found that Chinese English learners sometimes perceive the Russian trill [r] as the Mandarin Chinese lateral [l]. This is due to the fact that although the articulatory gestures for the two sounds differ in the position of narrowness in the oral cavity, the difference in the degree of narrowness is minimal, leading to an insufficient contrast between the two sounds and thus affecting the learners' ability to distinguish them. The position and degree of narrowness are two articulatory features of interest in articulatory phonology.

Lastly, there is the issue of phonotactic constraints, which pertains to how different segments combine to form a licit syllable type in the language. Faced with alien second language syllable types, learners tend to modify them to conform to familiar types by dint of the insertion, deletion, or replacement of segments. This phenomenon is particularly conspicuous when learners produce syllable-final consonants and consonant clusters (Eckman, 1991).

Prosody entails an array of suprasegmental characteristics such as word stress, tone, pitch accent, intonation, speech rhythm, as well as pauses between words and sentences. Word stress refers to the prominence of one syllable over another within a word, which is manifested with the agency of auditory cues such as pitch, loudness, duration, vowel reduction. The corresponding acoustic features embrace rudimentary frequency of sonic waves, intensity, duration, and formant patterns (Fry, 1958). Pitch accent implies one mora being more prominent in pitch than another within a word. A mora is a prosodic unit analogous to a syllable and situates between a single syllable and segment, primarily found in Japanese, where a mora can be a single vowel, the first consonant in a geminate consonant, or a syllable-final nasal (Ishikawa, 2002). Pitch accent can be employed to distinguish meanings, as in Japanese, where some homophones have different meanings depending on the placement of the pitch accent. Tone and intonation refer to variations in pitch at the single-syllable and sentence levels, with tone to distinguish meanings, as in Mandarin Chinese, where nearly every character i.e. monosyllabic word carries a tone. Differences in tone with the same syllable usually indicate different characters and meanings. Therefore, learners of Chinese characters must simultaneously master their tones, in so much as mispronunciation can trigger misunderstanding. Intonation, on the other hand, is capitalized upon to deliver distinctions in types, syntax, discourse, semantics, affects, and attitudes between phrases or sentences. For example, in English, a wh-question ends with a falling intonation, whereas a yes-no question with a rising intonation. Additionally, pauses between words or sentences, the pronunciation lengthening or strengthening of a word or syllable at a particular position within a sentence, can also serve functions parallel to pitch variations in achieving grammatical, affective, and other communicative purposes.

Rhythm is a perceptual pattern formed in virtue of the regular occurrence of certain speech units within an utterance, represented through timing. Depending on the language, an utterance can exhibit regularity in syllable length or stress intervals. In accordance to research by Abercrombie (1967), Ladefoged (1975), and Pike (1945), the rhythmic patterns of languages around the world can be classified into three sorts: syllable-timed, stress-timed, and mora-timed. Archetypical syllable-timed languages such as Spanish can produce equally long and stressed syllables in succession, giving the impression of a machine gun. In contrast, emblematic stress-timed languages such as English feature alternating patterns of long and short stressed or unstressed syllables, resembling Morse code (Ramus et al., 1999). Meanwhile, typical mora-timed languages such as Japanese tend to generate mora-based equivalently long speech sound. These three kinds of languages are unitarily timed by dint of monosyllable, stressed syllable and mono-mora respectively. Each type is associated with specific phonological structures (Dauer, 1983), with stress-timed languages often having more complex syllable structures, allowing for consonant clusters and prevalent vowel reduction in unstressed syllables. In contrast, syllable-timed mora-timed languages typically have more simplex syllable structures, generally disallowing consonant clusters and showing less distinction between stressed and unstressed syllables, as well as less frequent vowel reduction.

3. Segmental Transfer

Segments are the basic units for learners to perceive and produce speech during the process of second language acquisition; therefore, segmental transfer is a primary object of research in phonological acquisition. According to Barrios et al. (2016), there are two prevailing viewpoints regarding the causes of segmental transfer in second language acquisition. The first standpoint posits that spatial similarity is the principal cause of segmental transfer, meaning that the positions of native and second language segments in a continuous multidimensional physical space must be close. This physical space can be defined by perceptual parameters, articulatory positions, or acoustic parameters. The second stand argues that resemblance in phonological features underlies transfer. Segments are a collection of multiple dispersed features, and the similarity between two segments can be determined by the quantity of shared and distinctive features within the dispersed feature set. The more shared features, the more similar the segments; the more distinctive features, the more dissimilar the segments.

The first point of view focuses on the overall similarity of segments in continuous space, while the second emphasizes the similarity of segments at the discrete level of phonetic features. Representing the first perspective of spatial similarity is Flege's Speech Learning Model (1995) and Best's Perceptual Assimilation Model (1995). Both models hinge upon the similarity of segments, with the former suggesting that the consistency of the division of phonetic categories between the native and second languages directly influences the exactitude of second language phonetic acquisition, and the latter proposing that the degree of correspondence between the categories of second language contrastive segments and native language segments is the sine qua non for differentiating them. Both theoretical models take the similarity of phonetic categories between the native and second languages in perception as their starting point, suggesting that initial perceptual errors lead to the misclassification of second language phonetic categories as native language phonetic categories, ultimately resulting in transfer.

Chang (2015) alleges that determining the similarity between two phonemes requires deliberation from at least three levels: phonetic level, which can be objectively measured acoustically; allophonic level, which concerns whether contrastive phonemes in different languages can form similar allophonic variants in different phonetic environments; and phonemic level, which examines whether different phonemes in different languages share similar phonetic contexts and distributional patterns. For instance, Chang et al. (2011) assert that, acoustically, the Mandarin Chinese vowel /u/ and the English vowel /u/ are not identical; the English /u/ is acoustically closer to the Mandarin Chinese /y/ (the pronunciation of the pinyin "ü"), but they share the same phonemic features, both being high back rounded vowels and occurring in parallel phonetic environments. This study found that highly proficient Mandarin Chinese learners whose native language is English have better acquisition of /y/ but are swayed by their native language in acquiring /u/. Chang (2015) ascribes this chiefly to the phonemic consistency of /u/ in both languages, leading to negative transfer. Chang also discovered that more experienced second language learners are more susceptible to the impact of segmental similarity at the phonological level.

Empirical studies on segmental transfer have shown an intimate correlation between the difficulty of segment acquisition and the occurrence of segment transfer (Flege, 1987; Eckman et al., 2003; Bayarataroğlu, 2011). Among these, transfer resulting from segmental contrasts is the most common, especially when two contrasting segments in the second language both exist in the native language where antithesis turns out to be absent. Eckman & Iverson (2013) found that Korean learners of English often fail to discriminate elementary English word pairs such as "see" and "she" because the contrasting phonemes /s/ and /ʃ/ in English, albeit present in Korean, are merely two allophones [s] and [ʃ] of the Korean phoneme /s/. When followed by the high front vowel /i/ or the glide /j/,[s] will be changed into [ʃ]. Consequently, from the perspective of Korean learners of English, these two English phonemes are not in antithesis, eliciting confusion. This example illustrates that transfer at the allophonic level poses a barrier to acquisition.

Research on the segmental acquisition of second language by Chinese learners of English can be broadly categorized into two situations: segmental substitution caused by the phonetics per se and the one induced by the phonetic environment. For example, Jiang & Liu's (2014) study on Chinese learners' perception and production of English words containing th falls into the first category. The pertinent results indicate that Chinese learners tend to confuse these two English phonemes with the Mandarin Chinese /s/ and /d/. This signals that Chinese learners have not established new categories for these phonemes, possibly due to the onerousness of grasping the interdental articulation required (the positioning of tongue between the upper and lower teeth), while the corresponding Mandarin Chinese phonemes /s/ and /d/ bear some phonetic resemblance to them and are deemed ideal substitutes.

Li's (2008) study on the acquisition of English vowels followed by nasals by Chinese learners of English falls into the second case. The researcher found that Chinese learners often pronounce the English word down as Mandarin Chinese dàng, where the alveolar nasal /n/ in down sounds like the velar nasal /ŋ/. Since Mandarin Chinese also has the velar nasal, it seems not justifiable that Chinese learners are incapable of producing velar nasal. Rather, considering the impact of the phonetic environment, the combination of the diphthong /au/ and the alveolar nasal /n/ in English is more akin to that of the monophthong /a/ and the velar nasal /ŋ/ in Mandarin Chinese. From a co-articulation perspective, this is due to an inertia-induced carry-over effect during the movement of the tongue from a low front position to a high back position (i.e., from /a/ to /u/), turning the alveolar nasal into the velar nasal. Notwithstanding, there exist more complications.

4. Phonological Transfer

The phonological features, as constituent elements of phonemes, are intimately correlated to phonetic transfer. Phonological transfer often occurs when one or more phonological features of a phoneme change, resulting in the transformation of that phoneme into another. In the relevant theories, the most representative is the feature-based perceptual theory (Brown, 1998), whose cardinal argument is that the grammatical features of the learner's native language constrain the accurate perception of non-native contrastive phonemes. Providing that a certain distinctive feature present in the second language is absent from the learner's native

language phonetics repertoire, the learner will fail to perceive it, and consequently will be unable to acquire the contrastive phoneme containing that feature in the second language.

Empirical research on phonological feature transfer is essentially represented by consonants, such as studies on the acquisition of voice onset time (VOT) in plosives. VOT can occur before or after the closure of the articulatory organs, with the former referred to as VOT lead and the latter as VOT lag, manifesting as aspiration. Flege (1987) established pivotal hypotheses in speech learning models through an investigation of the onset times in the production of French /t/ and English /t/. Flege noticed that bilinguals, especially novices, were incapable of distinguishing between the VOT lag in French /t/ and English /t/ in their production, for they pigeonhole the former as equivalent to the latter. As learners' level in the second language ameliorated, they gradually became able to discriminate difference, but still could not behave proficiently as native speakers. They even failed to produce their native language /t/ correctly, erroneously adjusting it under the sway of the second language, thus sparking a gradual fusion of the two /t/ sounds.

Subsequent studies on plosives have largely been based on Flege's (1987) findings and have probed into different languages from varied perspectives. For instance, Holliday's (2015) long-term tracking study of Chinese learners' acquisition of Korean plosives is one such example. In Chinese Mandarin, plosives only have a contrast between aspirated and unaspirated pair. For example, in the pinyin representations of the plosives /kh, k/ in Mandarin Chinese, the former clearly has a VOT lag, i.e., a larger VOT value, while the latter theoretically has no VOT lag, i.e., a VOT value of zero (although there is actually a brief, imperceptible lag that can be disregarded). On the other hand, Korean plosives such as /k', k, kh/ have a contrast between strong, weak, and aspirated sounds, corresponding to short, medium-long, and long VOT lags. That is, Korean plosives have a contrast in the length of aspiration, and the vowel pitch following a weak plosive is lower than that following the other two plosives, which is a secondary distinctive feature of the three plosives. Due to the similarity between strong and weak plosives and the unaspirated plosives in Mandarin Chinese, according to speech learning models, learners are likely to categorize them into the same category. According to the perceptual assimilation model, Chinese learners tend to assimilate both strong and weak plosives into the unaspirated plosives in Mandarin Chinese and assimilate the aspirated plosives in Korean into the aspirated plosives in Mandarin. In line with the feature-based perceptual theory or the redeployment hypothesis, the temporal distinctive feature of VOT in Korean does not exist in Mandarin, and consequently, the distinction between the three Korean plosives is not easily discerned by Chinese learners.

Another phonological feature closely related to consonants is voicing. Cebrian (2002) studied the acquisition of the voiceless/voiced contrast of coda consonants in English syllable by native speakers of Catalan. The researcher predicted that the devoicing rule is easier to acquire than the voicing rule, as voiced consonants in coda position are more marked than voiceless ones. In many languages, the former occurs less frequently than the latter. Languages with voiced coda consonants always have voiceless coda consonants, and most voiced coda consonants in English undergo devoicing. The results confirmed the transfer of two phonological rules, and the scope of transfer was as predicted, with the output before pauses and voiceless consonants changing from voiced to voiceless.

5. Prosodic Transfer

The most instrumental model in second language prosodic perception is the Suprasegmental Perceptual Assimilation Model proposed by So & Best (2014) based on tone perceptual assimilation phenomena. This theory extends Best's (1995) perceptual assimilation model from the segmental level to the suprasegmental level. It presumes that learners will assimilate the prosodic types of the second language, such as tone type, pitch accent type, intonation type, and rhythmic timing type, to the corresponding prosodic types in their native language. The suprasegmental perceptual assimilation model purveys a theoretical basis for the transfer of various suprasegmental features.

So & Best (2014) established a suprasegmental perceptual assimilation model based on empirical research, investigating how two groups of Mandarin Chinese learners with English and French as their native languages respectively, associate the intonation of their native languages with the tones in Mandarin speech. The researchers suggested that the differences observed were due to the different stress patterns in the two languages: English has word stress while French does not. This led to English-speaking learners in the study needing to pay attention to both the target word stress and the intonation of the entire sentence, whereas French-speaking learners only needed to focus on intonation, allowing them to allocate more attention to discerning subtle pitch differences between different tones.

In early research on second language prosody, intonation acquisition was not a focal point and has only recently garnered attention. Discussions on the acquisition of second language intonation, stress, and tone primarily involve the pitch patterns of the prosodic system, while another key element is the duration pattern. Research on the duration patterns chiefly homes in on learners' mastery of different languages' rhythmic types. For example, in Japanese, the pronunciation of the kanji 本 /hon/ consists of two mora, a vowel and a nasal sound, while Chinese learners often pronounce it as the similar-sounding Chinese syllable /hoŋ/ due to their unfamiliarity with mora as a timing unit and their tendency to pronounce based on the number of syllables, as is done in Chinese. This example illustrates that even with highly similar phonetic combinations, different timing types lead to different pronunciation outcomes.

6. Conclusion

From research conducted over the past few decades, it is evident that the study of phonetic transfer has undergone two stages: the initial phase focused on segmental and phonological transfer, and the recent phase has centered on prosodic transfer. The early research primarily concentrated on phonetic transfer resulting from the similarity of segmental categories, while the recent focus initially revolved around the acquisition of second language stress, followed by a significant emphasis on the prosodic features of non-stress-timed languages, such as the acquisition of Mandarin Chinese tones and other second language intonation patterns.

Furthermore, as the scope of research continues to expand, new patterns of second language perception and production are constantly being discovered, all of which urgently require new theoretical frameworks for support. It is worth noting that,

although there is a plethora of research on Chinese second language learners with different dialectal backgrounds, most of it has essentially applied Western theories of second language acquisition, inevitably leading to a mismatch between theory and practice. Therefore, it is particularly imperative to construct models of second language perception and production that are suitable for the nature of Chinese in order to explain the phonological transfer phenomena observed in Chinese learners.

References

[1] Abercrombie, D. (1967). Elements of General Phonetics. Edinburgh: Edinburgh University Press.

[2] Altmann, H. (2006). The Perception and Production of Second Language Stress: A Cross-linguistic Experimental Study. Ph. D. Dissertation. Newark : University of Delaware.

[3] Altmann, H. & l. Vogel. (2002). L2 acquisition of stress: The role of L1. Paper presented at the DGfS Annual Meeting "Multilingualism Today ", Mannheim, Germany, March 2002.

[4] Archibald, J. (2005). Second language phonology as redeployment of L1 phonological knowledge. *Canadian Journal of Linguistics* 50(1):285-314.

[5] Barrios, S., N. Jiang & W. J. Idsardi. (2016). Similarity in L2 phonology: Evidence from L1 Spanish late-learners' perception and lexical representation of English vowel contrasts. *Second Language Research* 32(3):367-395.

[6] Bayarataroğlu, S. (2011). A model of classification of phonemic and phonetic negative transfer: The case of Turkish-English interlanguage with pedagogical applications. *Journal of Language and Linguistic Studies* 7(1):113-168.

[7] Best, C. T. (1995). A direct realist view of cross language speech perception. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-language Speech Research.* Timonium : York Press .171-204.

[8] Brière, E. J. (1966). An investigation of phonological interference. *Language* 42(4):768-796.

[9] Broselow, E. (1984). An investigation of transfer in second language phonology. *International Review of Applied Linguistics in Language Teaching* 22(4):253-269.

[10] Browman, C. P. & L. Goldstein. (1989). Articulatory gestures as phonological units. *Phonology* 6(2):201-251.

[11] Brown, C. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research* 14(2):136-193.

[12] Cebrian, J. (2002). Phonetic Similarity, Syllabification and Phonotactic Constraints in the Acquisition of a Second Language Contrast. Ph. D. Dissertation. Toronto: University of Toronto.

[13] Chang, C. B. (2015). Determining cross-linguistic phonological similarity between segments: The primacy of abstract aspects of similarity. In E. Raimy & C. E. Cairns (eds.). *The Segment in Phonetics and Phonology*. Oxford: Wiley - Blackwell .199-217.

[14] Chang, C. B., Y. Yao, E. F. Haynes & R. Rhodes. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America* 129(6):3964-3980.

[15] Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11(1):51-62.

[16] Eckman, F. R. (1991). The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition* 13(1):23-41.

[17] Eckman, F. R. & G. K. Iverson. (2013). The role of native language phonology in the production of L2 contrasts. *Studies in Second Language Acquisition* 35(1):67-92.

[18] Eckman, F. R., A. Elreyes & G. K. Iverson. (2003). Some principles of second language phonology. *Second Language Research* 19(3):169-208.

[19] Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15(1):47-65.

[20] Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-language Research.* Timonium: York Press .229-273.

[21] Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech* 1(2): 126-152.

[22] Holliday, J. J. (2015). A longitudinal study of the second language acquisition of a three-way stop contrast. *Journal of Phonetics* 50(1):1-14.

[23] Ishikawa, K. (2002). Syllabification of intervocalic consonants by English and Japanese speakers. *Language and Speech* 45(4):355-385.

[24] Jiang, F. & Y. B. Liu. (2014). The impact of L1 negative phonological transfer on L2 word identification and production. *International Journal of Linguistics* 6(5):37-50.

[25] Ladefoged, P. (1975). *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.

[26] Li, Y. (2008). Mandarin Speakers' Production of English and Mandarin Post-vocalic Nasals: An Acoustic Approach. M. A. Thesis. Victoria: University of Victoria.

[27] Major, R. C. (2001). Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology. Mahwah: Lawrence Erlbaum Associates.

[28] Pike, K. L. (1945). *The Intonation of American English.* Ann Arbor: University of Michigan Press.

[29] Ramus, F., M. Nespor & J. Mehler. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* 73(3):265-292.

[30] So, C. K. & C. T. Best. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition* 36(2):195-221.

[31] Tarone, E. E. (1980). Some influences on the syllable structure of interlanguage phonology. *International Review of Applied Linguistics in Language Teaching* 18(2):139-152.

[32] Weinreich, U. (1953). *Languages in Contact: Findings and Problems.* New York: Linguistic Circle of New York.

[33] Yang, Y. X. & X. X. Chen. (2019). Within-organ contrast in second language perception: The perception of Russian initial /r-l/ contrast by Chinese learners. *The Journal of Acoustic Society of America* 146(2):EL117-EL123