ABSTRACT

The purpose of the present thesis is to provide a comprehensive test of a recent model of L2 speech perception, the Second Language Linguistic Perception (L2LP) model. A unique strength of L2LP is that it can provide computational simulations of L2 perceptual acquisition that are testable with empirical data. According to L2LP, three types of learning scenarios can be distinguished depending on the number of sound categories in L1 and L2: the SIMILAR scenario, where the number of L1 categories equates with that of L2 categories; the SUBSET scenario, where the number of L1 categories exceeds that of L2 categories; and the NEW scenario, where the number of L1 categories falls short of that of L2 categories. The scenarios are associated with different levels of difficulty: SIMILAR < SUBSET < NEW. To test the model's predictions, the present thesis consists of three case studies (Studies 1, 2, and 3) that correspond to the three learning scenarios respectively. Each case study is summarized below.

Study 1 examined Japanese listeners' perception of American English (AmE) /iː/-/ɪ/, which are SIMILAR to Japanese /ii/-/i/. L2LP-based simulations predicted that Japanese learners of English would exhibit primarily duration-based perception for Japanese /ii/-/i/, while their perception for AmE /iː/-/ɪ/ would be more dependent on spectra and less on duration. The experiment, which manipulated the language context in two sessions, supported the predictions. Listeners showed the predicted shift in cue weighting between the sessions, despite the stimuli being identical. The result supports L2LP's separate grammars and perception modes hypotheses. However, duration still remained as the dominant cue for most learners, including those who had lived in the United States, which is at odds with L2LP's prediction that the SIMILAR scenario is of least difficulty.

Study 2 examined Australian English (AusE) listeners' perceptions of Japanese /ii/ and /i/, which constitute a SUBSET of their native vowels /iː, ɪ, ɪə/. The simulations predicted that AusE listeners would fully transfer their L1-like use of temporal and dynamic spectral cues to the perception of nonnative Japanese vowels. The prediction was borne out in the experiment, in which the listeners relied not only on duration but also on onset and offset formant frequencies in categorizing the Japanese vowels, the latter of which is important for the AusE vowel contrast but irrelevant to the Japanese length contrast. The result is compatible with L2LP's *Full Copying* hypothesis. However, the SUBSET scenario was not as difficult as L2LP claims because AusE learners of Japanese were expected to be capable of acquiring a native-like L2 length distinction with relative ease.

Study 3 investigated Japanese listeners' perception of a NEW vowel /æ/ in AmE by adopting segment- and feature-based versions of L2LP. While the segmental version had difficulty in explaining why AmE /æ/ is perceived as a 'deviant' exemplar of L1 Japanese categories and how this relates to new L2 category formation, the featural modeling provided a coherent explanation. Specifically, AmE /æ/ sounds deviant because it consists of /low/ and /front/ features that do not co-occur in Japanese (*/low, front/), and these features can be reorganized to form a new well-formed category in AmE. The experiment found that AmE /æ/ was perceptually represented as a 'fronted version of /a/' in Japanese listeners, which aligns well with the featural account. The result indicates that the NEW scenario is not difficult per se as L2LP proposes because the relatively inexperienced Japanese listeners without overseas experience could successfully establish a new sound representation. Instead, whether and how the relevant cues are utilized in the L1 may predict the relative difficulty of acquisition better.

In general, L2LP-based simulations for each particular learning scenario were found to be accurate, showing a close resemblance to real listeners' perception. Therefore, as an overall evaluation, the present thesis has verified the validity and usefulness of the model. However, the model's prediction concerning the difficulty of acquisition across scenarios was found to be inaccurate. As opposed to the prediction (SIMILAR < SUBSET < NEW), the attested levels of difficulty in the present thesis were SUBSET < SIMILAR ≦ NEW. This draws attention to a necessity to amend the current version of L2LP. While several suggestions can be proposed for extending the model, the most promising avenue for future research seems to be to incorporate the Bidirectional Phonology and Phonetics (BiPhon) framework, which would allow L2LP to model not only L2 perception but also L2 production for a better understanding of L2 speech acquisition.