L2 Vowel Learning

Juli Cebrian
Universitat Autònoma de Barcelona
juli.cebrian@uab.cat

Word count: 3000

Abstract

L2 vowel learning refers to the process of learning to perceive and to produce the vowel sounds of a second language (L2), that is, a language that is learned after acquiring a first language or mother tongue (L1). This process is greatly affected by the existence of the already learned language and by the specific characteristics of the L1 and the L2. Many factors that affect L2 vowel learning are thus linguistic in nature, such as the number of L2 (and L1) vowels, the perceived similarity between L1 and L2 vowels, the use that languages make of different phonetic cues to distinguish vowels, and the effect that neighboring sounds and suprasegmental features like stress and rhythm have on vowel perception and production. L2 vowel learning is also affected by the individual characteristics of the learners, such as age, years of learning, setting of learning, and psychological variables like motivation and aptitude, among others. L2 vowel learning can take place in a naturalistic setting through exposure to the target language but can also result from more direct interventions like pronunciation teaching, phonetic training and the use of modern technologies. Another factor that can affect L2 vowel learning is the influence of the orthography (spelling). The way these different factors may influence the process of L2 vowel learning will be described and discussed in this entry.

Keywords

Second language acquisition, vowels, L2 speech learning, speech perception and production, linguistic factors, learner factors

L2 vowel learning makes reference to the process of learning to perceive and to produce the vowel sounds of a second language (L2) after having acquired a first language or mother tongue (L1). The language that is being learned is often referred to as the target language (TL). The use of the word *learning* generally indicates that this process takes place after childhood, as opposed to acquiring an L1 from birth. Learning can take place in an instructional setting (e.g., learning a foreign language (FL) at school) or a naturalistic setting (e.g., residing in a TL-speaking country) and the terms *learning* and *acquisition* may refer to the distinction between the unconscious process of L1 acquisition from birth and the learning of additional languages at a later stage. However, this distinction is not always clear

cut, as learners may undergo instruction in a primarily naturalistic setting and L2 structures may be unconsciously acquired in an instructional setting. On the other hand, many individuals can speak more than two languages, either because they are bilingual speakers of two languages since childhood or because they have learned successive languages. For the sake of simplicity, the phrase *L2 vowel learning* will be used in this entry as an umbrella term meaning the learning of the vowels of a language other than the L1 and learned after the L1 (or L1s), and including both L2 learning and FL learning.

The focus of the entry is the learning of vowels, that is, speech sounds that are produced with unobstructed flow of air. In the articulation of a vowel, the air stream originated at the lungs travels through the larynx, making the vocal folds vibrate and creating a sound whose resonating quality is shaped by the configuration of the supralaryngeal vocal tract (degree of opening of the mouth, position of the tongue with respect to the roof of the mouth and to the front and back of the mouth, as well as the shape of the lips). There are as many articulatory configurations as there are vowels in the languages of the world, and languages vary in the number and type of vowels they have. For instance, in most languages only back vowels are produced with lip rounding (as /u/ in *boot*), but many languages like French, Swedish or Turkish have front rounded vowels, e.g., the French high /y/, high mid (/ø/) and low mid (/œ/) front rounded vowels in 'tu,' *you* SG., 'bleu,' *blue*, and 'cœur,' *heart*, respectively. Similarly, while all languages have oral vowels (vowels articulated with air flowing only out the mouth), a number of languages, like Cherokee, French, Polish, Portuguese and Yoruba, also have nasal vowels (produced with air flowing simultaneously out the mouth and the nose). L2 learners may have difficulty with L2 vowels, vowel contrasts and linguistic characteristics that are not present in their L1. Learner factors, that is, characteristics of the learners themselves, also play a role. The following sections elaborate on the topic of L2 vowel learning by developing these issues further.

[A] The Development of an L2 Vowel System

An important difference between acquiring the L1 and learning an additional language is that in the latter case the learner already has a language, which can exert an influence on the development of the new language. This influence is believed to be greater the later in life L2 learning begins, which also corresponds to a greater development of the L1. Thus it is generally accepted that the younger the starting age of learning, the more likely it is that the learner will approximate native speaker performance. Still, age alone cannot explain the different outcomes often observed across individuals. A number of other variables interact with age of learning, including amount of L1 (and L2) use, length of residence in a TL-speaking country, and years of L2 instruction. For example, in a study examining the perception of English vowels by several groups of Italian L1 speakers, Flege and MacKay (2004) found that while early L2 learners (with a pre-adolescence starting age of learning) outperformed late L2

learners, only early learners who reported a high use of their L1 (and a low use of their L2) reached accuracy levels comparable to those of native speakers' (see also Flege and Bohn's (2021) notion of full-time equivalent of L2 input, which combines years of learning and the proportion of L2 use). Still, there are other differences between individuals that also affect the learning of L2 vowels, including psychological variables like motivation to learn, language learning aptitude, phonological memory (i.e., ability to store and recall speech-based information in short-term memory), attention control or inhibitory control (ability to control our attention and focus on relevant information disregarding distracting stimuli), among others.

Studies on L2 vowel learning have shown that some non-native vowels are more readily learned than others. According to most accounts of L2 speech learning, TL sounds are typically equated to native categories, at least at the initial stages of learning. Thus the extent to which cross-linguistic mapping affects the ability to perceive and produce L2 vowels accurately is related to the degree of perceived similarity between L1 and TL sounds. For instance, the Native Language Magnet theory (NLM, Kuhl & Iverson, 1995) suggests that a TL sound is more likely to be perceived in terms of an L1 category the smaller the phonetic distance between the TL sound and the L1 category prototype. Models like the Perceptual Assimilation Model (PAM-L2, Best & Tyler, 2007), or the L2 Linguistic Perception model (L2LP, Elvin & Escudero, 2019) make predictions about accuracy of TL vowel discrimination and about possible learning scenarios based on the degree of similarity between the non-native and the L1 sounds. Similarly, the Speech Learning Model (SLM-r, Flege & Bohn, 2021) proposes that TL vowels that are perceived to be more different from native vowels are more likely to be perceived and produced accurately, given enough input and exposure to the TL, than vowels that are perceived to be similar to L1 vowels. The reason for this is that similar vowels are readily assimilated to L1 categories and are perceived and produced in terms of their L1 counterparts, while dissimilar vowels are more likely to be detected as different from L1 categories and may be eventually categorized accurately. For instance, Bohn and Flege (1992) found that the production of the new vowel /æ/ by German learners of English improved as a result of L2 experience (understood as length of residence in the U.S.), whereas English vowels like /i, ɪ, ɛ/, with a clear counterpart in German, were equally produced by learners of different levels of experience. However, not all vowel learning trajectories can be explained by the relative similarity between TL and L1 categories. A number of factors influence the development of L2 categories and interact with cross-language mappings, such as individual differences, and linguistic and metalinguistic factors such as the role of language instruction and orthography, as discussed in the next section.

Given the role of cross-linguistic perception, a related issue of theoretical importance is whether it is necessary for learners to perceive TL vowels correctly in order to produce them correctly. For example, is it necessary for English learners of French to be able to distinguish French /u/ and /y/ (as in 'tout,' *all*, and 'tu,' *you* SG.) perceptually before they can successfully produce the two vowels differently? Some studies have shown that learners' L2 perception ability can predict L2 production accuracy (e.g., Flege, Bohn & Jang, 1997). Still, research on this issue has not yielded conclusive results as empirical studies provide evidence of different types of interaction between perception and production and generally report only modest correlations between the two modalities. This has led recent theories to shy away from making strong claims about precedence requirements. For example, the SLM-r suggests that the two modalities are strongly connected but co-evolve without necessarily having one modality precede the other.

[A] Linguistic Factors Affecting L2 Vowel Learning

A number of linguistic factors may influence the learning of L2 vowels. For example, the L1 and the TL may differ in the number of vowels they have, that is, the size of the vowel inventory. We may expect that speakers of languages with small vowel inventories (e.g., the three-vowel system (/i, u, a/) of Quechua or Kalaallisut or the five-vowel system (/i e a o u/) of Japanese, Greek and Spanish) will have trouble learning to perceive and produce vowels from TLs with larger vowel inventories. For example, in a study that compared the perception of English vowels by speakers of four languages that varied in vowel inventory size, Iverson and Evans (2007) found that speakers of large inventory languages like Norwegian and German (18 and 15 vowels, respectively) were more accurate at identifying English vowels than speakers of languages with smaller inventories like French and Spanish (11 and 5 vowels respectively). Still, studies also show that speaking an L1 with a large vowel inventory does not always offer an advantage when learning an L2, as a smaller TL inventory may include vowel contrasts that are not present in the L1 (Colantoni, Steele & Escudero, 2016). Furthermore, languages with small vowel inventories may actually have a number of diphthongs (like the vowel sequences in Spanish 'ley' /lei̯/, *law*, or 'hay' /ai̯/, *there is*), which also play a role in TL to L1 perceptual mappings. On a related note, studies have also shown that the degree of compactness of the L1 categories may also affect the categorization of TL vowels such that speakers with comparatively compact L1 vowel categories (i.e., with little within category variability) may be more likely to detect differences between native and TL vowels and consequently categorize TL vowels differently from native vowels (Flege & Bohn, 2021; Colantoni et al., 2016).

Languages may also differ in the use of acoustic cues to signal vowel distinctions. For example, languages like Estonian, Japanese or Swedish distinguish some pairs of vowels by their duration. Thus,

the contrast between a short and a long /i/ distinguishes the Japanese words ‘imasu’ (to exist) and ‘iimasu’ (to say). In principle, duration differences are more likely to be learned by speakers of languages that use duration distinctively. Still, even speakers of languages that don’t use vowel duration as a cue to vowel identity may detect and learn to use temporal cues to learn to distinguish vowel contrasts whose spectral differences are not easily detected. For example, Spanish learners of English have been found to make use of duration in their distinction of the English /i/-/ɪ/ contrast, found to have no clear counterpart in the L1, while duration is not resorted to in the case of /æ/ and /ɛ/, which map onto Spanish /a/ and /e/. By contrast, duration is used by German learners of English for the /æ/-/ɛ/ contrast, with no clear L1 counterpart, but not relied upon for the /i/-/ɪ/ contrast, which is assimilated to the German /i/-/ɪ/ pair (Flege et al., 1997).

L2 vowel learning may also be affected by the phonetic context, i.e., the neighboring sounds and syllabic position. For instance, the surrounding consonants have been found to affect the degree of perceived similarity between TL and L1 vowels. English native speakers perceive the French high front rounded vowel /y/ to be closer to the English high back rounded vowel /u/ when the French vowel is next to an alveolar consonant (/t, d/) than when next to other consonants (e.g., Levy, 2009). This is explained by the fact that English /u/ is more fronted, and hence more /y/-like, in this context. Moreover, vowels may be affected by suprasegmental phenomena like stress, rhythm and tone. For example, vowels in unstressed position in English are often articulated as a reduced mid central vowel (i.e., /ə/, as the first vowel in *again* or the middle vowel in *melody*). Thus contextual factors and suprasegmental features need to be considered when assessing L2 vowel learning.

Adult learners tend to be exposed to written language (through instruction, reading or everyday life situations) as much as to spoken language. Hence, the way sounds are represented orthographically may affect learners’ categorization of those sounds due to L1-TL differences in spelling or because of TL spelling inconsistencies. For example, many languages that use the Roman alphabet, like Italian and Spanish, use letter <i> to represent primarily vowel /i/, while in English this letter most frequently stands for the diphthong /aɪ/ as in *dine* and *highlight* and /ɪ/ as in *pin* and *criminal*. Thus, the mispronunciation of English *bit* as *beat* may be due not only to the difficulty in distinguishing /i/-/ɪ/ but also to the effect of L1 orthography. Similarly, English learners may pronounce the word *son* differently from *sun* due to the difference in the spelling.

Finally, the likelihood that a given L2 vowel contrast may be learned may also be affected by how productive that contrast is in the TL, i.e., its functional load. For instance, the English /i/-/ɪ/ contrast has a high functional load, as evidenced by the numerous pairs of words that are distinguished by these two

vowels (e.g., minimal pairs like *leave-live*, *deep-dip*, etc.). This may facilitate the learning of this contrast through exposure to minimal pairs, and as a result of attention directed to this contrast in language instruction. By contrast, a pair like English /u/-/ʊ/ (as in *pool-pull*), with a comparatively low functional load, may be less easily detected by L2 learners, and the failure to produce this contrast is less likely to result in miscommunication.

[A] Final Remarks and Conclusions

A good understanding of the process of L2 vowel learning needs to take a number of linguistic and learner factors into account. Research shows that learners tend to assimilate TL vowels to existing L1 categories. Thus L2 vowel learning studies should have as their starting point an analysis of the similarity between L1 and L2 vowels so that predictions about accuracy of L2 perception and production can be made. Further, a whole-system approach, assessing the learning of the whole L2 vowel system, can provide a more accurate picture of the process of L2 vowel learning as opposed to simply focusing on individual vowels or specific contrasts. In addition, while generalizations can be made for learners from the same L1 background, it is clear that learning trajectories may vary greatly from individual to individual as a result of individual factors and differences in L1 development.

There are other issues that are relevant to L2 vowel learning that have not been discussed in this entry for the sake of brevity. For instance, dialectal variation often affects the pronunciation of vowels; consequently, the choice of specific TL variety is not a simple one. Similarly, vowels are not learned in isolation but in words and thus vowel learning cannot be fully understood without considering word learning as well. Most outcomes discussed in this entry stem from cross-sectional studies; more comprehensive studies on L2 vowel learning are needed from a longitudinal approach (e.g., Munro & Derwing, 2008). L2 vowel learning can take place in a naturalistic setting through exposure to the target language but can also result from more direct interventions like pronunciation teaching, phonetic training and the use of modern technologies. Finally, vowel learning and teaching need not be limited to auditory perception and production; visual stimuli and gestural representations may also play a role.

Cross-references: wbeal0292.pub3, wbeal20625, wbeal20639, wbeal20641, wbeal20642, wbeal20646

References

Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn, & M. Munro (Eds.), *Language Experience in Second Language Speech Learning – In honor of James Emil Flege* (pp. 13-34). John Benjamins Publishing Company.

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14, 131-158.

Colantoni, L., Steele, J., & Escudero, P. (2016). *Second Language Speech. Theory and Practice*. Cambridge University Press.

Elvin, J., & Escudero, P. (2019). Cross-Linguistic Influence in Second Language Speech: Implications for Learning and Teaching, in *Cross-Linguistic Influence: From Empirical Evidence to Classroom Practice*. edited by M. J. Gutierrez-Mangado, M. Martinez-Adrian, and F. Gallardo-del-Puerto (Springer, Cham, Switzerland), pp. 1-20.

Flege, J. E., & Bohn, O. S. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: theoretical and empirical progress* (pp. 3–83): Cambridge University Press.

Flege, J., Bohn, O-S., & Jang, S. (1997). The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics* (1997) 25, 437—470.

Iverson, P. and Evans, B. G. (2007) Learning English vowels with different first-language vowel systems: perception of formant targets, formant movement, and duration. *Journal of the Acoustical Society of America* 122(5): 2842-2854.

Kuhl, P., & Iverson, P. (1995). Linguistic experience and the "Perceptual Magnet Effect". In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross Language Research* (pp. 121-154). York Press.

Levy, E. S. (2009). Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *Journal of the Acoustical Society of America* 125, 1138–1152.

Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning* 58:3, 479–502.

Suggested Readings

Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. Language and speech 53(3), 343-365.

Ladefoged, P., & Ferrari Disner, S. (2012). *Vowels and Consonants*. 3rd Edition (Wiley-Blackwell, Chichester, UK).

Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29, 191-215.

Strange, W. (2007). Cross-language phonetic similarity of vowels: Theoretical and methodological issues. In O.-S. Bohn, & M. Munro (Eds.), *Language Experience in Second Language Speech Learning – In honor of James Emil Flege* (pp. 35-55). John Benjamins Publishing Company.

Tyler, M. D. (2021). Phonetic and phonological influences on the discrimination of non-native phones, in *Second language speech learning: Theoretical and empirical progress*, edited by R. Wayland (Cambridge University Press, Cambridge, UK), pp. 157-174.

Contributor bio

Juli Cebrian is Associate Professor at the department of English Philology at Universitat Autonoma de Barcelona, where he teaches undergraduate and graduate courses in English phonetics and phonology, pronunciation, oral skills, second language acquisition and research methods. His research investigates the acquisition of the phonology of a second/foreign language by adult learners, and his interests extend to experimental phonetics, bilingualism, psycholinguistics and pronunciation teaching. He has collaborated in and led a series of funded research projects on the acquisition of second language speech, cross-linguistic perception and phonetic training. His publications include numerous contributions to leading international journals and specialized books.