

**ASSESSMENT OF CROSS-LANGUAGE SPEECH PERCEPTION
IN ADULTS: STRONGLY CONSIDERING PHONETIC
AND ACOUSTIC CONTRIBUTIONS**

Luke L. JACKSON, *Ph.D*
Department of English
Faculty of Art
University of Sheffield
Sheffield, South Yorkshire, England
United Kingdom

ABSTRACT

Cross-language research on adult speech perception demonstrates a strong effect of linguistic experience on consonant perception but not precisely on vowel perception. First, identification and goodness functions for the high front quadrant of the vowel space were mapped for speakers of Swedish, English, and Spanish. Second, speakers performed a discrimination task for one vector in this vowel space. Stimuli along this vector were identified by Swedish speakers as belonging to the Swedish front rounded vowel series /ç:/- /ö:/. However, English and Spanish speakers reported that the stimuli were not in their language. Significant differences in the discriminability of these stimuli were observed across speakers of different languages. When translators or interpreters have this level of language competence, they demonstrate the ability to communicate between languages using complex sentence structures, a high level of vocabulary, and the ability to describe concepts or words when they do not know the actual word or phrase. Our results show that linguistic experience plays a significant role in vowel discrimination.

KEYWORDS: Cross-language, Roles, Challenges and Adults

Introduction

The development of models of cross-language speech perception and production has been shaped by the disparate goals of the authors proposing them. Flege (1995) and Flege (2003) intend the Speech Learning Model (SLM) to be a theory of second language acquisition after the critical period, and the author's research often focuses on how factors such as age of acquisition or length of residence affect speakers' ability to perceive and produce sounds in their second language. Learning a second or foreign language (L2) after childhood is a difficult task for a variety of reasons. In addition to learner-related factors such as age of learning and amount of first

and second language use (Piske et al., 2001; Flege et al., 2003), the existence of a first language (L1) phonological system in part accounts for the fact that adult L2 speakers rarely sound like native speakers (Trubetzkoy, 1939; Rochet, 1995; Strange, 1995).

Escudero's (2005) Second Language Linguistic Perception Model (L2LP) posits that L2 phones that have a similar L1 counterpart are easier to learn than new phones with no clear counterpart in the L1. The L2LP claims that the L2 is, originally, a copy of the L1, and this copy may then evolve toward more target-like values. Thus, learning a similar sound involves an adjustment of the category boundaries and their acoustic properties, whereas acquiring a new sound requires establishing a new L2 to L1 mapping prior to the process of phonetic adjustment. Previous studies have found that L1 Spanish/Catalan speakers have difficulty with the English /i/-/ɪ/ contrast but have not been able to relate it to the degree to which each vowel is assimilated into a L1 category. Cebrian (2006) found that two groups of L2 English speakers (20 Catalan speakers living in Canada and 30 Catalan undergraduate students of English living in Spain) perceptually assimilated Canadian English /i, ε, eɪ/ to Catalan English /i, ε, ei/, respectively, whereas English /ɪ/ obtained lower assimilation scores and goodness ratings than Catalan /e/. Perception of L2 sounds was examined with an identification test involving a synthetic continuum from /i/ to /ɪ/ to /ε/ varying in vowel quality and vowel duration. The two groups performed like native speakers in their perception of English /ε/, but differed from native speakers in their perception of English /i/ and /ɪ/, which showed a predominant reliance on temporal cues, unlike native Canadian English speakers, who relied mostly on spectral differences.

Concept of Cross-Language Speech

Cross-language speech perception and production (CLSP) is the study of how speakers produce and perceive sounds, sequences, prosody, and tones that are not found in their native language. This has been examined from more than a few perspectives, notably phonetic, theoretical/phonological, and psycholinguistic. The development of models of cross-language speech perception and production has been shaped by the disparate goals of the authors proposing them. Flege (1995) and Flege (2003) intend the Speech Learning Model (SLM) to be a theory of second language acquisition after the critical period, and the author's research often focuses on how factors such as age of acquisition or length of residence affect speakers' ability to perceive and produce sounds in their second language. The goal of cross-language is to build search engines that use a query expressed in one

language (e.g., English) and find content that is expressed in some other language (e.g., French).

Cross-language studies include studies of how speakers confront non-native sounds and structures even when they are not actively learning the language under study. Cross-language has a number of purposes, including getting a glimpse into the initial stages of adult second language (L2) acquisition, gaining a better understanding of the structure of the native phonology by understanding how speakers and listeners modify structures that are not allowed by their native phonology, and probing factors that affect loanword adaptation. Cross-language speech perception focuses on "selective perception routines," which refer to highly practiced behaviors used to perceive native sounds.

Cross-Language Speech Perception in Adults

Cross-language studies have shown that foreign consonant contrasts vary in the degree of perceptual difficulty with which they present adult non-native listeners. Phonemic, phonetic, and acoustic factors have been considered important in accounting for this variability. A very common belief concerning second language acquisition (SLA) is that children are better and faster learners than adults. This belief is especially strong in relation to phonology. Although some adult learners of a second language may attain a relatively high or even a complete mastery of the language, they do not seem to be able to get rid of their native accent. In fact, differential perception of length contrasts can already be observed in 18-month-old infants learning a language with phonemic vs. non-phonemic length, such as Dutch (Dietrich, Swingley, & Werker, 2007) or Japanese (Mugitani, Pons, Fais, Werker, & Amano, 2008). Previous cross-linguistic research on vowel production by native adult speakers of different languages has demonstrated systematic differences in how shared vowels (vowels represented by common phonemic symbols) are produced. Yang (2006) also found cross-linguistic differences between shared vowels produced by adult English and Korean speakers. English /u/ had higher F2 values than Korean /u/, and English /a/ had lower F2 values than Korean /a/. Both observations suggest that the concept of "shared" vowels does not account for subtle differences in vowel production across different languages.

Roles of Cross-language speech in Adults

The challenge of making sound predictions has become increasingly evident, however, and critics of contrastive analysis have emphasized cases where cross-linguistic comparisons fail to predict actual difficulties and where

difficulties predicted do not always materialize. Cook (2000), for example, observes that transfer and cross-linguistic influence spuriously suggest some kind of movement. Jarvis and Odlin (2000) note that cross-linguistic influence might involve actual phonological forms or simply the semantic structures represented by the form; it might involve either production or comprehension, or both; it might involve inflectional as well as derivational morphology, or simply one or the other; and it could involve either positive or negative transfer. Cross-language can also go a long way towards:

Development of Grammar: The most compelling evidence for cross-language variation begins between 18 and 20 months (on average), when grammatical development is finally underway. Ironically, early cross-linguistic work on language acquisition was based on the assumption that grammar (as opposed to phonology or the lexicon) would prove to be the bastion of language universals. Some secondary sources still claim that all children acquire language on the same schedule, in the same way, and this putative fact has led to further claims about a universal bioprogramme that governs language acquisition in children as well as the emergence of new languages from pidgin codes (i.e., creolization) (Bickerton 2004). In this scenario, all children (and all creoles) begin their linguistic careers with single uninflected words, followed by telegraphic combinations of uninflected words in ordered strings, with inflections and function words acquired only after this syntactic base has been established. None of these proposed universals has held up in cross-linguistic research.

Word Comprehension and Production: We have learned a great deal in the past few years regarding cross-linguistic similarities and differences in early lexical development, due in part to the development and proliferation of new parent report instruments that are low in cost but high in reliability and validity (Fenson et al., 2004, 2000). Grammatical development does begin with something like a one-word stage in every language, but there are cross-language variations in the form of one-word speech. For example, infant speakers of Western Greenlandic start out by producing little pieces of the large and complex words of their language (in which a sentence may consist of a single word with 10–12 inflections). In other richly inflected languages (e.g., Turkish), children often produce inflected nouns and verbs late in the one-word stage, before they have produced any word combinations at all. Some of these inflected forms may be accomplished by rote, but when there are multiple examples in which the same word appears with several contrasting inflections, it seems reasonable to infer that some kind of productive process is underway. By tapping into parental knowledge, researchers have charted means and variations in word comprehension and

production in children between 8–30 months of age, with instruments that are now available in more than a dozen languages [Afrikaans, American Sign Language, Catalan, Chinese, Croatian, Danish, Dutch, English (British and New Zealand), Finnish, French (Canadian), Greek, Hebrew, Icelandic, Italian, Japanese, Korean, Malawian, Polish, Sign Language of the Netherlands, Spanish (Mexican and Spain), Swedish]. These parental inventories rely on recognition memory rather than recall (using checklists of words that are among the first 600–700 words to be acquired in that language), and they are used only within the age ranges in which parents can give reliable reports of newly emerging behaviors (e.g., word comprehension can only be assessed with these methods between 8–18 months; word production can be assessed reliably between 8–30 months). Briefly summarized, two universal conclusions have emerged from this multinational effort: (a) Average onset times appear to be the same across languages for word comprehension (8–10 months) and word production (11–13 months); (b) huge variation in lexical growth is found in every language and appears to be equivalent across languages in shape and magnitude (e.g. a range from no word production at all to production of more than 500 words at 24 months).

Speech Production: Despite ample evidence for the early acquisition of language-specific contrasts in speech perception, we know relatively little about the emergence of corresponding contrasts in speech production. For most children, canonical or reduplicative babbling begins between 6–8 months, with short segments or longer strings that are punctuated by consonants (e.g. “dadada”). Boysson-Bardies and colleagues (2004) have reported that babbling “drifts” toward the particular sound patterns of the child’s native language between 6–10 months (i.e. native speakers can discriminate at above-chance levels between babble by Chinese, Arabic, English, or French infants). However, the phonetic basis of these adult judgments is still unknown. Critics of this research have argued that there are hard maturational limits on the infant’s ability to control the detailed gestures required for speech production, suggesting that babbling and early words are relatively immune to language-specific effects until the second year of life (Eilers, et al., 2003).

Challenges of Cross-language Speech

When a language barrier exists between qualitative researchers and their participants, the research becomes a cross-language qualitative study with unique challenges related to language (Temple, 2002; Temple & Young, 2004). Inconsistent or inappropriate use of translators or interpreters can threaten the trustworthiness of cross-language qualitative research and, subsequently, the applicability of the translated findings to participant

populations (Edwards, 2008). Poorly translated concepts or phrases will change what themes emerge from the analysis and may not reflect what the participant actually said. This threatens, for example, the credibility and dependability of the cross-language study and forms part of the study's limitations. Using that standard with cross-language research, then, how researchers describe the way they use translators in cross-language qualitative research reflects their competence in addressing language as a method issue. Several method articles broaching the issue of cross-language research have appeared since the year 2000. All provide salient points about tackling issues related to cross-language research, but no methodological consensus has emerged from them. Thus, researchers who fail to systematically address the methodological issues translators present in a cross-language qualitative research design can decrease the trustworthiness of the data and the overall rigor of the study (Edwards, 2008). When translators or interpreters have this level of language competence, they demonstrate the ability to communicate between languages using complex sentence structures, a high level of vocabulary, and the ability to describe concepts or words when they do not know the actual word or phrase. With this level of language competence, translated data is less likely to have errors related to translation (Jandt, 2003).

Concept of Phonetic Contributions

Phonetics is a part of the English language that helps us understand sounds of various alphabets. How an alphabet should sound is taught to us with the help of Phonetics. Phonetics play a very important role in improving our communication. All the alphabets and words must sound correctly, or else the content as well as our communication will lack luster and sound unimpressive. According to O'Grady (2005), phonetics is a branch of linguistics that studies how humans produce and perceive sounds, or in the case of sign languages, the equivalent aspects of sign. The field of phonetics is traditionally divided into three sub-disciplines based on the research questions involved, such as how humans plan and execute movements to produce speech (articulatory phonetics), how various movements affect the properties of the resulting sound (acoustic phonetics), or how humans convert sound waves to linguistic information (auditory phonetics). Phonetics broadly contributes to two aspects of human speech:

- **Production:** The ways humans make sounds
- **Perception:** The way speech is understood.

The communicative modality of a language describes the method by which a language produces and perceives languages. Languages with oral-aural

modalities such as English produce speech orally (using the mouth) and perceive speech aurally (using the ears). Language production consists of several interdependent processes that transform a non-linguistic message into a spoken or signed linguistic signal. In the same vein, phonetics has brought new insights to the teaching of pronunciation. Concerned with the study of the physical properties of sounds and their place and manner of articulation in the vocal tract, Broughton et al. (2009) stated that phonetics forms a tool of paramount importance that is used in the teaching of pronunciation. For instance, in any description of the English sound system, speech sounds are categorized into consonants and vowels.

Concept of Acoustic Contributions

Acoustic phonetics is a subfield of phonetics, which deals with acoustic aspects of speech sounds. Acoustic phonetics investigates time domain features such as the mean squared amplitude of a waveform, its duration, its fundamental frequency, or frequency domain features such as the frequency spectrum, or even combined spectrotemporal features and the relationship of these properties to other branches of phonetics (e.g. articulatory or auditory phonetics), and to abstract linguistic concepts such as phonemes, phrases, or utterances. Acoustic phonetics uses the frequencies of these sound waves to precisely analyze speech. All sounds are produced by movement of molecules (typically of air); they move in a vibrating-like fashion, creating periods of rarefaction (farther apart) and compression (closer together) between molecules. The air, they create many cycles that, as a whole, can be referred to as a wave. The speed of the vibrations (compression + rarefaction) is called the frequency and is a very common measure of sound.

The unit of measurement for frequency is Hertz (Hz) and 1 Hz is equal to 1 cycle/second. Humans are typically able to detect sound waves with frequencies of 20-20,000 Hz traveling through air. Acoustic phonetics uses the frequencies of these sound waves to precisely analyze speech. Acoustic phonetics is the study of the physical features of speech sounds, in particular those that are linguistically relevant and can be detected by the human ear, and the medium in which they travel. here are many possible speech sounds that can be used in a language but not every sound is used in every language, and the same sounds can be used in different ways depending on the language or dialect. In the latter case, acoustic phonetic properties of a word can change from language to language, and even dialect to dialect. For example, the acoustic characteristics at the end of the word “checker” would differ in British English compared to American English. The speech sounds actually used in a particular language are considered relevant, while any

others are not considered when studying acoustic phonetics of that language.

An example of a sound not heard in English but used in other languages are clicks, such as the (post)alveolar click. There are also sounds that can be made by the vocal tract, like coughs and burps that are not considered relevant speech sounds in any language. To sum up, the relevance of a speech sound depends on the language you are interested in, and as a general rule, acoustic phonetics looks at the speech sounds meaningfully produced within a language.

Conclusion

When translators or interpreters have this level of language competence, they demonstrate the ability to communicate between languages using complex sentence structures, a high level of vocabulary, and the ability to describe concepts or words when they do not know the actual word or phrase. Also, inconsistent or inappropriate use of translators or interpreters can threaten the trustworthiness of cross-language. Finally, linguistic experience plays a significant role in vowel discrimination.

Recommendations

1. Translators or interpreters should not use inappropriate interpretations less they threaten the trustworthiness of cross-language subsequently.
2. Researchers should not relent in put forward salient points about tackling issues related to cross-language research.

REFERENCES

- Bickerton D. (2004). The language bioprogram hypothesis. *Behav. Brain Sci.* 7:173–87.
- Boysson-Bardies B, Sagart L, Durand C. (2004). Discernible differences in the babbling of infants according to target language. *Journal of Child Language.* 11:1–15.
- Broughton, G., Brumfit, C., Flavell, R., Hill, P., & Pincas, A. (2009). *Teaching English as a Foreign Language*. New York: Routledge.
- Cebrian, J. (2006). Experience and the Use of Non-Native Duration in L2 Vowel Categorization. *Journal of Phonetics.* 34, 372–387.
- Crossed wires (2002). Interpreters, translators, and bilingual workers in cross-language research. *Qual Health Res.* 12(6):844-54.
- Dietrich, D. Swingley, J. F. Werker (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences*, 104, pp. 16027-16031
- Edwards R. (2008). A critical examination of the use of interpreters in the qualitative research process. *Journal of Ethnic and Migration Studies.* 24:197–208.
- Eilers R, Oller D K, Levine S, Basinger O, Lynch M, Urbano R. (2003). The role of pre- maturity and socioeconomic status in the on- set of canonical babbling in infants. *Infant Behaviour Development* 16:297–316.
- Escudero's, P. (2005). *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization*. PhD Thesis. Amsterdam (Netherlands): Netherlands Graduate School of Linguistics.
- Fenson L, Bates E, Dale P, Goodman J, Reznick J S, Thal D. (2000). Measuring variability in early child language: Don't shoot the messenger. Comment on Feldman et al. *Child Development.* 71:323–28.
- Fenson L, Dale P A, Reznick J S, Bates E, Thal D, Pethick S J. (2004). Variability in early communicative development. *Monogr. Soc. Res. Child Dev. Ser.*, 242, 59 (5).
- Flege, J. E. (1995). Second Language Speech Learning: Theory, Findings and Problems. In *Speech Perception and Linguistic Experience: Issues in*

- Cross-Language Research*. Editor W. Strange (Baltimore, MD: York Press), 233–277.
- Flege, J. E. (2003). *Assessing Constraints on Second-Language Segmental Production and Perception*. In *Phonetics and Phonology in Language Comprehension and Production*. Editors A. Meyer and N. Schiller (Berlin, Germany: Mouton de Gruyter), 319–355.
- Jandt F. (2003). *An introduction to intercultural communication: identities in a global community*. Sage; Thousand Oaks, C A.
- Mugitani, F. Pons, L. Fais, J.F. Werker, S. (2008). Amano Perception of vowel length by Japanese- and English-learning infants. *Developmental Psychology*, 45(1), 236-247
- O'Grady, William (2005). *Contemporary Linguistics: An Introduction* (5th ed.). Bedford/St. Martin's.
- Piske, T., MacKay, I. R. A., and Flege, J. E. (2001). Factors Affecting Degree of Foreign Accent in an L2: A Review. *Journal of Phonetics*. 29 (2), 191–215.
- Rochet, B. (1995). "Perception and Production of Second-Language Speech Sounds by Adults," in *Speech Perception and Linguistic Experience: Issues in Cross Language Research*. Editor W. Strange (Timonium, MD: York Press), 379–410
- Strange, W. (1995). *Speech Perception and Linguistic Experience: Issues in Cross Language Research*. Timonium, MD: York Press.
- Temple B, Young A. (2004). Qualitative research and translation dilemmas. *Qualitative Research*. 4 (2):161–78.
- Trubetzkoy, N. (1939). *Principles of Phonology*. Berkeley, CA: University of California Press.
- Yang, B. (2006). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*. 24, 245–261.