

# CVC syllables for investigating the phonetic sensitivity of Mandarin and English speakers

MILES MUNRO

*University of California, Irvine, California*

Although many individual speech contrasts pairs have been studied within the cross-language literature, no one has created a comprehensive and systematic set of such stimuli. This article justifies and details an extensive set of contrast pairs for Mandarin Chinese and American English. The stimuli consist of 180 pairs of CVC syllables recorded in two tokens each (720 syllables total). Between each CVC pair, two of the segments are identical, whereas the third differs in that a segment drawn from a "native" phonetic category (either Mandarin, English, or both) is partnered with a segment drawn from a "foreign" phonetic category (nonnative to Mandarin, English, or both). Each contrast pair differs by a minimal phonetic amount and constitutes a meaningful contrast among the world's languages (as cataloged in the UCLA Phonological Segment Inventory Database of 451 languages). The entire collection of phonetic differences envelops Mandarin and English phonetic spaces and generates a range of phonetic discriminability. Contrastive segments are balanced through all possible syllable positions, with noncontrastive segments being filled in with other "foreign" segments. Although intended to measure phonetic perceptual sensitivity among adult speakers of the two languages, these stimuli are offered here to all for similar or for altogether unrelated investigations.

In an ongoing series of investigations into individual differences in second language acquisition of speech sounds, I have developed a set stimuli used to test listener aptitude for detecting small phonetic differences within their first language (L1) and second language (L2) categories. Tasks measuring differential listener sensitivity are compared with measures of idiosyncratic productive accent, in an attempt to identify one or more cognitive correlate(s) of spoken accent. The stimulus set, composed entirely of CVC syllables, was designed to assess L2 speakers of American English with Mandarin Chinese as their L1. The segments that compose each CVC are compiled from speech contrasts that are nonnative to the phonetic categories of these two languages and that mark meaningful distinctions among a substantial fraction of the world's languages, as cataloged within the UCLA Phonological Segment Inventory Database of 451 languages (UPSID451; Maddieson & Precoda, 1990). The location of the paired contrasting segment is rotated through syllable positions; the corpus of 180 pairs, with two tokens recorded for each (a total of 720 CVCs), is subdivided by target language(s). The impetus for this publication is to substantiate these stimuli and to make them available with the prospect of life in other investigations. Thus, first, the motivation for these stimuli will be provided, along with a discussion of their rationalization based on the current related L2 speech framework. This is then followed by a description of the CVC pairs, including how they were generated. Finally, this article concludes with a series of

suggestions of other ways that these stimuli might be employed in similar, or even unrelated, research.

## Background and Motivation

Accurate acquisition of L2 phonetic categories is governed primarily by how well L2 contrastive pairs phonetically map onto the particular L1 inventory. Specifically, the most difficult L2 speech sounds for late-learners to acquire are those in which a target phonetic category distinction is native within the L2 but nonnative within the L1, yet still resides within the same domain of phonetic space (i.e., the L1 and L2 sounds are very similar phonetically). However, not all late-learners are alike; the maturational state of the individual attempting to acquire such speech sounds imposes considerable variability onto the exactness of the assimilation of the target sounds. Although such individual differences have received little attention, the comparison of L1 against L2 phonetic categories has been the subject of numerous studies of perception, which can be broadly characterized as follows: A group of listeners with a shared L1 make perceptual judgments of two or more L2 sounds that, although contrastive within the L2, due to phonetic proximity are categorizable as a single subset of categories within the L1.

In the simplest case, in which two L2 sounds correspond to a single L1 sound, the dual assimilation can be described as an "even split," in which case both sounds in the L2 contrast fit the L1 phonetic category equally well, disregarding any insignificant asymmetry (Figure 1, top left). Alter-

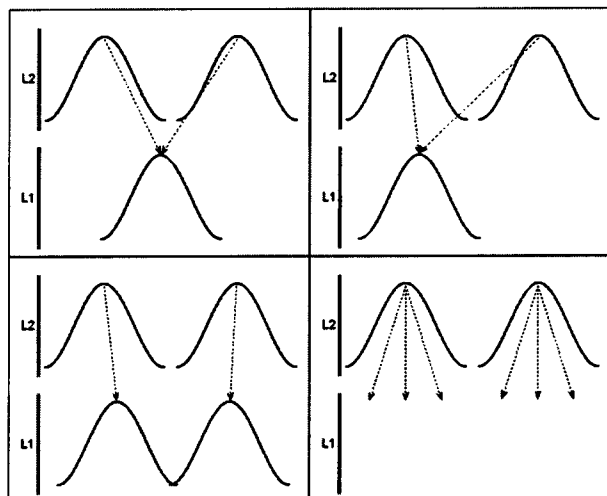


Figure 1. Problematic single-category (top left) and category goodness (top right), as well as nonproblematic two-category (lower left) and nonassimilable (lower right), cross-linguistic contrast assimilation patterns.

natively, a “third wheel” relationship can result, in which case one L2 sound of the contrast pair corresponds well to the L1 phonetic category, whereas the other maps substantially less precisely onto that same category (Figure 1, top right). Respectively, these patterned interactions of L1 against L2 phonetic categories are more commonly labeled as a *single-category (SC) assimilation* and *category goodness (CG) difference* (Best, 1994, 1995; Best & Strange, 1992). Other possible cross-language sound patterns are possible, too, such as two-category (TC; Figure 1, bottom left) assimilation and nonassimilable<sup>1</sup> (NA; Figure 1, bottom right) sounds, but the latter assimilations are considered mostly nonproblematic, comparatively speaking.

The cross-linguistic comparison literature offers several paradigmatic examples of each “problematic” contrast correspondence type. For SC interactions, the illustrious English contrast [ɹ]:[l], in which both phones correspond nearly equally well/poorly to the single Japanese /r/ category, has been examined by copious studies (to name only a few recent examples, Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Best & Strange, 1992; Logan, Lively, & Pisoni, 1991; McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002; Yamada & Tohkura, 1992). Somewhat less famously, the CG pattern can be typified by the French and German [u]:[y] contrast with English /u/ (see, e.g., Flege & Hillenbrand, 1984; Gottfried, 1984; Polka, 1995; Rochet, 1995), in which the phone [u] manifests much greater (although not perfect) category goodness to English /u/ over the phonetically similar [y]. The contrasts studied are not limited to this impoverished sampling. In fact, numerous studies have looked at various other SC and CG contrasts within an assortment of L1s and L2s.

One major drawback to testing phonetic similarity and individual differences in such studies is that, for any given language, at most only a few of the native phonetic categories have been examined within the entirety of the available L1 inventory. Similarly, given any particular L1 phonetic

category, only a small number of the potential nonnative contrasts have been considered. Furthermore, in almost the entirety of L2 contrast studies, the question is always one of differentiability of a particular L2 contrast within an L1 category or of maturational effects for either infants or adults. This situation has had both positive and negative effects. On the one hand, collectively such studies have provided a fairly convincing demonstration that individual L2 contrasts range somewhere between easy and hard, generally becoming more difficult with age. On the other, limiting the contrasts investigated per study to only a few (at most) has restricted the extent to which generalizations of phonetic discriminability or examinations of individual variability can be achieved. Moreover, any individual differences fortuitously observed within this general pattern have been underemphasized, despite numerous remarks, including the one by Strange (1995) over a decade ago that “[p]erformance across a variety of nonnative phonetic categories ranges from near native-like levels of accuracy to chance performance” (p. 35).

The range of severity of spoken accent among late-learners parallels this pattern. Any relation may be due to the hypothesis that success at acquiring contrasts underlies at least part of productive and perceptual L2 accent. More specifically, failure to distinguish a pair of L2 sounds results in their perception and (hence) pronunciation as the single L1 category to which they both correspond. The lack of distinction between L2 segments is manifest in production as a recognizable spoken accent, which perceptually can lead to difficulty understanding heard speech. In the case of SC (Figure 1, top left), this may result in a significantly deviant representation of both L2 sounds, whereas in the case of CG (Figure 1, top right), only one of the phones in the L2 contrast will likely be appreciably distorted (Best, 1994, 1995; Best & Strange, 1992). Having examined a wide range of cross-language segment pairs, Best and colleagues found “a gradient of performance across nonnative contrasts, ranging from near-chance to near-ceiling” (Best, McRoberts, & Goodell, 2001, p. 775). This gradient spans the earlier categorization taxonomy and is derived from the interaction of two factors: how fittingly the two sounds map to native phonetic categories, and how disparate the two sounds are from each other within phonetic space. This “discriminability gradient” is ideal for assessing idiosyncratic sensitivity.

The “impressive unexplained individual differences” (Jenkins, Strange, & Polka, 1995, p. 320) that seem to sporadically populate cross-linguistic comparison studies cannot be accounted for once age of immersion, length of immersion, and language(s) have been controlled for. This is not the case with contrast-pair difficulty, as well as early-/late-learner differences, for which there are successful explanations backed with supporting evidence. Best’s (1994, 1995; Best & Strange, 1992) perceptual assimilation model and Burnham’s (1986) robust-fragile hypothesis accurately describe differences in discriminability of L2 contrast pairs, whereas Flege’s L2 model (1995; presented in an early form as the “phonological translation hypothesis,” 1981) and Kuhl’s (1993; Kuhl & Iverson, 1995) native language magnet model integratively explain child-adult differences in acquisition suc-

cess. However, no cognitive framework, and almost no empirical research, has yet accounted for the differential sensitivity and acquisition of L2 speech sounds among late-learners. To meet the objective of an instrument with which to scrutinize idiosyncratic differences, I have developed the following expanded set of L2 contrasts intended to range in perceptual discriminability and designed to span Mandarin and English inventories.

**Stimulus Generation**

In order to maximize the accuracy of the L1 categories, the first stage of stimulus production involved deciding upon phonemic inventories (Table 1) through a cross-comparison of several phonology sources for Mandarin Chinese<sup>2</sup> (Duanmu, 2000; Ladefoged & Wu, 1984; Li, 1999; Maddieson, 1984), including the UPSID451 database (Maddieson & Precoda, 1990), and for American English (Edwards, 1997; Ladefoged, 1999, 2005). The question of what constitutes the complete phonemic inventory of a language is often difficult and contentious; this is especially the case with Mandarin (Duanmu, 2000; Li, 1999). Furthermore, there is also the question of how to define any particular category with respect to its constituent allophones and dialectal variations. This is manifest in the famous English [ɹ]:[ɻ] contrast for the Japanese /r/ category, in which the status of each of the three categories and their phonetic realizations are somewhat uncertain among phoneticians and phonologists (for a review, see Jenkins et al., 1995). With regard to Mandarin, although most

sources are in agreement for the majority of consonants, there are a few notable problems with respect to category identity. Despite the uncertain status of /ŋ/ as either a phonetic category or a category variant, it was included here as its own category. Next, the highly contentious existence of /ɹ/ (or /r/) led to its exclusion. Third, although inconsistent in its place of articulation—as /x/, /χ/, or /h/—the backmost voiceless fricative was included as /x/. Finally, the family of /tɕ, tɕʰ, ɕ/ was determined to be allophonic with one or more other categories, and thus excluded. With respect to English, the fairly questionable, or at least dialectal, phonemic status of the proposed categories /x, ʌ, ɲ/ led to their exclusion.

The problem of category definition is an even bigger problem for vowels in a tonal language such as Mandarin, in which phonetic category identity is blurred by extreme allophony, the presence of syllabic consonants and diphthongs, and the fact that linguistic estimates of the monophthong inventory size have ranged from zero to seven (Duanmu, 2000; Mok, 2005). The final inventory included only five monophthongs (the modal inventory size), since the remaining two were poorly agreed upon in characterization or existence. With respect to these five, although there is almost no disagreement with regard to the phonetic categories /i, y, u/, the single mid and low vowels are fairly well disputed. Ultimately, the inventory in this study included /ɜ/ as the modal mid vowel category and a category that is labeled here as /ɘ/ (consistent with the UPSID451 notational system, representing a low, cen-

**Table 1**  
**Inventory of Phonetic Categories**

Mandarin Only			
p <sup>h</sup>	[vc'less aspirated bilabial plosive]	ɕ	[vc'less retroflex sibilant fricative]
t <sup>h</sup>	[vc'less aspirated dental/alveolar plosive]	x	[vc'less velar fricative]
k <sup>h</sup>	[vc'less aspirated velar plosive]	ɻ	[vc'd dental/alveolar lateral approximant]
tɕ	[vc'less dental/alveolar sibilant affricate]	ɥ	[vc'd labial-palatal approximant]
tɕ <sup>h</sup>	[vc'less aspirated dental/alveolar sibilant affricate]	ɹ/ʒ	[vc'd retroflex sibilant fricative]
tʂ	[vc'less retroflex sibilant affricate]	y	[high front rounded vowel]
tʂ <sup>h</sup>	[vc'less aspirated retroflex sibilant affricate]	ɤ	[higher mid back unrounded vowel]
ʂ	[vc'less dental/alveolar sibilant fricative]	ɘ	[low central unrounded vowel]
Mandarin and English			
p	[vc'less bilabial plosive]	ŋ	[vc'd velar nasal]
t	[vc'less dental/alveolar plosive]	j	[vc'd palatal approximant]
k	[vc'less velar plosive]	w	[vc'd labial-velar approximant]
f	[vc'less labio-dental fricative]	i	[high front unrounded vowel]
m	[vc'd bilabial nasal]	u	[high back rounded vowel]
n	[vc'd alveolar nasal]		
English Only			
b	[vc'd bilabial plosive]	ʒ	[vc'd palato-alveolar sibilant fricative]
d	[vc'd alveolar plosive]	h	[vc'less "h"]
g	[vc'd velar plosive]	ɹ	[vc'd alveolar approximant]
tʃ	[vc'less palato-alveolar sibilant affricate]	l	[vc'd alveolar lateral approximant]
ɟʂ	[vc'd palato-alveolar sibilant affricate]	ɪ	[lowered high front unrounded vowel]
v	[vc'd labio-dental fricative]	ʊ	[lowered high back rounded vowel]
θ	[vc'less dental fricative]	e	[lower mid front unrounded vowel]
ð	[vc'd dental fricative]	ʌ	[lower mid back unrounded vowel]
s	[vc'less alveolar sibilant fricative]	ɔ	[lower mid back rounded vowel]
z	[vc'd alveolar sibilant fricative]	æ	[raised low front unrounded vowel]
ʃ	[vc'less palato-alveolar sibilant fricative]	ɑ	[low back unrounded vowel]

Note—Phonetic descriptions are based on the UPSID451 system. The "dental/alveolar" parameter indicates segments for which either there is no exact specification for the place or there are conflicting claims. This uncertainty is notated as parentheses around the dental diacritic.

tral, unrounded vowel), which is fairly evenly in between the almost equally named /a/ and /ɑ/ categories. The final vowel inventory did not include the syllabic consonant /ɹ/ (or /z̥/), since most Mandarin scholars believe this sound (along with /z̥/) is a variant of /i/, and since very few UPSID languages incorporate this category. With regard to English, the only questionable monophthong, which was ultimately excluded, was the category /ə/, which occurs more frequently as a “reduced” allophone of other categories than it does on its own. Diphthongs and syllabic consonants in both languages were excluded, on the grounds that they are often considered allophones and that very few languages within the UPSID451 possess them or versions of them similar to the target those in the target languages.

Once the final Mandarin and English inventories were established, they were parameterized on the basis of the phonetic characterizations employed in the UPSID451 (Mandarin already existed, but was edited in some instances because of the assessments above). All languages within the database were then searched exhaustively for instances of the Mandarin and English phonetic categories. Each language was then further searched for “CG counterparts”: sounds that could potentially be categorized as being perceptually equivalent to the Mandarin phonetic category. From the literature review above, this was operationally defined as a second contrastive segment that differed with respect to the matching Mandarin segment by either a single primary and/or secondary phonetic feature. The search was constrained to include all but click sounds (velaric ingressesives), given the past findings by Best and colleagues (Best & Avery, 1999; Best, McRoberts, & Sithole, 1988) that clicks are very resistant to assimilation in languages that lack them. In addition to CG pairs, the database was also searched for pairs that would have a high likelihood of being perceived as SC correspondences. Only a few of these were produced. Recurrences of a particular contrast among more than one language were tabulated as a count score and stored as a measure of the robustness of the contrast. Given Burnham’s (1986) theory that the more salient or perceptible a contrast, the more widely it occurs among the world’s languages, along with the subsequent findings by Best (1995) and colleagues that back this view, contrast pairs with the lowest frequency counts within the database should be considered as marginally distinguishable, and thus should be excluded as being too difficult.

Consequently, the rather large output of contrasts was first trimmed by removing all L2 contrasts, by setting a frequency cutoff of seven or more of the UPSID languages. The number seven was chosen for practical reasons of ultimate set size and because there is currently no metric for robustness based on frequency. Next, all contrast pairs from English that mapped onto a Mandarin category, and vice versa, were eliminated from the target population, given that the stimuli were created for Mandarin–English bilinguals. Finally, the frequency of segments was trimmed for redundancy to maintain the widest sampling possible of sounds and to balance consonant and vowel occurrences. The final set of contrasts is composed of 8 vowel and 16 consonant pairs, corresponding to 5 vowel

and 15 consonant categories in Mandarin only, 12 vowel and 24 consonant pairs (mapping to 2 vowel and 8 consonant categories) in Mandarin and English, and 10 vowel and 20 consonant pairs (parallel to 5 vowel and 11 consonant categories) in English only (see Appendix A, included in the stimulus archive).

The carrier CVC syllables are composed in pairs in which only one of the segments is contrastive, whereas the others temporarily perform a noncontrastive role (Figure 2). This role then rotates around, so that all segments function contrastively and noncontrastively in turn. When serving noncontrastively, it is the nonnative partner of each pair that fills this role within the CVC. In cases of an “even split,” the noncontrastive role is chosen at random. The contrastive/noncontrastive positions are randomly permuted to produce the final set of 180 CVC contrast pairs organized into three groups: 48 Mandarin only, 72 English–Mandarin, and 60 English (see Appendix B, included in the stimulus archive). Each consonant contrast occurs in onset and coda position three times, and each vowel occurs as the nucleus six times; each pair occurs contrastively one third of the time within their respective language group.

In order to fulfill the discrimination design, two tokens of each CVC (four per pair), for a total of 720 CVC utterances total, were produced in a single, male voice by a phonetically trained speaker, using the International Phonetic Association prototypical description of each phone (Ladefoged, 1999, 2005; Ladefoged & Maddieson, 1996). Although this procedure is not ideal with regard to faithfulness to each L2 phonetic category, it would not have been neither practical nor logical to record the speech sounds of each contrast from native speakers. Each CVC is a composite of contrasts from different languages, and therefore would have had to be either spliced together or uttered by a single speaker. Splicing would have been detectable and also would have carried traces of the original phonetic environment, thus violating the new environment. Furthermore, the speakers of any given L2 could have been expected to produce authentically the phones of another language without the appropriate training, and even the coarticulatory effects of the surrounding phonotactically foreign CVC environment would have removed any genuineness. In fact, authentic phonetic realizations of each L2 category are not essential in order for a pair of L2 phones to assimilate to a single L1 category. Crucially, in order to effectuate assimilation, it should be sufficient if each phonation falls very near the target phonetic category. Finally, the aim of the creation of these stimuli was not to investigate the linguistic discriminability of the contrast pairs themselves (as has been widely done already), but rather to examine the behavioral sensitivity of the nonnative listeners.

Tokens were recorded digitally in an anechoic room through a single channel at a sampling rate of 44,100 and stored with a depth of 16 bits in the WAV format utilizing a Shure 588SD unidirectional cardioid microphone through a Shure M268 mixer for preamplification. Subsequently, the tokens were postprocessed, first by removing any dead air before and after the utterance, then by high-pass filtering at 55 Hz to remove electronic noise, and finally by

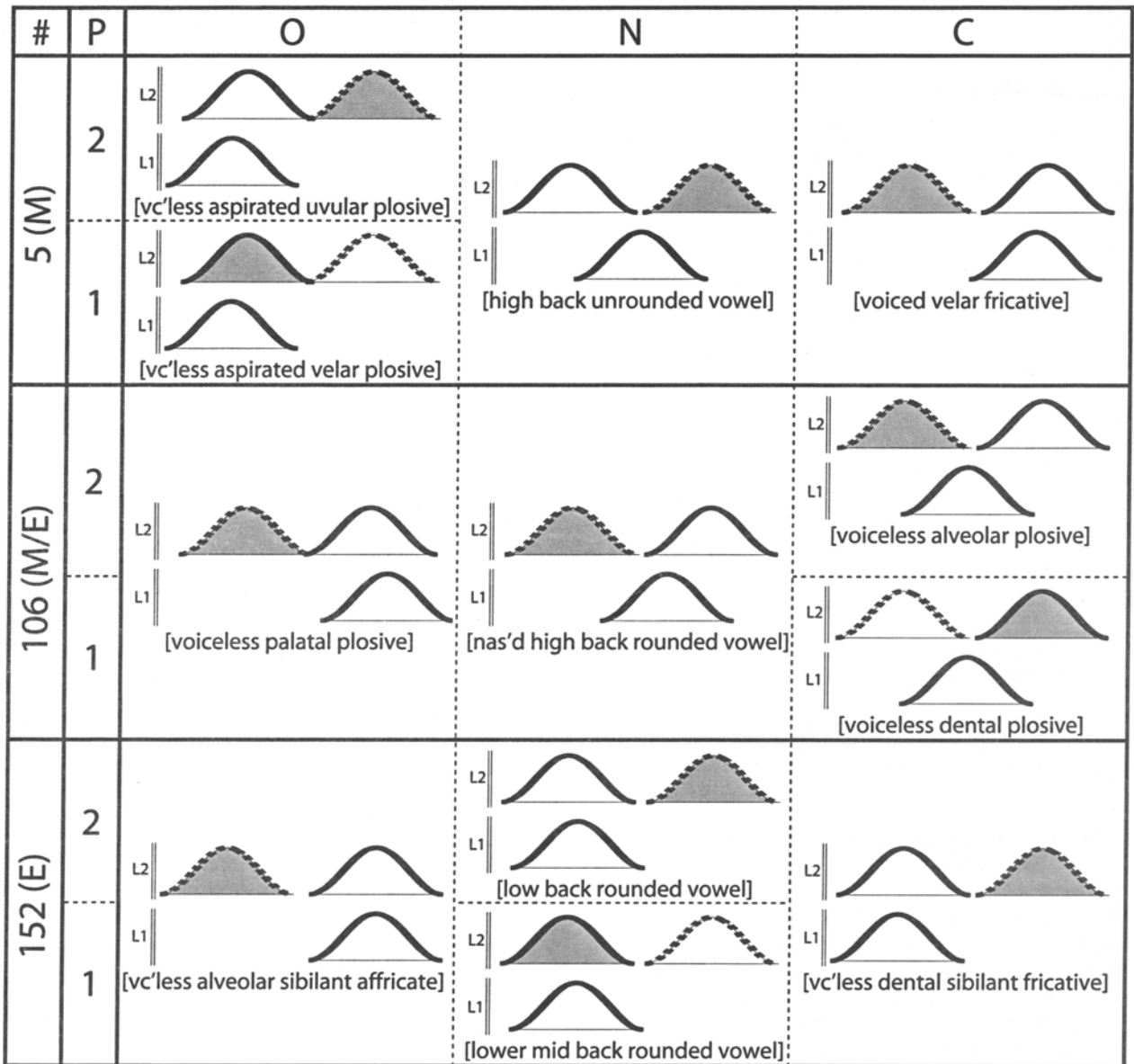


Figure 2. Examples of CVCs demonstrating single-category and category goodness types of contrast in onset (O), nucleus (N), and coda (C) positions, for Mandarin-only (M), Mandarin-English (M/E), and English-only (E) stimulus subgroups. Solid outlines represent sounds that correspond best to the target phonetic categories; in the case of the single-category stimuli, this distinction is arbitrary. The numbers at the left correspond to the set list in the online Appendix B.

employing a noise reduction algorithm with a gate level of -40 dB to remove any equipment hiss or hum; none of these procedures degraded the recordings noticeably.

**Conclusion**

I have described a set of cross-linguistic speech contrast stimuli drawn from an array of languages and designed for Mandarin speakers who may or may not have experience with the English sound system. More specifically, phonetic contrast pairs were selected from a wide range of the world's languages in such a manner that between-category phonetic differences from the sample language correspond to within-category differences in Mandarin,

English, or both languages. These stimuli are currently being employed in the investigation of differential late-learner L2 accent as measures of idiosyncratic sensitivity for nonnative speech contrasts, wherein ability is assessed through the range of discriminability that these stimuli offer with respect to a single L1. In general, this research measures performance by comparing judgments<sup>3</sup> between a CVC matched (1) with its L2 phonetic category contrast partner, (2) with the second token of the same L2 phonetic category instancing a slightly different waveform, or (3) with itself as identical waveforms (Figure 3).

These contrasts could be investigated using alternative discrimination, identification, or even imitation tasks.

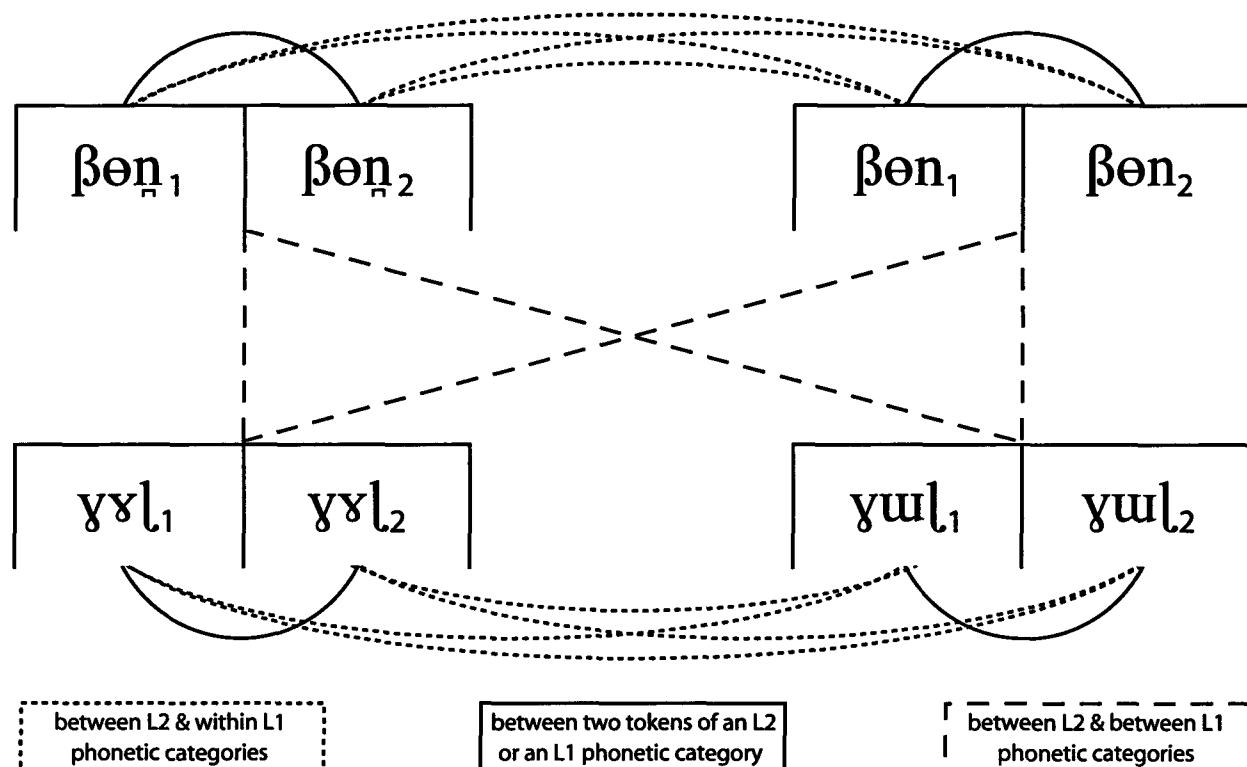


Figure 3. The CVC stimuli were designed for discriminations “between L2 and within L1 phonetic categories” and “between two tokens of an L2 or an L1 phonetic category,” but they could also be used for comparisons “between L2 and between L1 phonetic categories” or (not displayed) “any item paired with itself” as a control/catch condition.

One might also compare and contrast them with whatever method by means of other possible configuration patterns, as shown in Figure 3. One might easily employ them to study early Mandarin learners of English. In addition, although they were designed to test Mandarin (L1)–English (L2) speakers, by splitting up the set they could be equally well suited to probing individual differences among native English or Mandarin speakers who have learned other L2s. Linguistically speaking, they could also be used to examine the correspondences between phonetic features and variable contrast difficulty/robustness.

Outside the domain of cross-linguistic comparison research, these stimuli have the potential to be utilized in other experimental frameworks, such as lexical acquisition, nonword processing, various learning and memory designs, as well as neuroscientific investigations of brain activation in response to speech sounds. The latter proposals may be especially appropriate if one were to separate out the stimuli that correspond to English phonetic categories. Therefore, I hope that these CVC stimuli will be a resource for other researchers, whatever their undertakings.

#### Availability

The CVC syllable stimuli described in this article may be accessed by contacting the author at [mmunro@uci.edu](mailto:mmunro@uci.edu). A link to a ZIP file of the entire set, along with appendices, will be sent in the reply.

Two appendices are included with the CVC stimuli. Appendix A details the inventory of phonetic contrast pairs

used to construct the CVCs, specifying the segments as IPA symbols and in terms of phonetic parameters. This appendix also specifies which language(s) the “native” segments of each contrast pair occur in, whether the segment is a consonant or vowel, and the frequency with which each contrast occurs in a language within the UPSID451 database.

Appendix B lists, for all of the CVC pairs, the number of the matching WAV files, the corresponding language(s) from which the target native category segment was drawn, and the syllable position of the contrast.

#### AUTHOR NOTE

I thank Mary-Louise Kean, Mark Steyvers, and Bernard Tranel for their useful discussions in the design of these stimuli. I also thank Teresa Griffith for assistance in identifying and procuring several of the Mandarin phonological resources, as well as Kenny Vaden and Bruce Berg for their assistance in providing recording equipment and space. Correspondence relating to this article may be sent to M. Munro, Department of Cognitive Sciences, 3151 Social Science Plaza A, University of California, Irvine, CA 92697-5100 (e-mail: [mmunro@uci.edu](mailto:mmunro@uci.edu)).

#### REFERENCES

- AOYAMA, K., FLEGE, J. E., GUION, S. G., AKAHANE-YAMADA, R., & YAMADA, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32, 233-250.
- BEST, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

- 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 research* (pp. 171-206). Baltimore: York Press.
- BEST, C. T., & AVERY, R. A. (1999). Left-hemisphere advantage for click consonants is determined by linguistic significance and experience. *Psychological Science*, 10, 65-70.
- BEST, C. T., McROBERTS, G. W., & GOODELL, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109, 775-794.
- BEST, C. T., McROBERTS, G. W., & SITHOLE, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception & Performance*, 14, 345-360.
- BEST, C. T., & STRANGE, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, 20, 305-330.
- BURNHAM, D. K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics*, 7, 207-240.
- DAI, H., VERSFELD, N. J., & GREEN, D. M. (1996). The optimum decision rules in the same-different paradigm. *Perception & Psychophysics*, 58, 1-9.
- DUANMU, S. (2000). *The phonology of standard Chinese*. Oxford: Oxford University Press.
- EDWARDS, H. T. (1997). *Applied phonetics: The sounds of American English* (2nd ed.). San Diego: Singular.
- FLEGE, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly*, 15, 443-455.
- 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* (pp. 233-277). Baltimore: York Press.
- FLEGE, J. E., & HILLENBRAND, J. (1984). Limits on pronunciation accuracy in adult foreign language speech production. *Journal of the Acoustical Society of America*, 76, 708-721.
- GOTTFRIED, T. L. (1984). Effects of consonantal context on the perception of French vowels. *Journal of Phonetics*, 12, 91-114.
- HAUTUS, M. J., & COLLINS, S. (2003). An assessment of response bias for the same-different task: Implications for the single-interval task. *Perception & Psychophysics*, 65, 844-860.
- JENKINS, J. J., STRANGE, W., & POLKA, L. (1995). Not everyone can tell a "rock" from a "lock": Assessing individual differences in speech perception. In D. Lubinski & R. V. Dawis (Eds.), *Assessing individual differences in human behavior: New concepts, methods, and findings* (pp. 297-325). Palo Alto, CA: Davies-Black.
- KUHL, P. K. (1993). Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics*, 21, 125-139.
- KUHL, P. K., & 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). Baltimore: York Press.
- LADEFOGED, P. (1999). American English. In International Phonetic Association (Ed.), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 41-50). Cambridge: Cambridge University Press.
- LADEFOGED, P. (2005). *Vowels and consonants: An introduction to the sounds of languages* (2nd ed.). Malden, MA: Blackwell.
- LADEFOGED, P., & MADDIESON, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.
- LADEFOGED, P., & WU, Z. (1984). Places of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics*, 12, 267-278.
- LI, W.-C. (1999). *A diachronically-motivated segmental phonology of Mandarin Chinese*. New York: Peter Lang.
- LOGAN, J. S., LIVELY, S. E., & PISONI, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, 89, 874-886.
- MACMILLAN, N. A., & CREELMAN, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- MACMILLAN, N. A., KAPLAN, H. L., & CREELMAN, C. D. (1977). The psychophysics of categorical perception. *Psychological Review*, 84, 452-471.
- MADDIESON, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
- MADDIESON, I., & PRECODA, K. (1990). Updating UPSID. *UCLA Working Papers in Phonetics*, 74, 105-111.
- MCCANDLISS, B. D., FIEZ, J. A., PROTOPAPAS, A., CONWAY, M., & MCCLELLAND, J. L. (2002). Success and failure in teaching the /r/-/l/ contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, & Behavioral Neurosciences*, 2, 89-108.
- MOK, P. K. P. (2005, December). *Reassessing the notion of vowel phonemes in Cantonese and Mandarin*. Paper presented at the 10th International Conference on Yue Dialects, Hong Kong.
- POLKA, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America*, 97, 1286-1296.
- ROCHET, B. L. (1995). Perception and production of second-language sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 379-410). Baltimore: York Press.
- STRANGE, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 3-54). Baltimore: York Press.
- YAMADA, R. A., & TOHKURA, Y. (1992). The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception & Psychophysics*, 52, 376-392.

## NOTES

1. Best (1995; Best, McRoberts, & Goodell, 2001) usually includes two other assimilation categories, uncategorized-categorizable (UC) and uncategorized (UU), included here under the single NA umbrella (as the paradigm was first formulated; Best, 1994; Best & Strange, 1992), since all three involve a failure to assimilate both L2 phones to a single L1 category.

2. The reason for choosing Mandarin as the candidate L1 for generation of these stimuli is threefold: Although a major world language, Mandarin is studied relatively little in the L1/L2 comparison literature; Mandarin phonetic categories (especially those of vowels) are somewhat diffuse, given the tonal emphasis of the language; and finally, the research institution of this author and the local community provide a large investigational population.

3. Future publications will survey the discriminability results, in both *p*(c) and *d'*, obtained with a panel of both Mandarin late-learners of English and English-speaking monolinguals using these stimuli, as well as introduce methods I have derived for sorting and selecting on the basis of these discriminability scores. One note of mention here with regard to detection-theoretic estimations, given the unique nature of these stimuli and any AX discrimination tasks employed to investigate them, I have opted for a same-different analysis under a roving rather than a fixed design (assuming a differencing-over-independent-observations model), although the data are not archetypical of either (Dai, Versfeld, & Green, 1996; Hautus & Collins, 2003; Macmillan & Creelman, 2005; Macmillan, Kaplan, & Creelman, 1977).

(Manuscript received October 12, 2006;  
revision accepted for publication March 15, 2007.)