

Perceiving non-native speech: Word segmentation

Michèle Mondini and Joanne L. Miller

Northeastern University, Boston, MA

BACKGROUND

Word segmentation

Listeners use a number of strategies and sources of information to segment speech into discrete words. One important source of information is allophonic variation at word boundaries.

Allophonic information

Articulatory differences in the way in which phonetic segments are produced within and across word boundaries lead to systematic allophonic variation (Lehiste, 1960).

Listeners are perceptually sensitive to such allophonic variation, especially word-initially, and use it as one source of information for word segmentation (Nakatani & Dukes, 1977).

Non-native speech

Non-native talkers typically impose their native-language phonetic norms onto their second language. This leads to articulatory patterns that depart from the second language's phonetic norms and, as a result, to systematic acoustic differences between native and non-native speech.

Native listeners can detect these acoustic differences, and perceiving non-native speech (foreign-accented speech) can be more difficult than perceiving native speech (Bürki-Cohen, Miller, & Eimas, 2001; Flege, 1984; Munro & Derwing, 1995).

CURRENT EXPERIMENT

The current experiment is part of a larger project that examines whether word segmentation problems contribute to the overall difficulty of perceiving non-native speech.

The experiment focuses on the allophones of the English voiceless stop consonants /p t k/, produced by native speakers of English and native speakers of French.

In English, word-initial voiceless stop consonants are aspirated with relatively long voice-onset-times (VOTs) (Lisker & Abramson, 1964). These long VOTs serve as a marker that the consonant is word-initial.

In contrast, word-initial voiceless stop consonants in French are unaspirated and are produced with much shorter VOTs (Delattre, 1965; Caramazza & Yeni-Komshian, 1974).

As a consequence of this language difference, French-accented English is typically characterized by word-initial VOTs that are somewhat shorter than those of native English speech (Flege, 1984).

We hypothesized that this would result in word segmentation problems for native English listeners when listening to word-initial English voiceless stop consonants that were produced by native French speakers. The current experiment tested this hypothesis.

METHOD

Stimulus material

- 3 sets for each of the juncture consonants /p/, /t/, and /k/
- Each set contained:
 - 2 word pairs with a contrastive juncture consonant: C#V and V#C
 - 1 word pair with a geminate juncture consonant: C#C
 - 1 word pair with no juncture consonant: V#V

Stimulus Type

	C#V	V#C	C#C	V#V
/p/	wipe ink grape ale keep art	why pink gray pail key part	wipe pink grape pail keep part	why ink gray ale key art
/t/	neat ape great eyes might ache	knee tape gray ties my take	neat tape great ties might take	knee ape gray eyes my ache
/k/	bike oil weak ash make oat	buy coil we cash may coat	bike coil weak cash make coat	buy oil we ash may oat

Talkers, recording procedures, and stimulus sets

- Three female and three male native English talkers, and three female and three male highly fluent, native French talkers who learned English as a second language, were recorded saying numerous instances of the word pairs.
- The word pairs were spoken in the neutral sentence frame "he writes..." and were excised for presentation to listeners.
- Two stimulus sets, Native-English and French-Accent, were created. Each set consisted of 216 stimuli (one instance of each of the 36 word pairs x 6 talkers).
- Overall duration of the word pairs was matched across the two stimulus sets.

Listeners and testing procedure

- Forty native English listeners served as subjects. Twenty were tested on the Native-English stimulus set, and twenty were tested on the French-Accent stimulus set.
- Listeners were instructed to identify the word pair they heard on each trial by making a selection among four possible choices displayed on a computer screen.
 - For example, if the auditory stimulus was *buy coil*, the four choices on the screen were *buy coil*, *bike oil*, *bike coil*, *buy oil*.
- Stimuli were presented and responses were collected using Psyscope software.
- Percent correct was taken as a measure of word segmentation.

PREDICTIONS

Overall prediction

Word segmentation should be better for the native-English speech than the French-accented speech.

Specific predictions

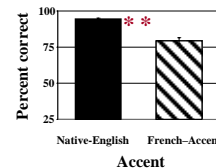
Given expected differences in VOT for word-initial consonants across native-English and French-accented speech, coupled with the known importance of word-initial information for word segmentation:

- Performance on stimuli with an intended word-initial junctural consonant (the V#C and C#C stimuli) should be most affected by the French accent.
- Performance on stimuli with no junctural consonant (the V#V stimuli) should be least affected by the French accent.
- Performance on the C#V stimuli would likely fall somewhere in between.

RESULTS

Mean percent correct across trials was computed for each listener, for each of the four stimulus types.

The means were submitted to a two-way ANOVA, Accent (Native-English, French-Accent) X Stimulus Type (C#V, V#C, C#C, V#V), which yielded significant main effects of Accent and Stimulus Type, and a significant Accent X Stimulus Type interaction ($p < .0001$ in each case).



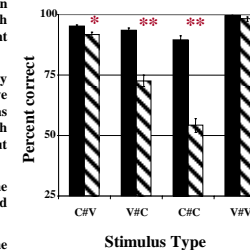
The graph on the left displays the main effect of Accent. Listeners were 94% correct on the Native-English stimuli and 79% correct on the French-Accent stimuli.

The graph on the right displays mean percent correct separately for each stimulus type, for each of the two accent conditions.

For native-English speech, accuracy across all stimulus types was above 89%, and performance was significantly better for native-English than French-accented speech for all but the V#V stimuli.

Moreover, the largest effect of the French accent was for the V#C and C#C stimuli.

Thus, there was an overall effect of the French accent on word segmentation, with stimuli containing an intended word-initial junctural consonant being especially problematic.



Stimulus Type

* $p < .001$, ** $p < .0001$

CONCLUSION

The results of this experiment show that native English listeners are substantially more accurate in segmenting English words produced by native English speakers than in segmenting English words produced by native speakers of French.

Thus one contribution to the processing cost associated with listening to non-native speech may be problems with word segmentation, due at least in part to the presence of altered allophonic information at word boundaries.

We are currently carrying out acoustic measurements on the speech of both the native English talkers and the French-accented talkers to explore in detail the relation between altered allophonic information at word boundaries and errors in word segmentation.

Additional word segmentation experiments currently underway:

- compare the acoustic properties and processing costs associated with French-accented versus German-accented speech.
- investigate the differential effects of background noise on word segmentation for native and non-native speakers.
- investigate whether processing speed, as well as accuracy, is affected by a foreign accent.
- Taken together, such studies should contribute to our understanding of how detailed acoustic-phonetic information influences speech processing in both native and non-native listeners.

REFERENCES

- Caramazza, A. & Yeni-Komshian, G. (1974). Voice-onset-time in two French dialects. *Journal of Phonetics*, 2, 239-245.
- Bürki-Cohen, J., Miller, J.L., & Eimas, P.D. (2001). Perceiving non-native speech. *Language and Speech*, 44, 149-169.
- Delattre, P. (1965). *Comparing the phonetic features of English, French, German and Spanish: An Interim Report*. Groos: Heidelberg.
- Flege, J.E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, 76, 692-707.
- Lehiste, I. (1960). An acoustic-phonetic study of internal open juncture. *Phonetica, Supplementum ad Vol. 5*, 1-54.
- Lisker, L. & Abramson, A.S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.
- Munro, M.J. & Derwing, T.M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38, 289-306.
- Nakatani, L.H. & Dukes, K.D. (1977). Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America*, 62, 714-719.

ACKNOWLEDGEMENTS

Supported by NIH grant DC00130 from the National Institute on Deafness and Other Communication Disorders.