Andries W. Coetzee and Kevin B. McGowan. **University of Michigan**

Allophonic cues to syllabification

Allophonic realization rules are often stated in terms of syllable structure. For instance, in English /l/ is "dark" in codas and "light" in onsets. In this study, we show that attention to positional allophones influences how listeners parse an acoustic signal into syllables.

To show this, we use the positional allophones of English voiceless stops. Rule (1) gives a simplified statement of the realization of /p, t, k/, and captures the fact that aspirates appear only in syllable initial position. If listeners use this kind of information when parsing acoustic signals into syllables, they could use aspiration on a stop as a cue to insert a syllable break directly before the stop.

We assume that when a listener perceives an acoustic signal, his/her perceptual system considers different perceptual hypotheses that differ *inter alia* in syllable structure. If it is possible to parse the acoustic signal faithfully into a legal syllable structure, the perceptual system is likely to settle on an acoustically faithful percept. However, if no syllabically legal parse is possible for a percept that is faithful to the acoustic input signal, then the perceptual system will opt for a percept that is unfaithful to the acoustic stimulus, but that can be parsed into legal syllables. Syllabic well-formedness of the percept is hence more important than identity between the percept and the acoustic input signal. One of the acoustic manipulations that the perceptual system can perform is perceptual epenthesis, where the listener "perceives" a vowel that is not acoustically present (Dupoux *et al*. 1999).

Table (2) shows the hypotheses that would be entertained for acoustic inputs containing [s + aspirated/ unaspirated stop] in word-initial and word-medial position. In "aspirated" forms, an acoustically faithful percept can be parsed into legal syllables only for "medial" tokens, where the [s] can be parsed into the coda of the preceding syllable. This is not possible for the "initial" tokens. The perceptual system will hence be more likely to settle on an acoustically unfaithful percept for the "initial aspirated" than the "medial aspirated" forms. In "unaspirated" forms, a syllabically well-formed parse of an acoustically faithful percept is possible both for "initial" and "medial" tokens – since the unaspirated stop can be preceded by an [s] in the same syllable. An acoustically faithful percept is hence expected to be equally likely for both "initial unaspirated" and "medial unaspirated" tokens.

To test these predictions, we conducted a same/different experiment in which we presented listeners with non-word pairs. One member of each pair had the sequence [s + schwa + aspirated stop]. The other member lacked a schwa, but either had or lacked aspiration on the stop. Table (3) contains representative examples, and shows if we expect more "same" or "different" answers for each pair. We calculated d'-scores for each condition – d' is an index of the perceptibility of a difference between two forms. Figure (4) represents results. The confidence intervals of d' for "unaspirated" tokens in "initial" and "medial" position overlap, indicating that there is no perceptibility difference for "unaspirated" tokens in the two positions. However, d' for "aspirated medial" tokens is significantly higher than that for "aspirated initial" tokens, showing that perceptual epenthesis was more likely in "initial" than "medial" position for "aspirated" tokens. This can be explained if we assume that listeners inserted a syllable break before aspirated stops, which necessitates perceptual epenthesis only in initial position.

These results show that English listeners do rely on the positionally determined allophones when parsing an incoming acoustic signal into syllables. When an acoustic signal contains an aspirated stop, this is taken as evidence of the start of a new syllable. In an acoustic signal like [sp^{h}ika], this results in stranding the word-initial [s], which is resolved by the perceptual system by means of perceptual epenthesis. Consequently, a [sp^{h}ika] acoustic signal is unfaithfully perceived as [sə.p^{h}i.ka].

More generally, these results show that the mapping from acoustic signal to linguistic percept includes amongst other things the parsing of the signal into syllable structure. It also provides evidence for the psychological reality of syllables and of allophonic realization rules. It opens up new avenues for investigating syllable structure generally, as well as the properties of the linguistic perceptual system.

(1) **Some positionally determined allophones of /p, t, k/**

/p, t, k/ → [p^h, t^h, k^h] / σ[__ …

[p, t, k] / σ[s __ …

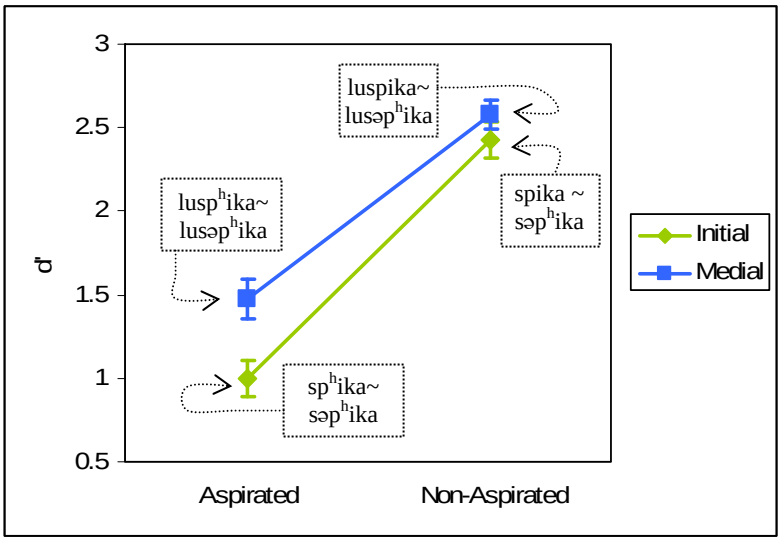
(2) **Perceptual hypotheses considered** (blacked cells mark syllabic ill-formedness)

			Perceptual hypotheses		
			Acoustically faithful		Perceptual epenthesis
Acoustic signal			Syllable break before stop	Syllable break before [s]	Syllable break before stop
Aspirated	Initial	[sp^{h}ika]	s.p^{h}i.ka	sp^{h}i.ka	sə.p^{h}i.ka
	Medial	[lusp^{h}ika]	lus.p^{h}i.ka	lu.sp^{h}i.ka	lu.sə.p^{h}i.ka
Unaspirated	Initial	[spika]	s.pi.ka	spi.ka	sə.pi.ka
	Medial	[luspika]	lus.pi.ka	lu.spi.ka	lu.sə.pi.ka

(3) **Example of tokens used in experiment**

	Token 1	Token 2	Expected	Reason
[+asp]	lusəp^{h}ika	lusp^{h}ika	Different	Legal parse for Token 2 possible
	səp^{h}ika	sp^{h}ika	Same	No legal parse for Token 2. Perceptual epenthesis expected.
[-asp]	lusəp^{h}ika	luspika	Different	Legal parse for Token 2 possible
	səp^{h}ika	spika	Different	Legal parse for Token 2 possible

(4) **Mean d'-scores in the four conditions – error bars show 95% confidence intervals**



References

Dupoux. Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier and Jacques Mehler. (1999) Epenthetic Vowels in Japanese: A Perceptual Illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25(6):1568-1578.