If segments are real, what are they?

Sean A. Fulop
Chris Golston
California State University Fresno

Three kinds of ‘real’

Physically real:
identifiable in motor production or in segments of the acoustic speech signal

Theoretically real:
needed to adequately model observable grammatical patterns (e.g., reduplication).

Cognitively real:
discoverable in an experimental setting, language processing, language acquisition, etc.

Can we find a physical speech segment?

• The answer has been “no” for over 100 years.

Can we find a physical speech segment?

Can we find a physical speech segment?

Can we find a physical speech segment?

Can we find a physical speech segment?

Are segments theoretically important?

No – all linguistic facts derived using segments can be derived without them, by relying solely on distinctive features and prosodic structure. (We think.)

An affricate is definable as a particular sequencing of distinctive features within a syllable onset or coda.

On this account, the codas of batch and bats would be described identically, modulo place.

German rhymes an affricate [ts] with derived [t+s]

Hops and malt,
God save it!
Oowekyala (Wakashan, Pacific NW) does have the propensity to contrast affricates with corresponding stop + fricative sequences (Howe 2000, Lincoln and Rath 1980)

[tsala] ‘to cut through water’
[tʰsala] ‘pushing’

But Oowekyala also disallows most obstruent clusters within syllable components, generally giving each obstruent its own piece of a syllable or even its own syllable entirely (Howe 2000):

[tsa.la] ‘to cut through water’
[tʰsɔ.la] ‘pushing’

This is not coincidental, but is because segments aren’t real in Oowekyala, though they may seem to be at first glance.

Are segments cognitively real?

Literature on child development, adult psycholinguistics, and speech errors suggests ‘probably not’.

Jusczyk (1992) summarizes the earlier literature on infant speech perception; numerous studies found little evidence that infants organize speech into segments.

‘Currently, there are no data that demonstrate that infants actually do perceive similarities between segments occurring in different positions’ within syllables.

‘Children may not make such associations much before they learn to read.’

Experiments in adults suggest that different languages may have different basic perceptual units.

The syllable is most important in French
(Cutler et al. 1986)

A stress-based foot is most important in English
(Cutler and Norris 1988)

No language has been definitively shown to have the segment as a perceptual unit.

Cheung et al. (2001) cite literature on children doing onset deletion tasks: it was easier to delete an entire onset than part of one, even when the two deletion targets were the same.

Easier to delete
[s] from [sep] (ie, delete an onset), than
[s] from [slap] (ie, delete a ‘segment’)

When literature can be found that argues for the cognitive reality of segments, closer analysis generally reveals that the ‘evidence for segments’ is equally ‘evidence for prosody’.

E.g. Werker and Pegg (1992) state that if a child can discriminate bog from dog, we have evidence that they represent the difference between a bilabial and an alveolar consonant.

Of course this only demonstrates the ability to distinguish bilabial from alveolar place of articulation. It does not show whether there are distinct segments, distinct onsets, distinct syllables, etc.
O’Seaghdha et al. (2010) claim that English and Dutch speech perception benefits from preparation with ‘the first segment of the syllable.’

They purport to argue the importance of segments in those languages; but all segments tested were also onsets. The findings are equivocal.

<table>
<thead>
<tr>
<th>Set</th>
<th>[Ø]</th>
<th>[W]</th>
<th>[Y]</th>
<th>[J]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneous</td>
<td>Night-day</td>
<td>Swap-boy</td>
<td>Give-up</td>
<td>Pig-cook</td>
</tr>
<tr>
<td>1</td>
<td>Year-day</td>
<td>Crack-pow</td>
<td>Light-yay</td>
<td>Stay-now</td>
</tr>
<tr>
<td>2</td>
<td>Bread-squash</td>
<td>Cork-paw</td>
<td>Tennis-man</td>
<td>Cape-right</td>
</tr>
<tr>
<td>3</td>
<td>Meet-lake</td>
<td>High-paw</td>
<td>Like-tune</td>
<td>Tail-nay</td>
</tr>
<tr>
<td>4</td>
<td>Meet-take</td>
<td>High-pace</td>
<td>Like-purr</td>
<td>Tail-nay</td>
</tr>
</tbody>
</table>

Most errors are ambiguous, don’t tell us much (Brown 2004).

Fish grotto  ➔  frish grotto  move r or move [RTR]

Brake fluid  ➔  blake fluid  swap r/l or move [lateral]

No speech errors unambiguously swap segments
  *fits ➔  fist, *stomp ➔  tamps
  *split ➔  split, *split ➔  plist

But some do unambiguously swap features
  clear blue sky  ➔  glear blue sky (Fromkin, 1973)
  And some do unambiguously swap entire onsets
  squeaky floor  ➔  fleaky squoor (Fromkin, 1973)

Evidence for prosody, segments, features

Widely cited speech error findings are supposed to show that errors usually affect segments. (Shattuck-Hufnagel & Klatt 1979)

But much of this is sub-segmental:

‘[E]rrors which have been consigned to the phonemic, segmental, or feature levels could be reinterpreted as errors at the motor output level.’ (Mowrey & McKay 1990:1311)

Gradient activation results in partial or incomplete gestures. (Pouplier, Chen, Goldstein & Byrd 1999)

‘Unambiguous cases seem to support the psychological reality of distinctive features...and prosodic constituents like onsets... The remaining number of cases, which seem to be ambiguous in regard to unit activity (i.e., segments or syllable constituents) because of frequency of prosodic distribution, seem to be explained better in terms of onsets, nuclei, and codas...’ (Brown 2004:100)

Feet

stress is assigned by feet
  (Hayes 1980; 1995; Halle & Vergnaud 1987)
words are minimally a foot in many languages
  (McCarthy & Prince 1986, 1990; Crowhurst 1992)
  [µµ] in English, Latin, Greek, Fijian, etc.
  [no] in Mohawk, Pitta-Pitta, Wangkumara, etc.
reduplication copies a foot in many languages
  (McCarthy & Prince 1990)
feet are domains for phonological processes
  (Kiparsky 1982; Hayes 1982; Jensen 2001)
Syllables

lexical access/storage uses syllable count
(Cutler 1986; Levelt 1989)
tip-of-tongue state matches in terms of syllable count
(Brown & McNeill 1966; many others)
malapropisms match in terms of syllable count
(Fromkin 1973; Fay & Cutler 1977)
word, root, affix minima/maxima
(McCarthy & Prince 1986; Golston 1991 etc.)
language games that reverse syllables
(Clements 1986)

Syllables act as physiological unit of organization
(Krakow 1989; Browman & Goldstein 1995)
in articulation
(Fujimura 1990; Sproat & Fujimura 1993; Krakow 1999)
in word recognition
(Wingfield et al. 1997; Lindfield et al. 1999)
in acquisition by children and, arguably, the species

can be counted by native speakers
(Japanese)
form morabararies (syllabaries based on moras)
(Poser 1992)
are conserved in second language acquisition
(Broselow & Park 1995)
are conserved in compensatory lengthening
(Hayes 1989)

poetic meter based on syllables
(Chinese, Vietnamese, Hmong)
phonotactics
(too many to begin to list)
syllabaries
(cuneiform, Cherokee, etc.)
contrastive syllabification
(Clements 1986b; Ladefoged et al. 1998; Bosch 1998)
children’s URs seem to be syllabified
(Curtin 2001)
native speakers’ ability to count syllables
(except Japanese, where moras are more salient)

Moras

attract stress
(Trubetzkoy 1939)
model consonant and vowel length
(McCawley 1968)
model phonetic length
(Hubbard 1995)
license extraprosodic sounds
are used in language games and poetic meter
(Shiritori, Haiku, etc.)

Onsets

universally preferred, even by children
(Jakobson 1962)
commonest locus for speech errors
(Shattuck-Hufnagel 1986)
used in language games and poetic meter
(Pig Latin, Ubby-Dubby, Alliteration)
used in blends
(Bat El 2006)
Rhymes

phonotactic domain
(VV or VC in many languages)
domain for phonological processes
(nasalization, etc.)
used in language games and poetic meter
(Pig Latin, Ubby-Dubby; Rhyming)
used in blends
(Bat El 2006)

Segments

?

Features/Gestures

distinguish minimal pairs

spread in assimilation, dissimilation, harmony, etc.
directly observable (as gestures)

the bedrock of speech

Are segments physically real?
identifiable in motor production or in segments of
the acoustic speech signal? No.
Theoretically real?
needed to adequately model observable
grammatical patterns? Doesn’t look like it.

cognitively real?
discoverable in an experimental setting,
language processing, language acquisition?
Doesn’t look like it.

References (partial)