

Conclusion

Phonotactic typology can extend to (intramorphemic) vowel patterns beyond the syllable level. Many Australian languages more or less avoid the iCu sequence, and prefer IC1 and UCu. 

Background

The study of patterns of (segmental) phonemes within the word falls under phonotactics, morpheme (and word) structure constraints, and the study of (morpho)phonological processes. The domain of phonotactics has been taken to be the syllable, supplemented in some studies by sensitivity to word boundary (and sometimes morpheme boundary and even reduplication boundary). Notably, the surveys in WALS (Maddison 2013), and the World Phonotactics Database (Donohue et al. 2013) are entirely about the structure of the syllable. Any dependency across a word-internal morpheme boundary typically involves alternations and so is accounted for by (morpho)phonological rules (or active constraints), or effects of metrical structure. The standard view (as outlined) tacitly assumes the absence of phonotactic patterns spanning more than one syllable within a morpheme (but not necessarily spanning an entire polysyllabic morpheme). Even so, some such patterns have been long recognised, under the rubrics of vowel (or other) harmony, assimilation (or dissimilation), and OCP effects. Some early statistical studies of connected speech (and writing) were alert to trans-syllabic phoneme patterns insofar as they showed a relationship between word boundaries and the rise and fall of Markovian transition probabilities.

Previous studies

Some correlation between vowels in adjacent syllables had been noticed in particular Australian languages Gooniyandi (McGregor 1990:88–89) and Ngawun (Breen 1981:28). For (non-PIY) Bardi, Bowern (2012:98–100) showed interaction between V1 and V2: “there is a strong tendency for the V2 vowel to be identical to the V1 vowel” and also found more IC1 sequences than expected. Hercus (1994:54) noted the near absence of iCu sequences in Arubana. Macklin-Cordes 2015 studied transition frequencies between adjacent segments for Ngumpin–Yapa languages (and thus not V/V interaction); later extended to Yolpu Matha. In Warlpiri and Warlmanpa (Yapa languages), the sequence iCu occurs only when the consonant is p or w (Nash 1986:73-4). It might be expected that p would also participate in this but there are no examples of an iCu or pCu sequence. In Warlpiri, surface iCu is no more possible across a morpheme boundary than intra-morphemically, as vowel assimilation rules operate to change underlying iCu to IC1 or uCu. In Warlpiri there are no such vowel assimilation rules and iCu does occur across a morpheme boundary (as also in Arubana-Wangkangurru).

Further topics

• effect of intervening consonants on VCV patterns: e.g. only i[xlabial]u is allowed (as in Pamkala (Thara-Yura), Yapa languages)
• diachrony: loanword detection; inference on direction of change; role in Arandic developments
• lexical frequency interaction; intra-word interaction; statistics on corpora rather than lexicons
• phonetic motivation for interaction?: articulatory, auditory, acquisition; possibly related to word discrimination (‘in short, the evidence that infants rely on transitional probabilities between syllables to segment words from speech is quite convincing.’ Johnson 2012:59)
• assessing areal and phylogenetic components of the distribution of phonotactic patterns; see Macklin-Cordes

Scope

Pama-Nyungan languages (and many other Australian languages) conform to the simple assumptions here: • suffixed • initial stress • stems of at least 2 morae • typically 3 vowels /a/, o, u/ in stems (some languages also with /e/, /o/)

Other assumptions: vowel length ignored; study confined to first two syllables (trochees); compounds with monosyllabic first element counted only if written as one word

This study

Our study began with the typical 3 vowels /a/, o, u/ of stems in Pama-Nyungan languages resembling the Arandic subgroup in central Australia, and revealed a fairly widespread avoidance (more or less) of the high back vowel /u/ following the high front vowel /a/, across intervening (non-labial) consonants C (and independent of metrical structure). In other languages of the region the iCu sequences occur but at a lower frequency than would be expected just from the overall vowel frequencies; also, there is some complementary preference for different V in adjacent syllables (especially l and u). As a graded property the iCu dispreference shows a contiguous areal distribution; there is a similar weaker dispreference in the Western Desert dialect web, and rather different dispreference in other languages further afield, some more strongly avoiding oCu.

The indicator in each cell is

(E-O)* sign (E-O)

where

E is Expected
O is Observed
N is the total number of words

Key

<table>
<thead>
<tr>
<th>Language name</th>
<th>Lexicon size</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>36</td>
</tr>
<tr>
<td>V2</td>
<td>36</td>
</tr>
</tbody>
</table>

Pattern persists across lexicon size sampled, e.g.:

<table>
<thead>
<tr>
<th>Arandic</th>
<th>Pama-Nyungan</th>
</tr>
</thead>
<tbody>
<tr>
<td>iCu</td>
<td>IC1</td>
</tr>
</tbody>
</table>

 Acknowledgements

To LA Hercus, H Koch, J Simpson for discussion (1990), To the compilers of dictionaries used (via OzBib, AILC). To Claire Bowern for Australian Language Locations Poster presentation at the 12th Association for Linguistic Typology conference (ALT2017), 13 December 2017