The position of phonetics in grammars

Binary and Scalar Phonology

PHONETIC TRANSCRIPTION

Universal Phonetics

PHYSICAL REPRESENTATION

The SPE Model.

• Reading for next week: Johnson chapter 2.
• Assignment: Mandarin aspiration, due 9/29.
Phonetic and phonological representations

- The study of linguistic sound patterns is traditionally divided into two sub-fields: phonetics and phonology.
- Phonology specifies the sounds that a language uses, the distribution of those sounds, and alternations in the realization of morphemes (among other things).
- What is left for phonetics?
- If the ‘sounds’ whose distribution is specified in the phonology are characterized in sufficient physical detail, then phonology should describe all aspects of the sound structure of a language.
- But phonology traditionally operates in terms of rather coarse-grained descriptions of sounds, so a lot of detail is left out.
Phonetic and phonological representations

- Example – phonological representations in Chomsky and Halle (1968):
  - strings of segments, essentially as in IPA-style transcription.
  - each segment is specified as a matrix of binary feature specifications.
  - Features are defined phonetically, but in rather broad terms.
Phonetic and phonological representations

E.g. Halle and Clements (1983):

<table>
<thead>
<tr>
<th>Feature</th>
<th>Definition of the + value</th>
</tr>
</thead>
<tbody>
<tr>
<td>[syllabic]</td>
<td>'Constitute syllable peaks'</td>
</tr>
<tr>
<td>[consonantal]</td>
<td>'Sustained vocal tract constriction at least equal to that required in the production of fricatives'</td>
</tr>
<tr>
<td>[sonorant]</td>
<td>'Air pressure inside and outside the mouth is approximately equal'</td>
</tr>
<tr>
<td>[coronal]</td>
<td>'Raising the tongue blade towards the teeth or the hard palate'</td>
</tr>
<tr>
<td>[anterior]</td>
<td>'Primary constriction at or in front of the alveolar ridge'</td>
</tr>
<tr>
<td>[labial]</td>
<td>'With a constriction at the lips'</td>
</tr>
<tr>
<td>[distributed]</td>
<td>'With a constriction that extends for a considerable distance along the midsaggital axis of the oral tract'</td>
</tr>
<tr>
<td>[high]</td>
<td>'Raising the body of the tongue toward the palate'</td>
</tr>
<tr>
<td>[back]</td>
<td>'With the tongue body relatively retracted'</td>
</tr>
<tr>
<td>[low]</td>
<td>'Drawing the body of the tongue down away from the roof of the mouth'</td>
</tr>
<tr>
<td>[round]</td>
<td>'With protrusion of the lips'</td>
</tr>
<tr>
<td>[continuant]</td>
<td>'Allowing the air stream to flow through the midsaggital region of the oral tract'</td>
</tr>
</tbody>
</table>

Image by MIT OCW.

Phonetic and phonological representations

<table>
<thead>
<tr>
<th>[lateral]</th>
<th>'With the tongue placed in such a way as to prevent the air stream from flowing outward through the center of the mouth, while allowing it to pass over one or both sides of the tongue'</th>
</tr>
</thead>
<tbody>
<tr>
<td>[nasal]</td>
<td>'Lowering the velum and allowing air to pass outward through the nose'</td>
</tr>
<tr>
<td>[advanced tongue root]</td>
<td>'Drawing the root of the tongue forward'</td>
</tr>
<tr>
<td>[tense]</td>
<td>'With a tongue body or root configuration involving a greater degree of constriction than that found in their lax counterparts'</td>
</tr>
<tr>
<td>[strident]</td>
<td>'With a complex constriction forcing the air stream to strike two surfaces (sic), producing high-intensity fricative noise'</td>
</tr>
<tr>
<td>[spread glottis]</td>
<td>'With the vocal folds drawn apart, producing a non-periodic (noise) component in the acoustic signal'</td>
</tr>
<tr>
<td>[constricted glottis]</td>
<td>'With the vocal cords drawn together, preventing normal vocal cord vibration'</td>
</tr>
<tr>
<td>[voiced]</td>
<td>'With a laryngeal configuration permitting periodic vibration of the vocal cords'</td>
</tr>
</tbody>
</table>

Image by MIT OCW.

Phonetic and phonological representations

• So standard phonological representations can characterize speech to about the same level of detail as a broad phonetic transcription. The remaining detail is generally held to be the subject matter of phonetics.

• Chomsky and Halle proposed an intervening step: phonetic detail rules convert binary feature specifications into scalar values.
  – However, hardly anybody has actually adopted this proposal.

• The remaining detail is supposed to be a matter of universal phonetics, and therefore not really part of grammar.
Phonetics and phonology

• The question that Keating (1985) addresses is how much phonetic detail can be supplied by ‘universal phonetics’.

• The short answer: not much.

• Implication: Most aspects of phonetic realization must be specified in grammar, either in phonology or in a phonetic component.
Widespread tendencies are subject to language-specific variation

Case study: Voicing effects on vowel duration.

- Vowels are shorter before voiceless obstruents than before voiced obstruents or sonorants in many languages (Chen 1970)
  - E.g. English [ɛ] is shorter in ‘bet’ than in ‘bed’ and ‘ben’ (ratio is approx. 0.8).

- Language-specific variation:
  - Effect is greater in English
  - No effect in Polish, Czech, Saudi Arabic
  - Some evidence that the effect is conditioned by underlying voicing in Russian, German, English

- There are many more examples of language-specific realization of similar phonological categories (below).
Mechanical physiological effects

• What could give rise to universal phonetic effects?
• Keating: mechanical physiological effects.
  – If a pattern of realization is a consequence of basic physiology then it should be observed in all languages.
Keating: Mechanical physiological effects

Case study: Intrinsic vowel duration

- Across languages, lower vowels are longer, other things being equal.
  - If articulator velocities are constrained, lower vowels will take longer to produce.
- Hypothesized physiological basis: If only the magnitude, but not duration, of force input to jaw varies, low vowels will be longer (Lindblom 1967).
Keating: Mechanical physiological effects

Test: Electromyographic (emg) study of muscle activity in jaw lowering.

• Lower jaw position is correlated with longer duration in English
  – The difference in duration was due to difference in movement duration, not duration of steady state.

• But low vowels show longer and higher amplitude of muscle activity

• i.e. variation in duration is under the control of the speaker.

• The correlation between vowel height and duration could still be related to differences in movement distance, but the linkage does not have the hypothesized mechanical basis.
  – Could involve a dispreference for the effort involved in fast movements.
Phonetic and phonological representations

E.g. Halle and Clements (1983):

<table>
<thead>
<tr>
<th>Feature</th>
<th>Definition of the + value</th>
</tr>
</thead>
<tbody>
<tr>
<td>[syllabic]</td>
<td>'Constitute syllable peaks'</td>
</tr>
<tr>
<td>[consonantal]</td>
<td>'Sustained vocal tract constriction at least equal to that required in the production of fricatives'</td>
</tr>
<tr>
<td>[sonorant]</td>
<td>'Air pressure inside and outside the mouth is approximately equal'</td>
</tr>
<tr>
<td>[coronal]</td>
<td>'Raising the tongue blade towards the teeth or the hard palate'</td>
</tr>
<tr>
<td>[anterior]</td>
<td>'Primary constriction at or in front of the alveolar ridge'</td>
</tr>
<tr>
<td>[labial]</td>
<td>'With a constriction at the lips'</td>
</tr>
<tr>
<td>[distributed]</td>
<td>'With a constriction that extends for a considerable distance along the midsaggital axis of the oral tract'</td>
</tr>
<tr>
<td>[high]</td>
<td>'Raising the body of the tongue toward the palate'</td>
</tr>
<tr>
<td>[back]</td>
<td>'With the tongue body relatively retracted'</td>
</tr>
<tr>
<td>[low]</td>
<td>'Drawing the body of the tongue down away from the roof of the mouth'</td>
</tr>
<tr>
<td>[round]</td>
<td>'With protrusion of the lips'</td>
</tr>
<tr>
<td>[continuant]</td>
<td>'Allowing the air stream to flow through the midsaggital region of the oral tract'</td>
</tr>
</tbody>
</table>

Image by MIT OCW.

The nature of phonetic universals

• The failure of phonetic universals to place hard constraints on cross-linguistic variation is unsurprising. E.g:

• Phonetic universal: There are limits on the rate of \( f0 \) change.
  – Xu & Sun (2002) asked subjects to imitate a fast, alternating high-low pitch pattern, with various pitch ranges between H & L.
  – Mean time (ms) to complete a pitch change of \( d \) semitones:
    • Rising: \( t = 89.6 + 8.7d \)
    • Falling: \( t = 100.4 + 5.8d \)
  – E.g. a fall from 200 Hz to 100 Hz (12 s.t.) takes at least \(~170\) ms.
Tone coarticulation in Cantonese

• So it’s no surprise that there are transitions between tones of different levels.
Tonal coarticulation

• But there is no obvious physiological constraint that determines how f0 transitions should be timed with respect to the segmental string.

  during syll1?  during syll2?  across the boundary?

• Different patterns are observed in different languages.
Tone coarticulation in Cantonese

- In Cantonese transition towards a tone does not begin until the onset of the syllable containing that tone (Li et al 2004).
  - Also in Mandarin (Xu 1997), Thai (Gandour et al 1994), Vietnamese (Brunelle 2003).
Tonal coarticulation in Kinyarwanda

Kinyarwanda (Myers 2003), L, H tones:

• In an L.H sequence, the rising transition begins well before the onset of the H tone syllable, half way through the low syllable, or earlier.

• Falling transition carries over into a following L syllable.

© International Society of Phonetic Sciences, Karger. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.
Tonal coarticulation

• Transitions between tones are universal, but the timing of those transitions is language-specific and must be specified in the grammar.

• Physiological limitations constrain phonetic realization, but they cannot determine it.
Phonetic grammars

• Hypothesis: phonetic grammars operate by balancing conflicting constraints enforcing preferred properties for phonetic realizations.

• Two basic classes of constraints:
  – Minimize Effort - physiological limitations, etc.
  – Faithfulness constraints - require accurate realization of perceptual targets for speech sounds.

• Phonetic realizations are selected so as to minimize violation of these constraints.

• Cross-linguistic variation derives from different weightings of these constraints.
  – Weights may correlate with phonological properties.
Tonal coarticulation

- Tonal targets extend through the syllable (Xu & Liu 2006)
- Faithfulness: Do not deviate from tonal targets.
- Effort: limits on rate of change of $F_0$
- These constraints conflict - transitions result in deviation from the tone targets.
- Specific patterns of timing result from more specific faithfulness constraints that penalize deviations from particular kinds or parts of targets.
Tonal coarticulation

The Cantonese pattern:

- faithful realization of tonal targets is more important in the rime than in the onset (Flemming 2011)
  - The rime is generally the region of highest intensity periodicity, and therefore the place where tone is most perceptible (cf. Zhang 2004).

- Realizing transitions at the beginnings of syllables minimizes violation of faithfulness to rime targets:

  ✓ least deviation in rime
  ✗ greatest deviation in rime
  ✗ deviation in rime
Phonetic realization as an optimization problem

• This analysis can be given a precise formulation as an optimization problem.
• Effort: assume a maximum rate of $F_0$ transition.
• So transition between level tones takes the form shown below.
Faithfulness constraints

• Ident-T: The $F_0$ contour must match the tone target, $T$.
  – Cost of violation is proportional to the squared difference between the target and the actual $F_0$ contour, integrated over the duration of the target.
  – Multiplied by a language-specific weight.

• Ident-T-rime: The $F_0$ contour must match the tone target during the rime.
  – Cost calculated in the same way.
  – Also weighted.

• Select the timing of the $F_0$ transition that minimizes the summed violations of the faithfulness constraints, subject to the effort constraint on maximum rate of $F_0$ transition.
Example: Realization of H.L

- Ident-T has a weight of 1
- Ident-T-rime has a weight of 0

- Ident-T has a weight of 0.01
- Ident-T-rime has a weight of 0.99.
Tonal coarticulation

The Kinyarwanda pattern:

• faithfulness to H tone targets is more important than faithfulness to L tone targets.
  – Only one H tone per morpheme, so H tones are much more informative than L tones in distinguishing words from each other

• Realizing transitions during L tone syllables minimizes violation of faithfulness to H tone targets.

![Graphs showing least and greatest deviation from H target]
The nature of phonetic universals

- Phonetic universals are constraints, not patterns of phonetic realization.
- Patterns of phonetic realization derive from the interaction of multiple constraints.
- They are language-specific because the prioritization of constraints differs across languages.
The nature of phonetic universals

- Phonetic universal: full retroflexion is not compatible with a high front tongue position [i]. (The tongue tip/blade and tongue body cannot simultaneously form constrictions with the hard palate).

Images by MIT OCW.

The nature of phonetic universals

• This constraint has a variety of consequences in front vowel/retroflex sequences:
  – Kodagu (Emeneau 1970) - vowels are retracted preceding retroflexes.
  – Mantjiltjara (Marsh 1969) - retroflexion is ‘very weak’ after [i].
  – Gujarati - reduced retroflexion following [i] observable in palatograms in Dave (1977).

• Effort constraint: minimize peak velocity of articulator movements (tongue tip, body).

• Faithfulness to C targets, Faithfulness to V targets.

• Different patterns result from different constraint weights.
Evidence for language-specific phonetic detail
- Cross-linguistic variation in the realization of phonological categories

• Aspirated and unaspirated voiceless stops
  – [p vs. pʰ]
• Background: Voice Onset Time
Voice Onset Time

- English utterance-initial stops

Voiceless unaspirated

Voiceless aspirated

22 ms

86 ms

die
tie
VOT, closure voicing

- English intervocalic stops can be fully voiced
  - VOT is 0 ms in 2nd and 3rd stops

brigadoo(n)
VOT, closure voicing

- Hindi - three-way contrast
  - recordings from Ladefoged [http://www.phonetics.ucla.edu/vowels/chapter12/hindi.html](http://www.phonetics.ucla.edu/vowels/chapter12/hindi.html)

bal
‘hair’

pal
‘take care of’

pʰal
‘knife blade’
Cross-linguistic variation in voiceless stops

- Standard phonological representations use a single feature, [+/-spread glottis] to distinguish aspirated and unaspirated stops.
- If phonetics is universal, these categories should be realized similarly in all languages.
- In fact VOT of voiceless unaspirated and aspirated stops varies significantly between languages.
Cross-linguistic variation in voiceless stops


<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>CONTRASTING VELAR STOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aleut (Eastern)</td>
<td>k</td>
</tr>
<tr>
<td>Aleut (Western)</td>
<td>k, kʰ, k’</td>
</tr>
<tr>
<td>Apache</td>
<td>k</td>
</tr>
<tr>
<td>Banawá</td>
<td>k</td>
</tr>
<tr>
<td>Bowiri</td>
<td>k, g</td>
</tr>
<tr>
<td>Chickasaw</td>
<td>k</td>
</tr>
<tr>
<td>Dahalo</td>
<td>k, g</td>
</tr>
<tr>
<td>Defaka</td>
<td>k, g</td>
</tr>
<tr>
<td>Gaelic</td>
<td>k, kʰ</td>
</tr>
<tr>
<td>Hupa</td>
<td>k, kʰ, k’</td>
</tr>
<tr>
<td>Jalapa Mazatec</td>
<td>k, k’</td>
</tr>
<tr>
<td>Khonoma Angami</td>
<td>k, g</td>
</tr>
<tr>
<td>Montana Salish</td>
<td>k, k’</td>
</tr>
<tr>
<td>Navajo</td>
<td>k, kʰ, k’</td>
</tr>
<tr>
<td>Tlingit</td>
<td>k, kʰ, k’</td>
</tr>
<tr>
<td>Tsouì</td>
<td>k</td>
</tr>
<tr>
<td>Wari’</td>
<td>k</td>
</tr>
<tr>
<td>Yapese</td>
<td>k, k’, g</td>
</tr>
</tbody>
</table>

Figure removed due to copyright restrictions.
Cross-linguistic variation in voiceless stops

• Conflicting constraints?
  – Distinctiveness of voicing/aspiration contrasts (favors longer VOT for aspirated stops, shorter VOT for contrastively unaspirated stops).
  – Preference for fully voiced vowels (faithfulness to vowel targets).
  – And/or faithfulness to place of articulation targets: aspiration makes formant transitions less clear.
VOT duration/voiced vowel duration trade-off

• Port and Rotunno (1979) found that in English VOT increases with duration of the following vowel,
  – but VOT is not a fixed proportion of the vowel.

• This pattern could represent a trade-off between a preference for distinct voicing contrasts (long VOT) and faithfulness to vowel targets (preference for fully voiced vowels).

Cross-linguistic variation in the realization of phonological categories

- ‘VOT’ in ejectives (a.k.a. glottal lag)
- Examples: Montana Salish

\[
\text{p’úm} \quad \text{t’áq’ën}
\]
VOT in ejectives

- Navaho [k’aːʔ] vs. Hausa [k’aːrəː]

- Navajo 94ms vs. Hausa 33ms

![Waveforms illustrating differences between Navajo and Hausa ejectives. The arrows indicate the releases of the glottal closures.](Image by MIT OCW.)
VOT in ejectives

• Cho and Ladefoged (1999)

<table>
<thead>
<tr>
<th>Language</th>
<th>Bilabial</th>
<th>Alveolar</th>
<th>Velar</th>
<th>Uvular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apache</td>
<td></td>
<td>46</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Hupa</td>
<td></td>
<td>93</td>
<td>80</td>
<td>89</td>
</tr>
<tr>
<td>Montana Salish</td>
<td>81</td>
<td>65</td>
<td>86</td>
<td>81</td>
</tr>
<tr>
<td>Navajo</td>
<td></td>
<td>108</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Tlingit</td>
<td></td>
<td>95</td>
<td>84</td>
<td>117</td>
</tr>
<tr>
<td>Yapese</td>
<td>60</td>
<td>64</td>
<td>78</td>
<td></td>
</tr>
</tbody>
</table>


• Standard phonological features treat all ejectives as [+constricted glottis] stops.
  – but there are significant language-specific differences in phonetic realization within this category.
Degrees of retroflexion

Two (or more) degrees of retroflexion

- Apical post-alveolar, e.g. Hindi, vs. Sublaminal post-alveolar, e.g. Telugu (Ladefoged and Bhaskararao 1983)

- Phonologically: both [+coronal, -anterior, -distributed]?

X-ray tracings of the apical retroflex ç in Hindi and the sub-apical retroflex ç in Tamil and Telugu.

Image by MIT OCW.

Vowel quality

- Similar front vowels of Danish (dotted) and English (solid) (Disner 1978, 1983).

- Danish vowels are systematically higher than their English counterparts.

Image by MIT OCW. Adapted from Disner (1983).
Cross-linguistic variation in contextual phonetic effects

• Coarticulation
  – e.g. Nasalization adjacent to nasals (Cohn 1990, 1993).

Adapted from Cohn, A. Nasalization in English: Phonology or phonetics? Phonology 10 (1993): 43-81.
Language-specific variation in stop-vowel coarticulation

- F2 at the release of a stop, $F2(C)$, varies as a linear function of F2 at the middle of the following vowel, $F2(C)$ (1st lecture)

\[ F2(V) \]

Figure removed due to copyright restrictions.
Source: Figure 1, Fowler, Carol A. "Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation." Attention, Perception, & Psychophysics 55, no. 6 (1994): 597-610.
Language-specific variation in stop-vowel coarticulation

- $F2(C)$ after stops is a linear function of $F2(V)$ in CV sequences in all languages that have been studied, but the slope and intercept of that function differ from language to language for similar sounds.

- Thai [d] \[ F2(C) = 0.3F2(V) + 1425 \] (0.24-0.33)

- Urdu [d] \[ F2(C) = 0.5F2(V) + 857 \] (0.43-0.57)

- Sussman et al (1993). Averages over 6 and 5 speakers respectively. Both stops are reported to be dental.

- So the specific patterns must be specified in the grammars of these languages.
Closed syllable shortening

- Navajo: long vowel duration is similar in closed and open syllables (Zhang 2001)

- Thai: long vowels are substantially shorter in closed syllables (Morén & Zsiga 2006)


© Springer. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/. Source: Morén, Bruce, and Elizabeth Zsiga. "The lexical and post-lexical phonology of Thai tones." Natural Language & Linguistic Theory 24, no. 1 (2006): 113-178.
Summary

• There is language-specific variation in matters of relatively fine phonetic detail.
• Standard phonological representations cannot encode all of this detail.
• Therefore - either:
  – phonological representations need to be enriched, or
  – we should posit a language-specific phonetic component of grammar,
• Either way, phonetics is part of grammar.
References


References


