24.963

Linguistic Phonetics

The position of phonetics in grammars

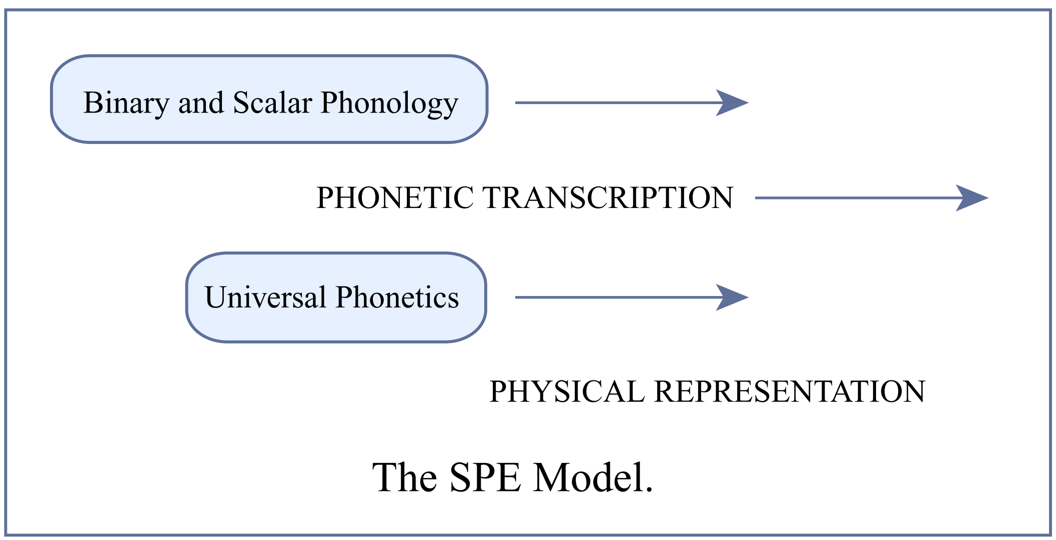


Image by MIT OCW.

Adapted from Keating, P. "Universal phonetics and the organization of grammars." In *PhoneticLinguistics*. Edited V. Fromkin. Indianapolis, IN: Academic Press, 1985, pp. 115-32.

- 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):

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

Image by MIT OCW.

Adapted from Halle, M., and N. Clements. *Problem Book in Phonology: A Workbook for Courses in Introductory Linguistics and Modern Phonology*. Cambridge, MA: MIT Press, 1983.

Phonetic and phonological representations

[lateral]	'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'
[nasal]	'Lowering the velum and allowing air to pass outward through the nose'
[advanced tongue root]	'Drawing the root of the tongue forward'
[tense]	'With a tongue body or root configuration involving a greater degree of constriction than that found in their lax counterparts'
[strident]	'With a complex constriction forcing the air stream to strike two surfaces (sic), producing high-intensity fricative noise'
[spread glottis]	'With the vocal folds drawn apart, producing a non-periodic (noise) component in the acoustic signal'
[constricted glottis]	'With the vocal cords drawn together, preventing normal vocal cord vibration'
[voiced]	'With a laryngeal configuration permitting periodic vibration of the vocal cords'

Image by MIT OCW.

Adapted from Halle, M., and N. Clements. *Problem Book in Phonology: A Workbook for Courses in Introductory Linguistics and Modern Phonology*. Cambridge, MA: MIT Press, 1983.

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 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.
- Lehiste (1970): low vowels generally involve greater articulatory movement from/to adjacent consonants.
  - 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):

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

Image by MIT OCW.

Adapted from Halle, M., and N. Clements. *Problem Book in Phonology: A Workbook for Courses in Introductory Linguistics and Modern Phonology*. Cambridge, MA: MIT Press, 1983.

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

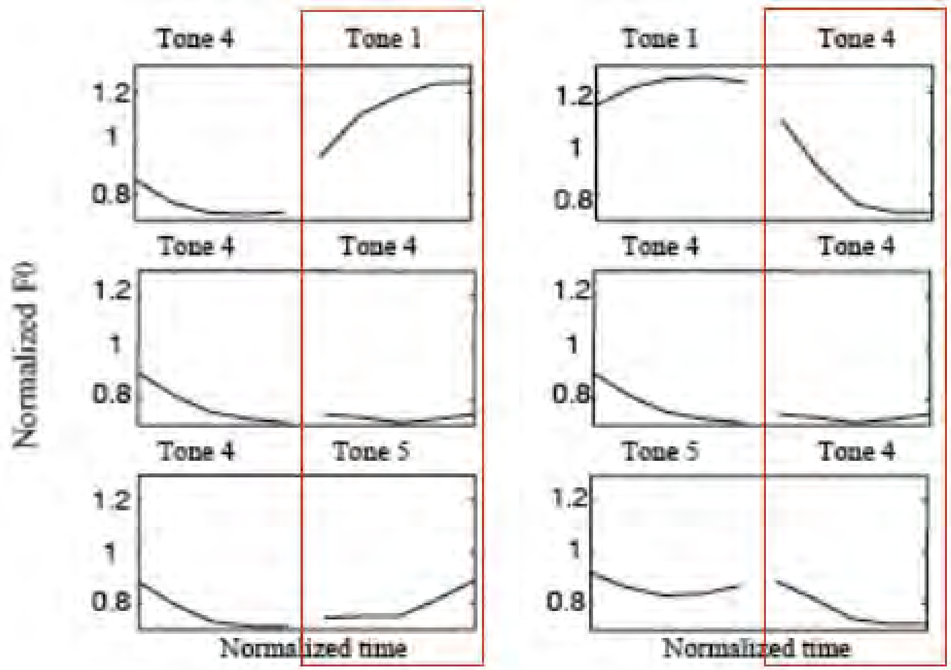


Fig. 13. Examples of co-articulated tone contours of disyllabic words.

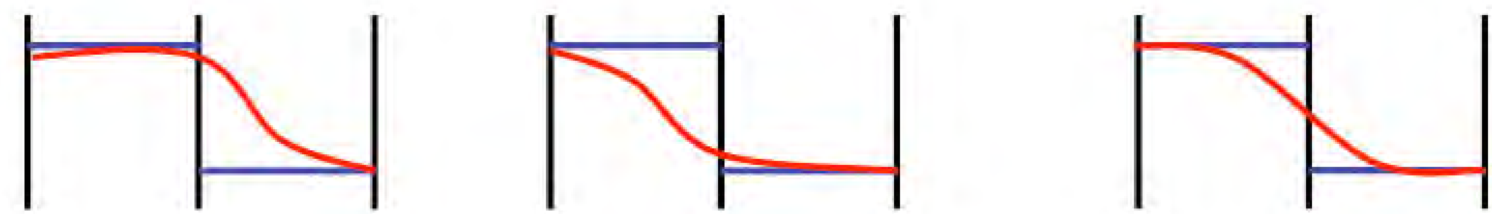
Source: Li, Lee & Qian (2002) "Acoustical F0 analysis of continuous Cantonese speech." International Symposium on Chinese Spoken Language Processing.

- 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

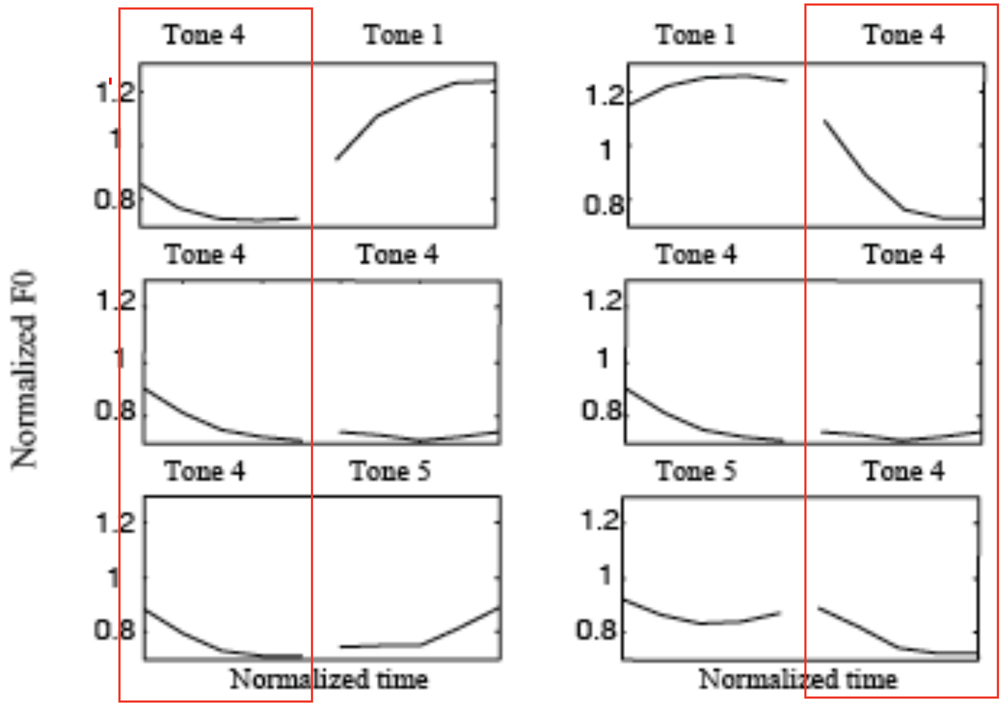


Fig. 13. Examples of co-articulated tone contours of disyllabic words.

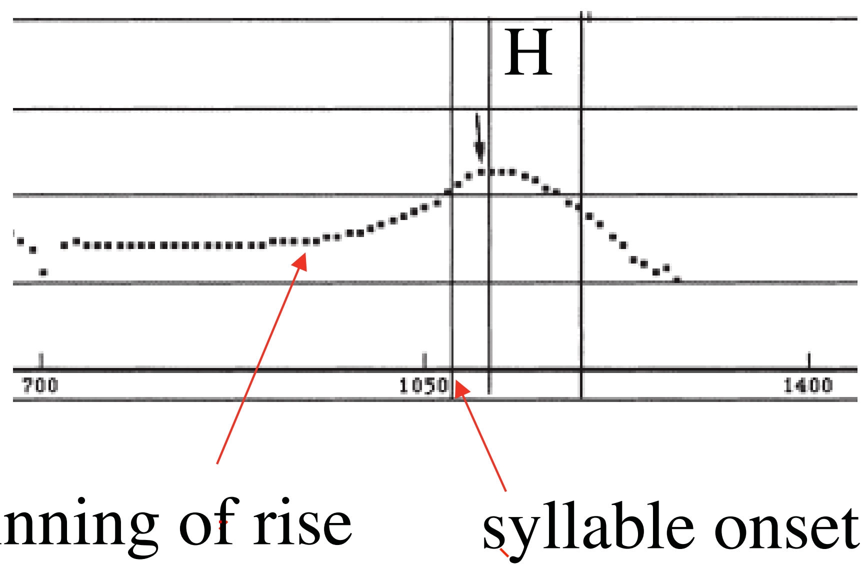
Source: Li, Lee & Qian (2002) "Acoustical F0 analysis of continuous Cantonese speech." International Symposium on Chinese Spoken Language Processing.

- 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.



Source: Myers, Scott. "F0 timing in Kinyarwanda." Phonetica 60, no. 2 (2003): 71-97.

Tonal coarticulation

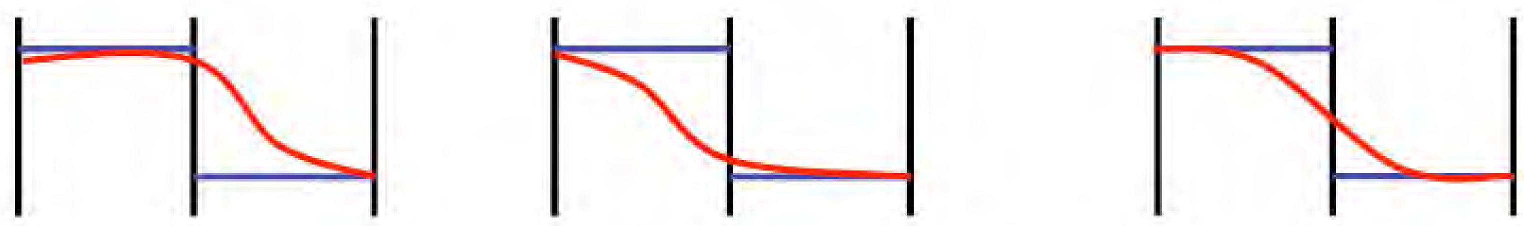
- Transitions between tones are universal, but the timing of those transitions is language-specific and must be specified in the grammar.
- Physiological limitations constrain phonetic realization, but they cannot determine it.

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 to 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

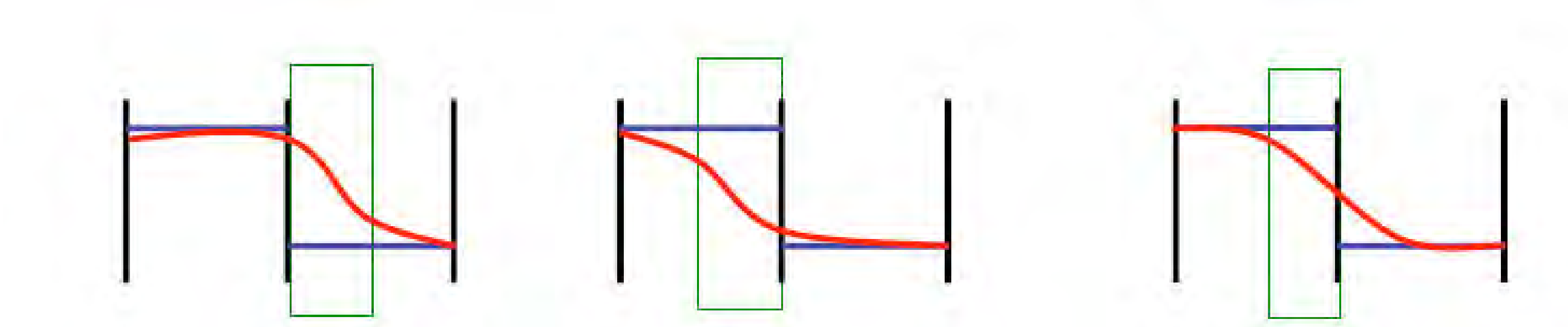
- Tonal targets extend through the syllable (Xu & Liu 2006)
- Faithfulness: Do not deviate from tonal targets.
- Effort: limits on rate of change of F_0
- These constraints conflict - transitions result in deviation from the tone targets.
- Specific patterns of timing result from more specific faithfulness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



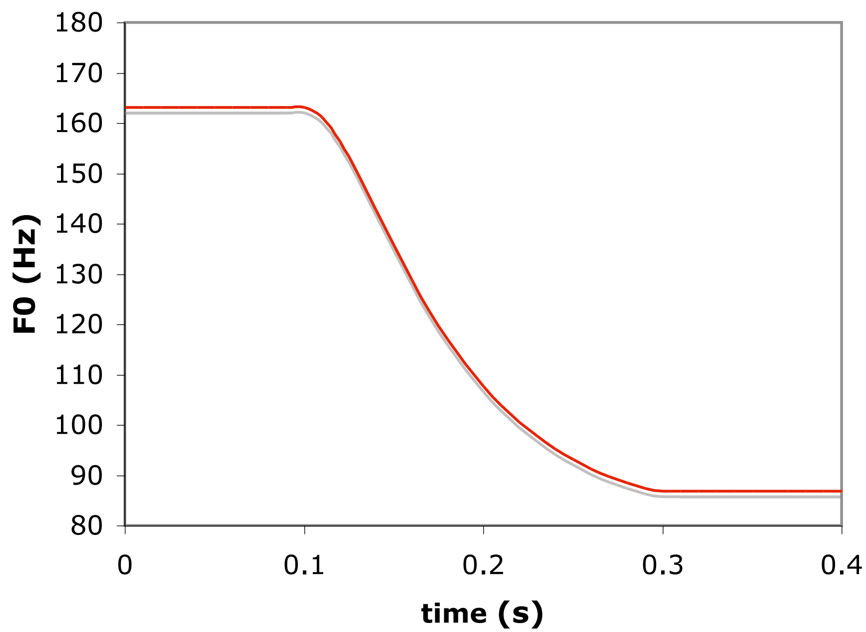
✓least deviation in rime

✗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.

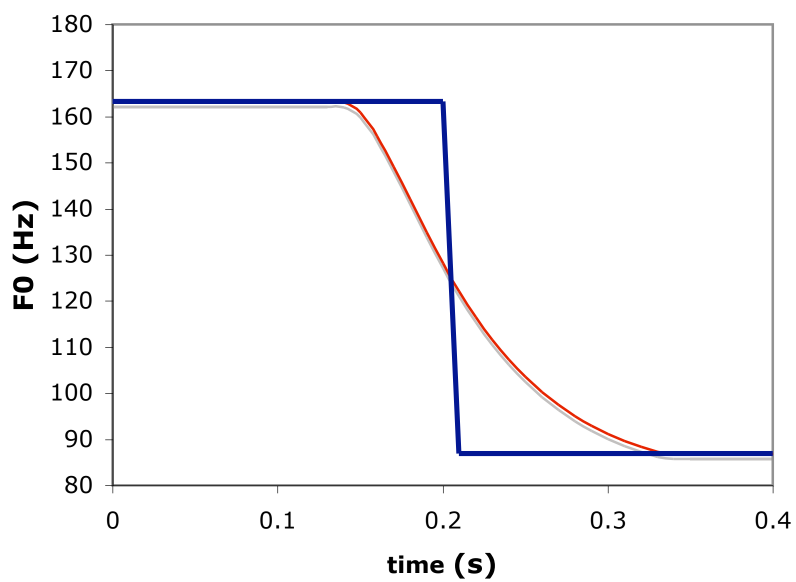


Faithfulness constraints

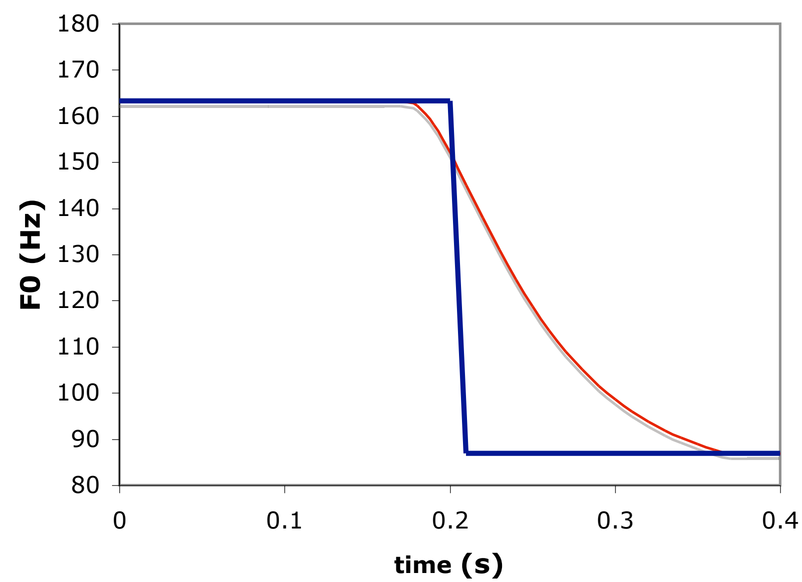
- Ident-T: The F_0 contour must match the tone target, T.
  - Cost of violation is proportional to the squared difference between the target and the actual F_0 contour, integrated over the duration of the target.
  - Multiplied by a language-specific weight.
- Ident-T-rime: The F_0 contour must match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



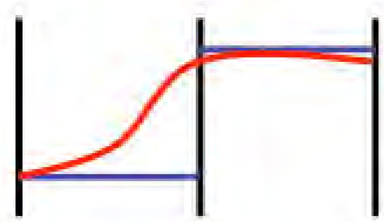
- Ident-T has a weight of 0.01
- Ident-T-rime has a weight of 0.99.



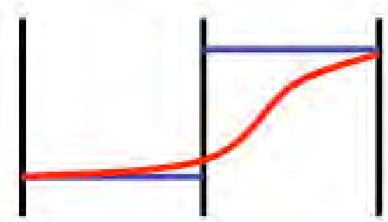
Tonal coarticulation

The Kinyarwanda pattern:

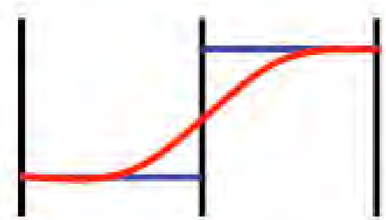
- faithfulness to H tone targets is more important than faithfulness to L tone targets.
  - Only one H tone per morpheme, so H tones are much more informative than L tones in distinguishing words from each other
- Realizing transitions during L tone syllables minimizes violation of faithfulness to H tone targets.



✓least deviation from H target



✗greatest deviation from H target



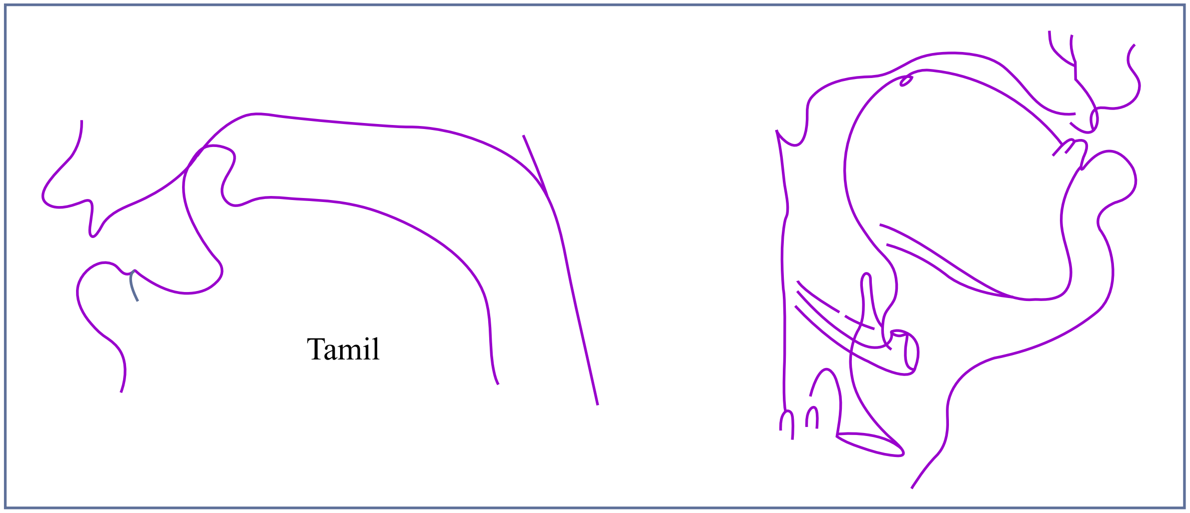
✗deviation from H target.

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.

Left figure adapted from Ladefoged, Peter, and Ian Maddieson. *The Sounds of the World's Languages*. Malden, MA: Blackwell, 1996. Right figure adapted from Stevens, Kenneth N. *Acoustic Phonetics*. Cambridge, MA: MIT Press, 1999.

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.
  - Gugada (Platt 1972) - partial backing and lowering of vowel [iəɖ] (cf. English).
  - 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.

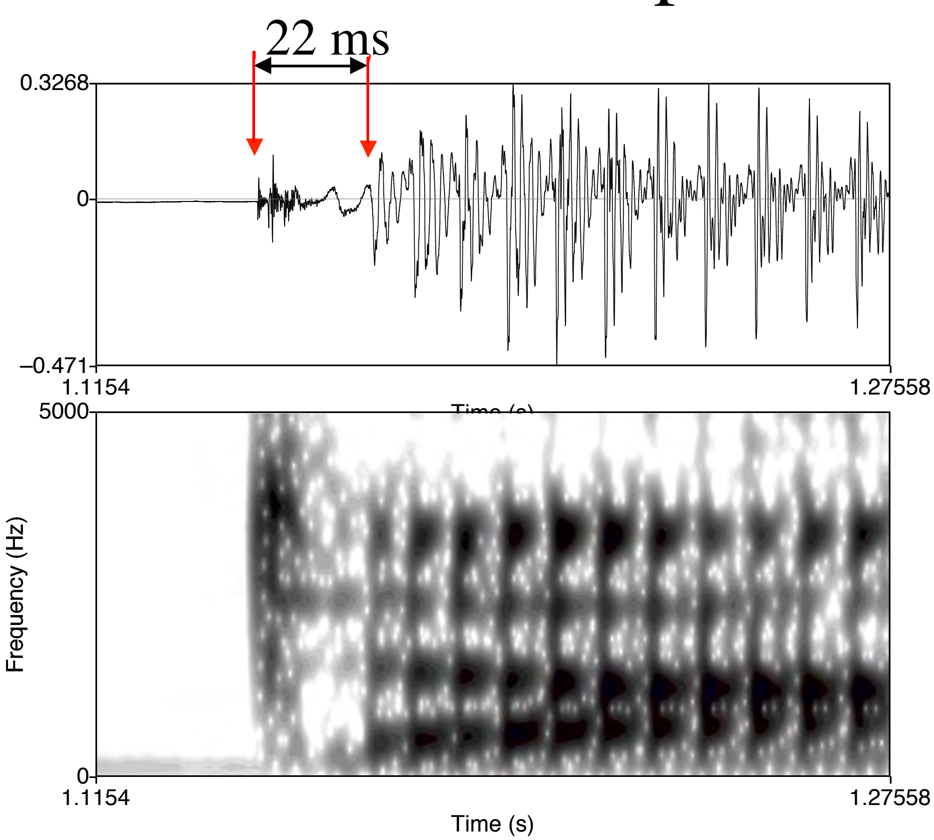
Evidence for language-specific phonetic detail - Cross-linguistic variation in the realization of phonological categories

- Aspirated and unaspirated voiceless stops
  - [p vs. p^h]
- Background: Voice Onset Time

Voice Onset Time

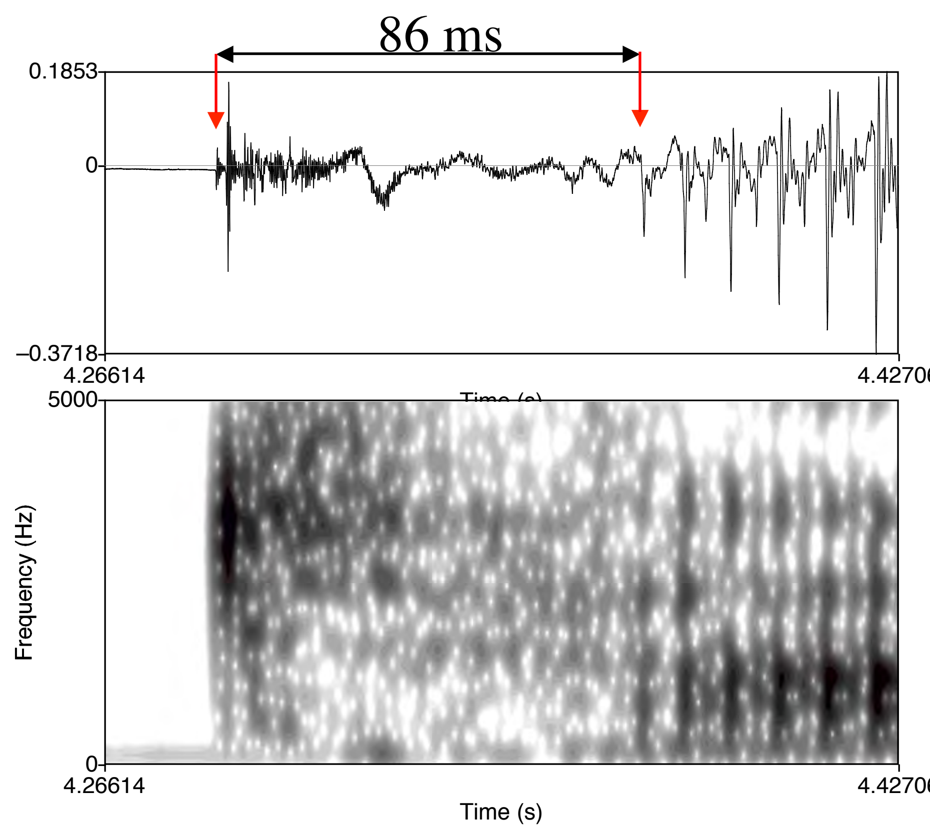
- English utterance-initial stops

Voiceless unaspirated



die

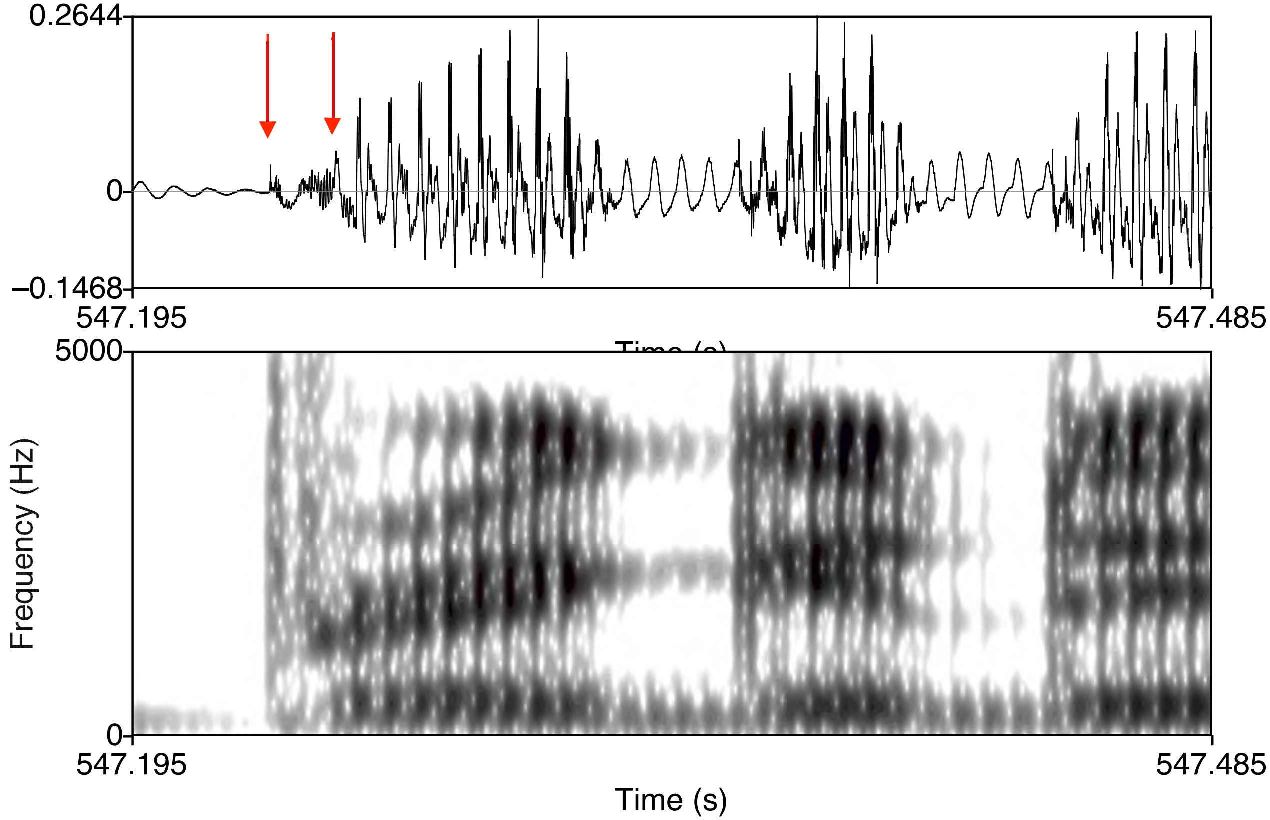
Voiceless aspirated



tie

VOT, closure voicing

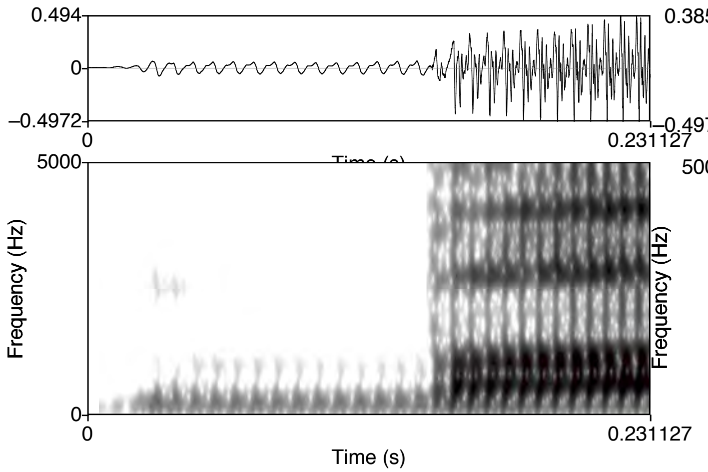
- English intervocalic stops can be fully voiced
  - VOT is 0 ms in 2nd and 3rd stops



brigadoo(n)

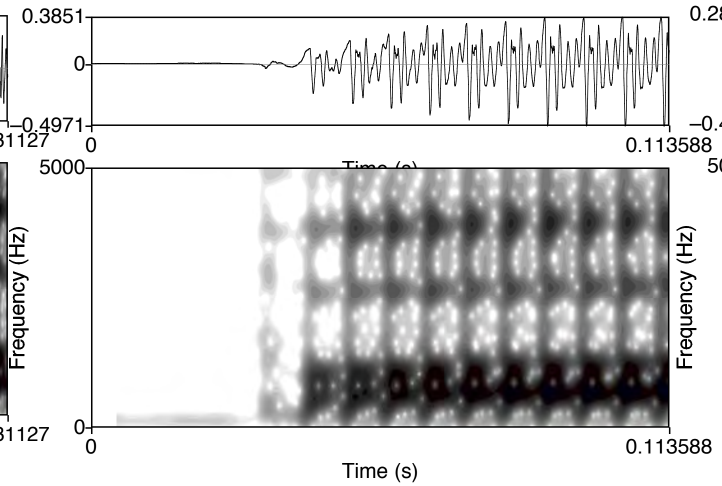
VOT, closure voicing

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



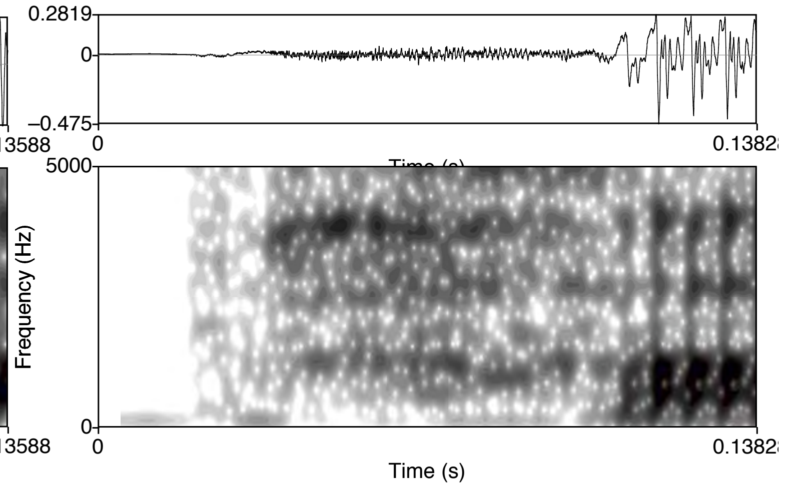
bal

'hair'



pal

'take care of'



p^hal

'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

- Voiceless aspirated and unaspirated velar stops (Cho and Ladefoged 1999, Ladefoged and Cho 2001).

LANGUAGE	CONTRASTING VELAR STOPS
Aleut (Eastern)	**k**
Aleut (Western)	**k**
Apache	**k, k^h, k'**
Banawá	**k**
Bowiri	**k, g**
Chickasaw	**k**
Dahalo	**k, g**
Defaka	**k, g**
Gaelic	**k, k^h**
Hupa	**k, k^h, k'**
Jalapa Mazatec	**k, k^h**
Khonoma Angami	**k, g**
Montana Salish	**k, k'**
Navajo	**k, k^h, k'**
Tlingit	**k, k^h, k'**
Tsou	**k**
Wari'	**k**
Yapese	**k, k', g**

Figure removed due to copyright restrictions.
Source: Ladefoged, Peter, and Taehong Cho. "Linking linguistic contrasts to reality: The case of VOT." UCLA Working Papers in Phonetics (2000): 1-9.

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).

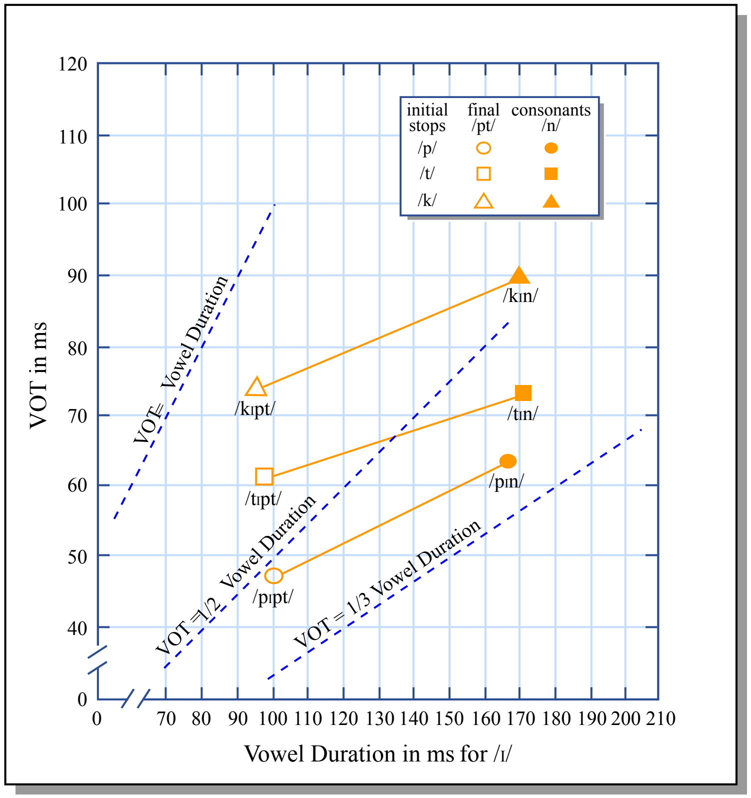
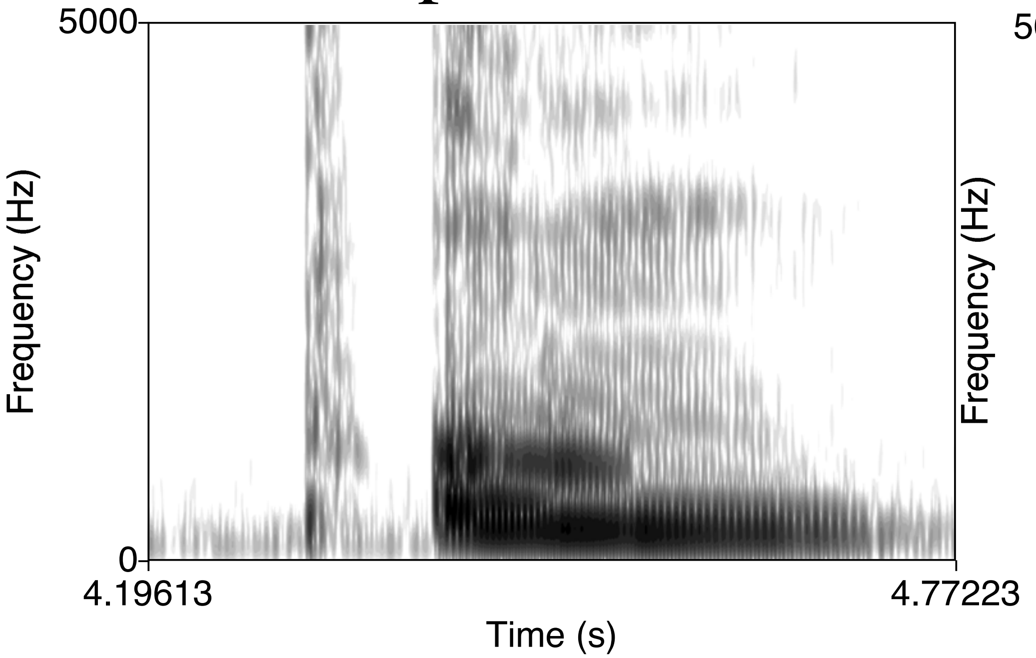


Image by MIT OCW.

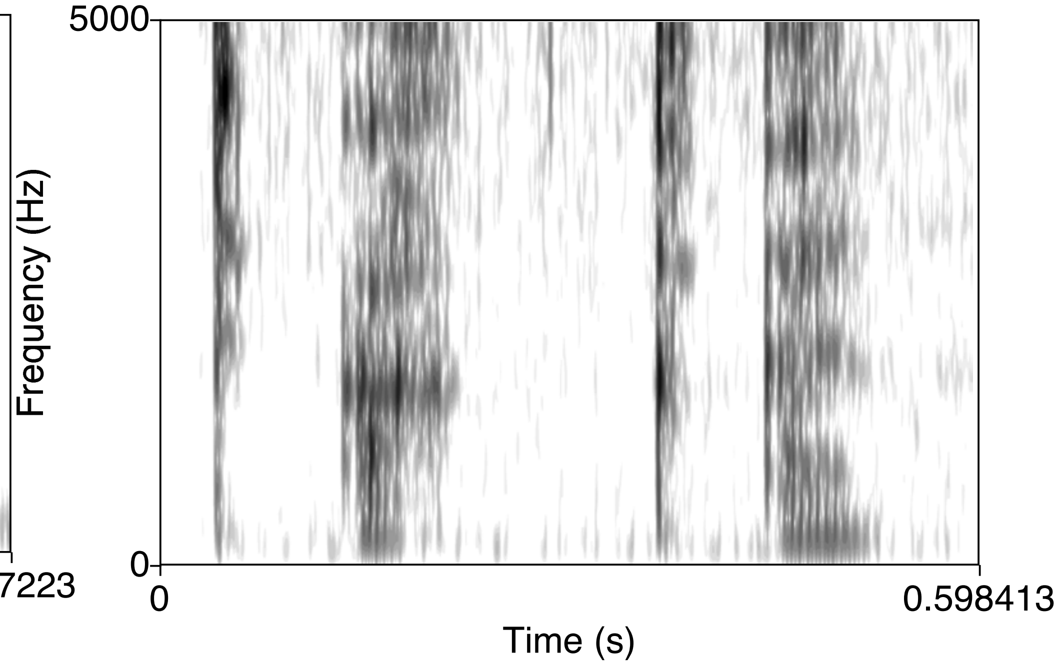
Adapted from Port, Robert F., and Rosemarie Rotunno. "Relation between voice-onset time and vowel duration." *The Journal of the Acoustical Society of America* 66, no. 3 (September 1979): 654-662.

Cross-linguistic variation in the realization of phonological categories

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



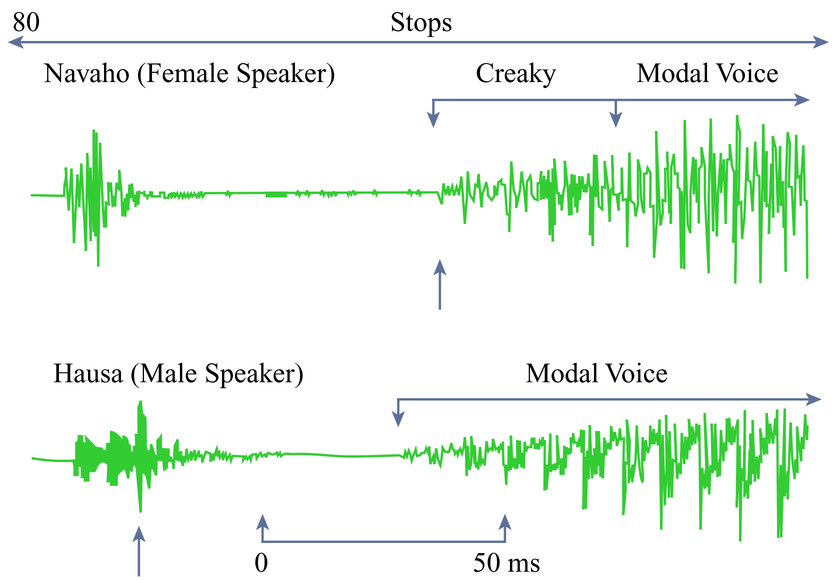
p'úm



t'áq'ən

VOT in ejectives

- Navaho [k’aːʔ] vs. Hausa [k’aːràː]
- 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)

Voice Onset Time (ms) for Ejectives in Six Languages				
Language	Bilabial	Alveolar	Velar	Uvular
Apache		46	60	
Hupa		93	80	89
Montana Salish	81	65	86	81
Navajo		108	94	
Tlingit		95	84	117
Yapese	60	64	78	

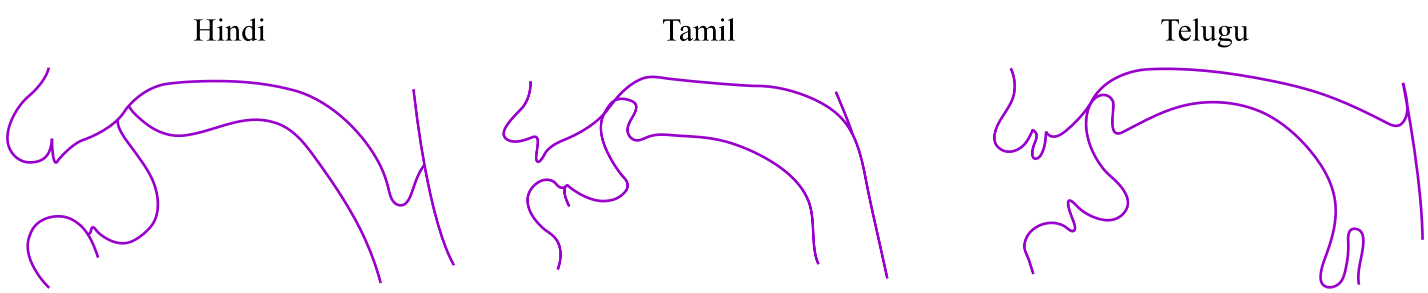
Adapted from.Cho, Taehong, and Peter Ladefoged. "Variations and universals in VOT: Evidence from 18 languages." *Journal of Phonetics* 27 (1999): 207-229

- 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 ḍ in Hindi and the sub-apical retroflex ḍ in Tamil and Telugu.

Image by MIT OCW.

Adapted from Ladefoged, Peter, and Ian Maddieson. *The Sounds of the World's Languages*. Malden, MA: Blackwell, 1996. Based on Ladefoged, Peter, and Peri Bhaskararao. "Non-quantal aspects of consonant production: A study of retroflex consonants." *Journal of Phonetics* 11 (1983): 291–302.

Vowel quality

- Similar front vowels of Danish (dotted) and English (solid) (Disner 1978, 1983).
- Danish vowels are systematically higher than their English counterparts.

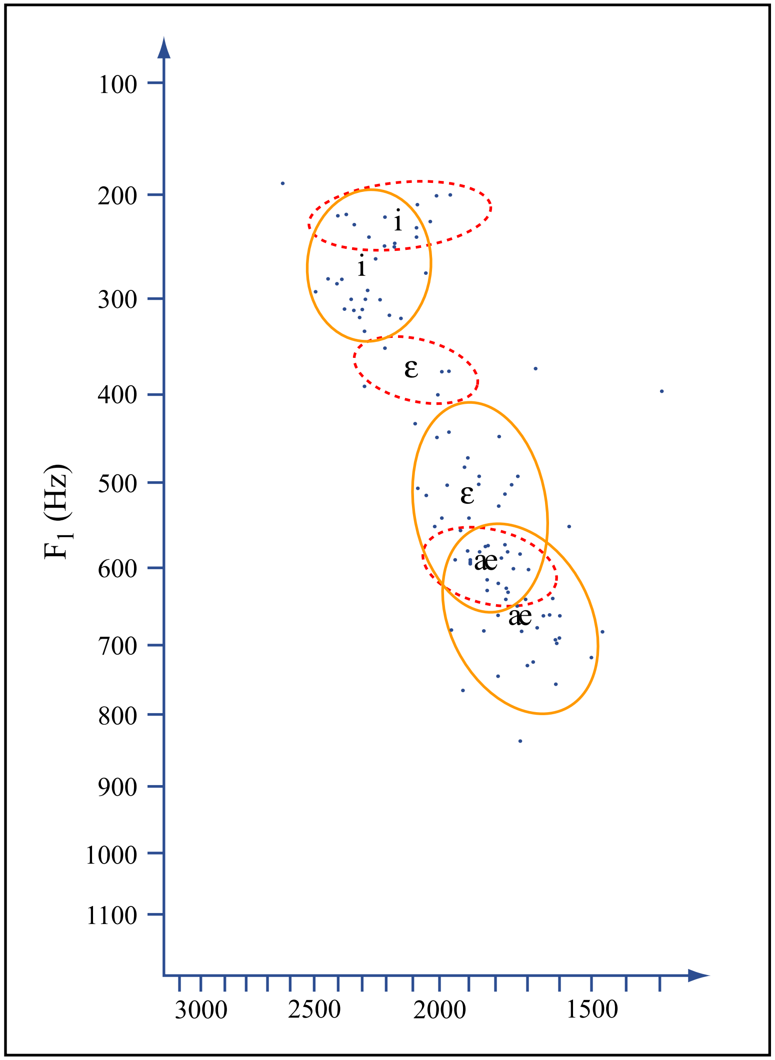


Image by MIT OCW.
Adapted from Disner (1983).

Cross-linguistic variation in contextual phonetic effects

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

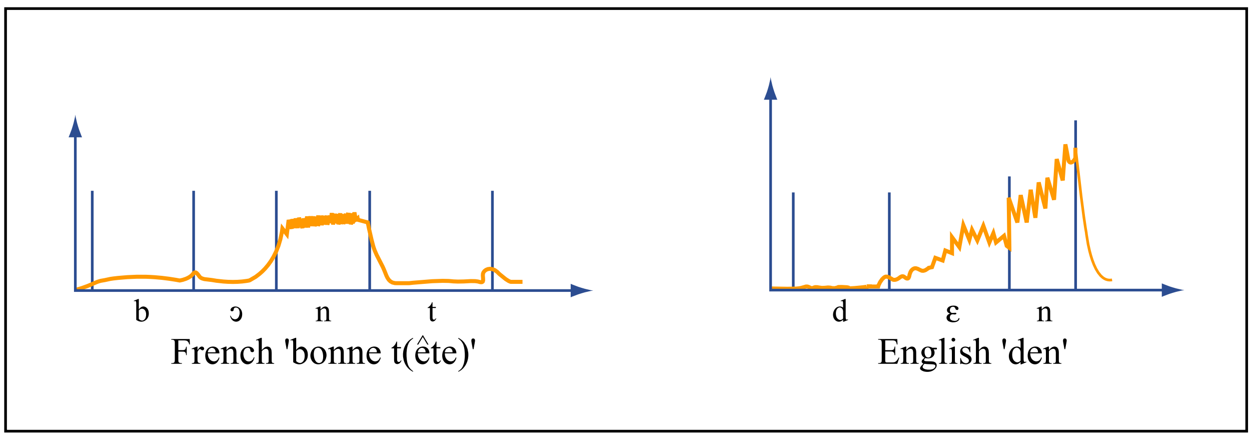


Image by MIT OCW.
Adapted from Cohn, A. Nasalization in English: Phonology or phonetics? *Phonology* 10 (1993): 43-81.

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)

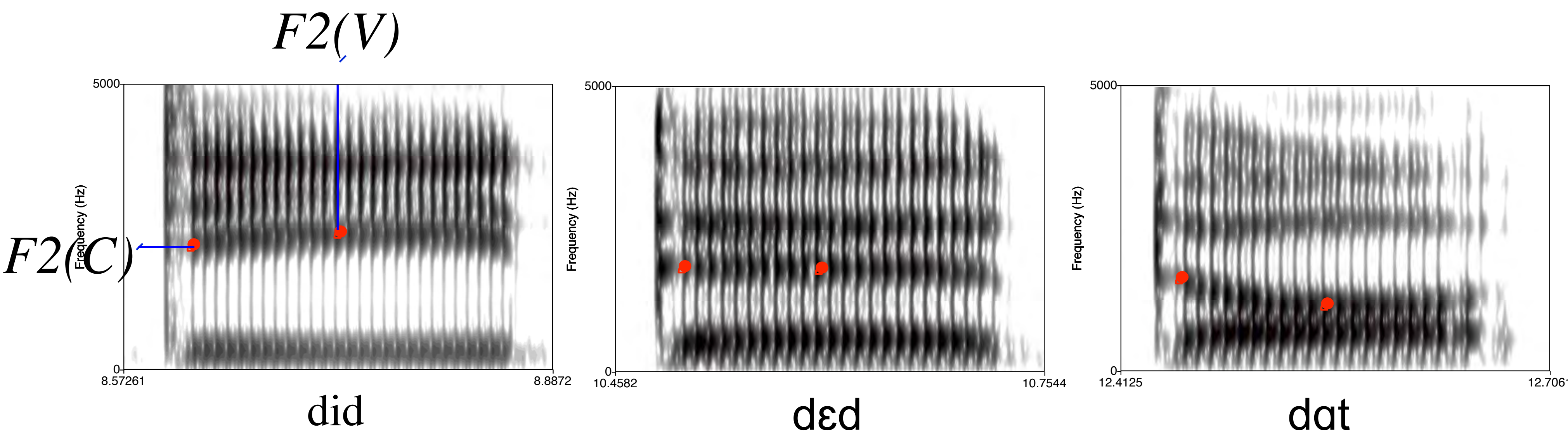


Figure removed due to copyright restrictions.
Source: Figure 1, Fowler, Carol A. "Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation." Attention, Perception, & Psychophysics 55, no. 6 (1994): 597-610.

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)

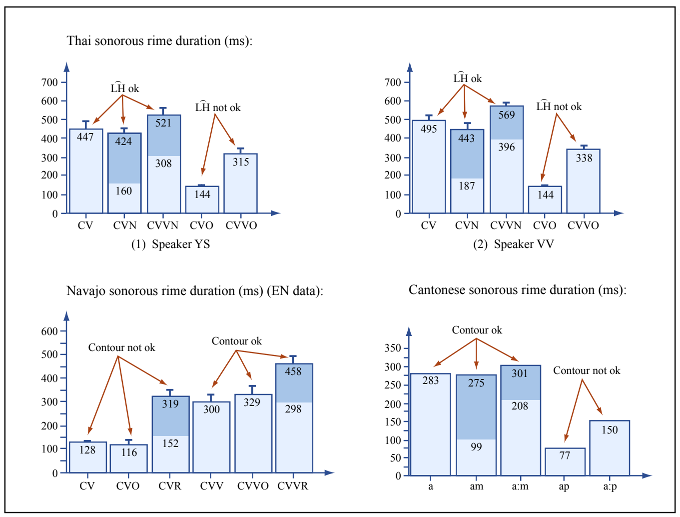
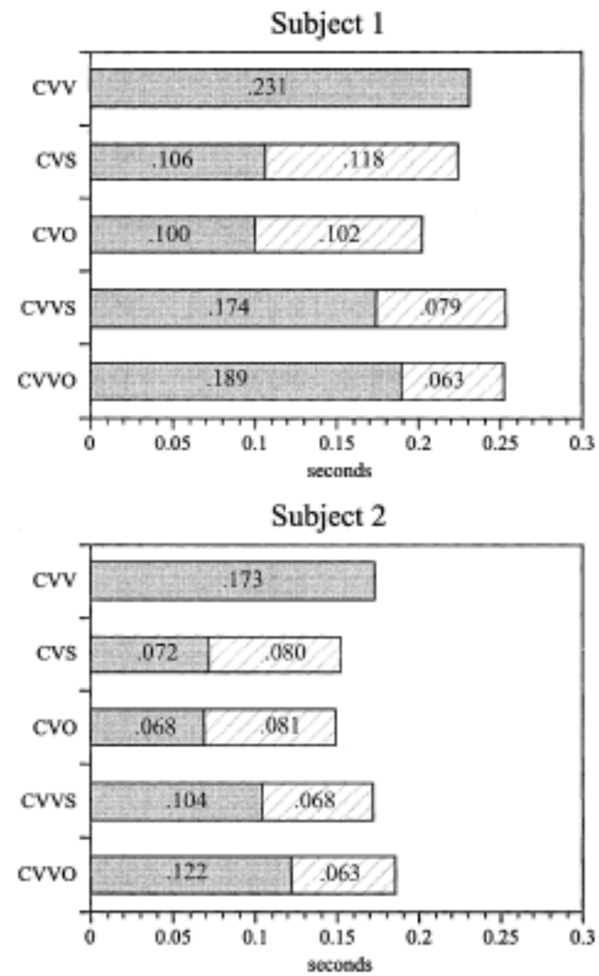


Image by MIT OCW.

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



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

- Cho, T., and P. Ladefoged (1999). Variations and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27, 207-229.
- Cohn, A. (1990) Phonetic and Phonological Rules of Nasalization. UCLA Working Papers in Phonetics 76.
- Cohn, A. (1993). Nasalization in English: Phonology or phonetics? Phonology 10, 43-81.
- Dave, Radhekant (1977). Retroflex and dental consonants in Gujarati. A palatographic and acoustic study. *Annual Report of the Institute of Phonetics, University of Copenhagen* 11: 27-156.
- Disner, S. (1978) Vowels in Germanic Languages. UCLA Working Papers in Phonetics 40.
- Disner, Sandra F. (1983). *Vowel quality: The relation between universal and language-specific factors*. Ph.D. dissertation, University of California, Los Angeles.
- Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18..
- Keating, P. (1985). Universal phonetics and the organization of grammars. V.Fromkin (ed.) Phonetic Linguistics. Academic Press, 115-32.

References

- Ladefoged, Peter, and Peri Bhaskararao (1983). Non-quantal aspects of consonant production: A study of retroflex consonants. *Journal of Phonetics* 11: 291–302.
- Ladefoged, P. & T. Cho (2001) Linking linguistic contrasts to reality: The case of VOT. In N. Gronnum & J. Rischel (eds.), Travaux Du Cercle Linguistique De Copenhague, vol. XXXI. (To Honour Eli Fischer-Forgensen.) C.A. Reitzel, Copenhagen.
- Li, Y, Lee, T, & Qian, Y. (2004). Analysis and modeling of f0 contours for Cantonese text-to-speech. ACM Transactions on Asian Language Information Processing, 3(3), 169-180.
- Marsh, J. (1969). Mantjiltjara phonology. *Oceanic Linguistics* 8.2.
- Platt, J.T. (1972). *An Outline Grammar of the Gugada Dialect: South Australia*. Australian Aboriginal Studies, No. 48, Australian Institute of Aboriginal Studies, Canberra.
- Zhang, Jie (2002). *The effects of duration and sonority on contour tone distribution--A typological survey and formal analysis*. Routledge, New York.