Preface

This set of materials is designed to be used as handouts accompanying an introductory course in phonology, particularly at the undergraduate level. It is specifically intended to be used in conjunction with Stephen Marlett’s 2001 textbook, *An Introduction to Phonological Analysis*. The latter is currently available for free download from the SIL Mexico branch website. However, this course packet could potentially also be adapted for use with other phonology textbooks. The materials included here have been developed by myself and others over many years, in conjunction with courses in introductory phonology taught at SIL programs in North Dakota, Oregon, Dallas, and Norman, OK. Most recently I have used them at GIAL. Two colleagues in particular have contributed significantly to many of these handouts: Jim Roberts and Steve Marlett, to whom my thanks. I would also like to express my appreciation and gratitude to Becky Thompson for her very practical service in helping combine all of the individual files into one exhaustive document, and formatting it for me. Many of the special phonetic characters appearing in these materials use IPA fonts available as freeware from the SIL International website.

Unless indicated to the contrary on specific individual handouts, all materials used in this packet are the copyright of Steve Parker. These documents are intended primarily for educational use. You may make copies of these works for research or instructional purposes (under fair use guidelines) free of charge and without further permission. However, republication or commercial use of these materials is expressly prohibited without my prior written consent.

Steve Parker
Graduate Institute of Applied Linguistics
Dallas, 2013
Table of contents: list of handouts included in this packet

Day 1 lecture topic: Introduction to phonology
- Lecture notes (PowerPoint slides) ................................................................. 7

Day 2 lecture topic: Overview of articulatory phonetics and phonological features
- Lecture notes (PowerPoint slides) ................................................................. 9
- Homework on the phonological features of consonants ............................... 11
- IPA and APA symbol charts ........................................................................ 13

Day 3 lab topic: Practice with phonetics and phonological features
- Homework on the phonological features of vowels ..................................... 15
- Classical generative distinctive features for vowels ................................... 17

Day 4 lecture topic: Contrast in Identical Environment and Complementary Distribution
- Lecture notes (PowerPoint slides) ................................................................. 18
- Homework instructions .................................................................................. 21
- Homework on English minimal pairs .......................................................... 22
- Homework on Italian nasals .......................................................................... 23
- Phonemics reading, Part 1 ........................................................................... 24

Day 5 lab topic: Phonemics
- Homework instructions .................................................................................. 31
- Hints for homework ....................................................................................... 32
- Classical distinctive features charts .............................................................. 33
- Practical premises for phonological analysis ................................................ 35

Day 6 lab topic: Phonemics
- Check Yourself 1 .......................................................................................... 37
- Huariapano (an extinct Peruvian language) ................................................. 39
- Phonemics – analytical procedures flow chart ............................................ 40
- Resources for studying phonology ............................................................... 41

Day 7 lab topic: Phonemics
- Homework on Chamicuro laterals (Peru) ....................................................... 42

Day 8 lecture topic: Suspect Pairs and Contrast in Analogous Environment
- Lecture notes (PowerPoint slides) ................................................................. 45
- Homework instructions .................................................................................. 47
- Phonetically Similar Segments chart ............................................................ 48

Day 9 lab topic: Suspect Pairs
- Homework instructions .................................................................................. 49
- Data for suspect pairs exercises .................................................................... 50
Day 10 lecture topic: Free Variation
- Lecture notes (PowerPoint slides) ................................................................. 51
- Homework instructions .................................................................................. 53
- Phonetic chart consonants ............................................................................ 54
- Phonetic chart vowels ..................................................................................... 55
- Phonemic chart consonants .......................................................................... 56
- Phonemic chart vowels .................................................................................. 57
- Phonemics reading (continued) ..................................................................... 58

Day 11 lab topic: Free Variation
(none)

Day 12 lab topic: Phonemics
(none)

Day 13 lecture topic: Syllables and syllabification
- Lecture notes (PowerPoint slides) ................................................................. 66

Day 14 lecture topic: Syllabification and interpretation
- Lecture notes (PowerPoint slides) ................................................................. 68
- Interpretation / resegmentation charts ......................................................... 70
- Interpretation: resegmentation reading ......................................................... 74

Day 15 lab topic: Syllable structure and interpretation
- Procedures for syllable structure exercises ................................................. 83
- Classical formal devices ................................................................................. 85

Day 16 lab topic: Syllable structure and interpretation
- Spanish consonant clusters ......................................................................... 91
- Isthmus Zapotec (Mexico) ............................................................................. 92

Day 17 lecture topic: Morphophonemics
- Lecture notes (PowerPoint slides) ................................................................. 93
- Procedures for morphophonemic analysis .................................................... 95

Day 18 lab topic: Morphophonemics
- Chamicuro morphophonemic data ............................................................... 96
- Check Yourself 2 ............................................................................................ 97

Day 19 lab topic: Morphophonemics
- Homework instructions ................................................................................ 99

Day 20 lecture topic: Distinctive features and their hierarchical organization
- Lecture notes (PowerPoint slides) ................................................................. 100
- Distinctive features (with Marlett’s feature geometry tree) ......................... 101
- Feature Geometry Tree (adapted) ................................................................. 102
Day 21 lab topic: Practice with distinctive features
- Practice with distinctive features and natural classes (in class) ........................................... 103
- Homework on distinctive features (for practice only) ................................................................. 106
- Homework on feature geometry trees to hand in on Day 22 ....................................................... 107

Day 22 lecture topic: Rule ordering
- Lecture notes (PowerPoint slides) ................................................................................................. 108
- Rule ordering homework: Czech (Marlett p. 174) ........................................................................ 110
- Some practical tips on writing formal rules using feature geometry trees ..................................... 112
- Rule ordering reading ..................................................................................................................... 119

Day 23 lab topic: Morphophonemics
- Fore (Papua New Guinea) .............................................................................................................. 128
- Lamba exercise ............................................................................................................................... 129
- What to do with your data .............................................................................................................. 130

Day 24 lecture topic: Suprasegmentals (stress and length)
- Lecture notes (PowerPoint slides) ................................................................................................. 135
- Evaluating analyses ......................................................................................................................... 137

Day 25 lecture topic: Redundancy
- Lecture notes (PowerPoint slides) ................................................................................................. 138
- Illustrative examples for redundancy ............................................................................................. 141

Day 26 lab topic: Morphophonemics
(none)

Day 27 lab topic: Suprasegmentals (stress and length)
- Latin stress ........................................................................................................................................ 143
- Navaho (USA) ................................................................................................................................. 145

Day 28 lecture topic: Tone
- Lecture notes (PowerPoint slides) ................................................................................................. 146
- Quote from Hale ............................................................................................................................... 148
- Suprasegmentals reading ................................................................................................................. 149

Day 29 lab topic: Tone
- Mandarin Chinese .......................................................................................................................... 162
- Suena (Papua New Guinea) ............................................................................................................ 163
- Mixteco of Santo Tomás Ocotepec (Mexico) .................................................................................. 164
- Shona tone problem ....................................................................................................................... 165
- Tone processes reading .................................................................................................................. 166

Day 30 lecture topic: Sonority
- Lecture notes (PowerPoint slides) ................................................................................................. 172
- Universal hierarchy of relative sonority .......................................................................................... 175
- Check Yourself 3 ............................................................................................................................ 176
Day 31 lab topic: Tone
-Western Apache (USA) ................................................................. 178
-A list of some of the factors which affect tone ........................................ 179

Day 32 lab topic: Morphophonemics
(none)

Day 33 lab topic: Morphophonemics
-Artificial language “J” .................................................................... 180

Day 34 lecture topic: Suppletion and abstractness
-Lecture notes (PowerPoint slides) ..................................................... 181

Day 35 lab topic: Morphophonemics
-Palauan exercise ............................................................................. 183

Day 36 lecture topic: Orthography
-Lecture notes (PowerPoint slides) ..................................................... 184
/Data and instructions for Katukina problem ......................................... 186
-“Brute Force” rules .......................................................................... 188
-Orthography reading ........................................................................ 189

Day 37 lab topic: Orthography
(none)

Day 38 lecture topic: Motivation for feature geometry
-Lecture notes (PowerPoint slides) ..................................................... 201
-Phonology humor ............................................................................ 204
-Prose rules to formalize .................................................................... 206

Day 39 topic: Final exam
(none)

Day 40 lecture topic: Praat and acoustic phonetics
-Lecture notes (PowerPoint slides) ..................................................... 208
Homework for Day 2:

- read the preface and chapter one in the Marlett textbook
- also, read from page 3 through the top of page 6 in the Schane book (on reserve in the GIAL library)

Day 1: The big picture

In today’s class, we will give a brief explanation of what phonology is, what its crucial areas of concern are, and how it differs from phonetics.

I. What is _________________?

A. short definition: the study of sound systems in natural languages
B. explain the notion of a system
C. patterns of rule-governed behavior — formalisms for representing sounds and rules
D. predictable (redundant) information vs. non-predictable (distinctive) information, e.g., [t] vs. [ð] and /t/ vs. /d/ in English
E. universals of language, e.g., every language has some form of the low vowel /a/
F. practical applications: orthography

II. Phonology vs. _________________

A. Branches of phonetics:
   1. articulatory
   2. acoustic
   3. auditory
B. Differences between the two:
   Phonetics is objective and is concerned with the physical properties of all speech sounds (in any language); concrete, involves the raw, uninterpreted data; an outsider’s view of the sound matter.
   Phonology studies the patterns of sounds within the system of individual languages; concerned with the structure and function of sounds; abstract; involves analyzing raw data to get a unified picture of how sounds behave in the framework of a whole language system; an insider’s view of the sound matter.
C. Phonetics is the point of departure for phonology.
   Phonology is the interface between phonetics and grammar (syntax).
D. Phonetics deals with nonsystematic surface forms.
   Phonology deals with an abstract, systematic level of structure, e.g., English [θ] vs. /θ/.
E. A phonological analysis attempts to capture the native speaker’s implicit knowledge of his or her language, and how we know how to use the rules of the language when confronted with a novel form or a novel utterance, in order to produce these correctly; e.g., the plural of wug or glarm.
III. Issues of phonology: three principal areas are involved; in each of these areas we seek to discover and describe the patterns and regular behavior of sounds.

A. __________________________ of sounds and sound units which occur in a particular language; cf. the traditional notion of the phoneme.

1. The units of phonological analysis in any given language occur in particular patterns that have to be discovered for each language; such regular patterns are the subject of a phonological description of the units a given language employs.

2. There tends to be symmetry among these basic units and their interrelationships or arrangements.

B. Permissible __________________________ or combinations of these sounds in syllables and words. Most languages have restrictions on which of these units can combine with each other, and in what particular order, etc., so there are patterns to be described here. For example, in English, if a word begins with three consonants in a row, which ones can they be?

C. __________________________ for adding, deleting, or changing sounds. Sounds tend to modify other sounds in their environment in regular, patterned ways, and these effects can likewise be captured by phonological rules. Phonetics provides us with the physical basis for this tendency, in the sense of movements of our articulatory organs: I’ll miss you.
Homework for Day 3:
- read Marlett chapter 6
- read Schane chapter 2 (on reserve)
- Marlett exercise 6.1 (for practice only)
- consonant exercise (to hand in)

Day 2: The big picture
In today’s class we are going to give an overview of articulatory phonetics as it relates to phonology, with an emphasis on the types of binary distinctive features we typically use in describing speech sounds and classifying them from a phonological point of view.

Describing

In describing consonant sounds in most languages, four principle parameters are often needed to distinguish them one from another:
- voicing
- point or place of articulation
- manner of articulation
- position of the soft palate (velic port)

For some consonants, additional features are necessary, but these four constitute a minimal set to adequately differentiate the inventory of consonants in most languages.

In terms of phonology, it is helpful to conceive of all speech segments as being comprised of a hierarchical tree-like structure in which these different features are related geometrically to each other:

C
   /   \\
 Laryngeal   Supralaryngeal
       /   \\
   [voice] Soft Palate Manner Place
       /   /   \\
 [nasal] stop fricative lateral glide etc.
            /   \\
         bilabial labiodental interdental alveolar etc.
All of the different types of articulation are interrelated in the following way:

- Speech sounds
  - Consonants
    - Obstruents
      - Stops
    - Sonorants
      - Nasals
      - Liquids
      - Glides
  - Vowels
    - Laterals
    - Rhotics

Describing ____________:

The four main parameters used in describing vowels are the following:

- Height of the tongue
- Frontness/backness of the tongue
- Rounding of the lips
- Tenseness/laxness (ATR = advanced tongue root)

Define the notion of a ____________________:

Informally, it is any group of two or more segments which share one or more phonological features in common. More technically, a natural class is any group of two or more segments for which it takes fewer distinctive features to specify the whole class than it does to specify any one single segment which is a member of that class.

Define some of the classes in terms of the formal, binary distinctive features. For example,

[+sonorant] = vowels, glides (approximants), liquids, and nasals

[-sonorant] = obstruents = stops, affricates, and fricatives
Homework on the phonological features of consonants

After reading the materials on phonetics, do the following exercises, to hand in on Day 3. Pay close attention to the sounds in these words, and be careful not to be confused by the spelling.

1. Circle the words that begin with a bilabial consonant sound (when pronounced):
   mat          gnat          sat          bat          rat          pat

2. Circle the words that begin with a velar consonant sound:
   knot          got          lot          cot          hot          pot

3. Circle the words that begin with a labiodental consonant:
   fat          cat          that          mat          chat          vat

4. Circle the words that begin with an alveolar consonant:
   zip          nip          lip          sip          tip          dip

5. Circle the words that begin with an interdental consonant:
   pie          guy          shy          thigh          thy          high

6. Circle the words that begin with an alveopalatal (or palatoalveolar) consonant:
   sigh          shy          tie          thigh          thy          lie

7. Circle the words in which the consonant in the middle is [+voice]:
   tracking         mother         robber         leisure         massive         stomach         razor

8. Circle the words that end with a fricative:
   race    wreath    bush    bring    breathe    bang    rave    rat    ray    rose    rough

9. Circle the words that end with a [+nasal] consonant:
   rain          rang          dumb          deaf          deal

10. Circle the words that end with an oral ([-nasal]) stop (plosive):
pill  lip  lit  limb  crab  dog  hide  laugh  back

11. Circle the words that begin with a liquid consonant:
    nut  lull  bar  rob  one

12. Circle the words that begin with a glide:
    we  you  one  run

13. Circle the words that end with an affricate:
    much  back  edge  ooze

14. Specify the voicing and place of articulation of the consonant sounds in the middle of each of the following words. The first one is already filled in as an example:

<table>
<thead>
<tr>
<th>voiced or voiceless</th>
<th>place of articulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ex. adder + voice</td>
<td>alveolar</td>
</tr>
<tr>
<td>a. father</td>
<td></td>
</tr>
<tr>
<td>b. singing</td>
<td></td>
</tr>
<tr>
<td>c. etching</td>
<td></td>
</tr>
<tr>
<td>d. robber</td>
<td></td>
</tr>
<tr>
<td>e. ether</td>
<td></td>
</tr>
<tr>
<td>f. pleasure</td>
<td></td>
</tr>
<tr>
<td>g. lodger</td>
<td></td>
</tr>
</tbody>
</table>

15. Each of the following groups of sounds consists of members of a natural class. A natural class of sounds all share one or more common properties (features), such as voiced sounds, fricatives, bilabials, nasals, and so on. Name the feature or set of features (as many as you can think of) that defines each class:

a. [ɡ], [b], [d]
b. [f], [p], [m], [v], [b]
c. [z], [v], [s], [ʒ]
d. [t], [z], [d], [s]
e. [m], [n], [ŋ]
f. [ɡ], [k], [b], [d], [p], [t]
<table>
<thead>
<tr>
<th>Syllable</th>
<th>Non-Syllabic</th>
<th>Palatal</th>
<th>Anterior</th>
<th>Posterior</th>
</tr>
</thead>
<tbody>
<tr>
<td>i [i]</td>
<td>e [e]</td>
<td>i [i]</td>
<td>a [a]</td>
<td>i [i]</td>
</tr>
<tr>
<td>u [u]</td>
<td>o [o]</td>
<td>u [u]</td>
<td>o [o]</td>
<td>u [u]</td>
</tr>
</tbody>
</table>

**Notes**

- The table represents a phonetic chart showing vowel sounds.
- The chart includes symbols for various phonetic elements such as fronted, back, and central vowels.
- The symbols are used in phonetic transcription.

**Diagrams**

- The diagrams illustrate the relationships between different phonetic elements.
- The symbols used in the diagrams correspond to the ones in the table.

**Suprasegmentals**

- Vowels
- Secondary stress
- Primary stress

**Other Symbols**

- Open
- Low
- High

**Directions**

- Where symbols appear in pairs, the one to the right represents a raised vowel.
AL 4303, Principles of Phonological Analysis

GIAL

Day 3

Homework on the phonological features of vowels

(to hand in on Day 4)

1. Circle the words that contain a high ([+high]) vowel sound (when pronounced):
   sat suit got meet mud

2. Circle the words that contain a low ([+low]) vowel sound:
   weed wad load lad rude

3. Circle the words that contain a front ([–back]) vowel:
   gate caught cat kit put

4. Circle the words that contain a central or back ([+back]) vowel:
   maid weep coop cop good

5. Circle the words that contain a [+round] vowel:
   who me us foot him

6. Each of the following groups of sounds consists of members of a natural class. A natural class of sounds all share one or more properties (features) in common, such as voiced sounds, fricatives, bilabials, nasals, and so on. Name the feature or features (as many as you can think of) that defines each of the following two classes:

   a. [æ], [u], [i], [e]
   b. [o], [ɔ], [u], [ʊ]

7. In each of the following pairs of words, the underlined sounds differ by one or more phonetic properties (features). State the features by which the sounds differ and also those that they have in common:

   example: phone
   different: –low, +ATR, +round
   same: vowels, +back, –high

   phonic
   different: +low, –ATR, –round
   same: vowels, +back, –high

(vowel HW.pdf)
<table>
<thead>
<tr>
<th></th>
<th>different</th>
<th>same</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>bath</td>
<td>bathe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>cool</td>
<td>cold</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>heal</td>
<td>health</td>
</tr>
</tbody>
</table>

8. A phonetic symbol is actually a “cover term” for a complex of phonetic properties (so called feature matrices) that define the sound. Define each of the symbols below by filling in the features in the matrix, as illustrated:

<table>
<thead>
<tr>
<th></th>
<th>U</th>
<th>I</th>
<th>Æ</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>height</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>backness</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>roundness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tenseness(ADR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Classical generative distinctive features for vowels

<table>
<thead>
<tr>
<th>-back</th>
<th>+back</th>
</tr>
</thead>
<tbody>
<tr>
<td>-round</td>
<td>+round</td>
</tr>
<tr>
<td>+high</td>
<td>+ATR</td>
</tr>
<tr>
<td></td>
<td>-ATR</td>
</tr>
<tr>
<td>-low</td>
<td>+ATR</td>
</tr>
<tr>
<td></td>
<td>-ATR</td>
</tr>
<tr>
<td>-high</td>
<td>+ATR</td>
</tr>
<tr>
<td></td>
<td>-ATR</td>
</tr>
</tbody>
</table>

\[ ù = y \quad ñ = u \quad ù = Y \quad ù = ø \]

(See Appendix A in the Marlett textbook for a formal definition of these features.)
Homework for Day 5:
This is listed and explained on one of the handouts in the course packet.

Day 4: The big picture
In today’s class we are going to explain and illustrate the two most important and basic phonological relationships that can exist between sounds: Contrast and Complementary Distribution. These form the heart of what is often called classical phonemic analysis.

Consider two English words: *sip* and *zip*. The only difference between these two words phonetically is in the beginning consonants: [s] is voiceless but [z] is voiced. This is called a minimal pair. The relationship between [s] and [z] in a minimal pair of words like these is called contrast. Specifically, this is called Contrast in Identical Environments or CIE.

When we observe contrast in a minimal pair, we say (conclude) that the two segments ([s] and [z]) are distinct __________ in this language.

Definition: A phoneme is a sound that can consistently be used to make a (minimal) distinction between two different words in the same language.

Now let’s consider the relationship between another pair of similar sounds in two different languages: oral [a] vs. nasalized [ã].

In Shipibo (a language spoken in the Amazon jungle of Peru), these two segments also contrast, as in the following minimal pair:

[ǐ'sa] ‘bird’
[ǐ'sã] ‘palm oil’

This is also a case of CIE, so we conclude that in this language, /a/ and /ã/ are two separate phonemes. This is confirmed by the fact that this type of contrast can also be found in many other words of the language, including in different phonological environments: unstressed syllables as well as stressed syllables, and word-initial and word-medial syllables as well as word-final position. Furthermore, the same type of relationship (contrast) can also be shown to exist for the other pairs of oral vs. nasalized vowels as well: /i/ vs. /ĩ/, /o/ vs. /õ/, etc. All of these facts are an indication that our analysis is on the right track.

One implication of this conclusion is that in Shipibo, it is probably necessary to write these two phonemes (/a/ and /ã/) in different ways in the practical orthography.

But here’s a very important phonological principle:

**Pike’s First Premise for Phonological Analysis:** Sounds tend to be ____________________________ by their environments.

Now let’s consider the relationship between the two phonetic segments [a] vs. [ã] in English:

<table>
<thead>
<tr>
<th></th>
<th>[a]</th>
<th>[ã]</th>
</tr>
</thead>
<tbody>
<tr>
<td>pot</td>
<td>raw</td>
<td>pawn</td>
</tr>
<tr>
<td>pod</td>
<td>law</td>
<td>pom-pom</td>
</tr>
<tr>
<td>pop</td>
<td>pa</td>
<td>Ron</td>
</tr>
<tr>
<td>Paul</td>
<td>ma</td>
<td>bomb</td>
</tr>
<tr>
<td>par</td>
<td>saw</td>
<td>Tom</td>
</tr>
<tr>
<td>rot</td>
<td>caw</td>
<td>prom</td>
</tr>
<tr>
<td>ball</td>
<td>ha</td>
<td>prom</td>
</tr>
<tr>
<td>top</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prod</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In English, the phone (segment) [a] occurs always and only before an oral consonant OR word-finally, while nasalized [ã] occurs always and only before a nasal consonant. This type of relationship is called complementary distribution. It is the opposite of contrast. In English, there can never be any minimal pairs contrasting oral [a] vs. nasal [ã] in the exact same phonological environment at the same time, as there are in Shipibo. In other words, it is not the case that [a] and [ã] are separate phonemes in English. Rather, they are __________________ of the same phoneme.

Definition: Allophones are predictable phonetic variants of the same phonemic unit.

Now let’s compare and contrast phonemes vs. allophones in the following chart:

<table>
<thead>
<tr>
<th>Phonemes (Shipibo /a/ vs. /ã/)</th>
<th>Allophones (English [a] and [ã])</th>
</tr>
</thead>
<tbody>
<tr>
<td>psychologically significant</td>
<td>psychologically insignificant to the native speaker</td>
</tr>
<tr>
<td>make a difference in meaning</td>
<td>do not make a difference in meaning</td>
</tr>
<tr>
<td>insider’s (native speaker’s) viewpoint</td>
<td>outsider’s (linguist’s) viewpoint</td>
</tr>
<tr>
<td>phonemic transcription/</td>
<td>[phonetic transcription]</td>
</tr>
<tr>
<td>underlying level of</td>
<td>surface level of representation</td>
</tr>
<tr>
<td>representation</td>
<td>phonetic level of representation</td>
</tr>
<tr>
<td>phonemic level of</td>
<td>can usually be written with the same grapheme</td>
</tr>
<tr>
<td>representation</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phonemes (Shipibo /a/ vs. /ã/)</th>
<th>Allophones (English [a] and [ã])</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a/ ≠ /b/</td>
<td>/a/ [x] [y]</td>
</tr>
</tbody>
</table>

How can we determine what the underlying form of a phoneme is?

Choosing underlying (phonemic) forms

All other things being equal, posit (select) as the underlying form of a phoneme that allophone which:

<table>
<thead>
<tr>
<th>Importance</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>most</td>
<td>(1) is least restricted in distribution</td>
</tr>
<tr>
<td></td>
<td>(2) is least modified / affected by its environment</td>
</tr>
<tr>
<td>medium</td>
<td>(3) is more frequent, common, natural, preferred,</td>
</tr>
<tr>
<td></td>
<td>basic, unmarked cross-linguistically (in the majority of the languages of the world)</td>
</tr>
<tr>
<td>least</td>
<td>(4) leads to a more concise, economical, and/or symmetrical phonemic inventory</td>
</tr>
<tr>
<td>importance</td>
<td>(5) the native speaker tends to pronounce in slow, deliberate, careful, emphatic speech</td>
</tr>
<tr>
<td></td>
<td>(6) is most psychologically real to the native speaker</td>
</tr>
</tbody>
</table>

Of these six criteria, the first two are the most crucial. Some of these six criteria may conflict with one another in certain cases; in such a situation it is the overall picture which is most important.

Note that statistical frequency of an allophone in a particular language (the most common one) is not necessarily a valid criterion for choosing an allophone as underlying, such as [pʰ] in English.

To clarify, every phoneme will always have at least one allophone: itself.

What is happening with English vowels is that the nasalization of the following nasal consonant is leaking or spreading backwards onto the preceding vowel.
This phenomenon could be called nasalization, and it is an example of what we call __________________________: a phonological process in which a sound changes and becomes more like (similar to) its environment. Assimilation is the most common and natural type of phonological process which exists among the languages of the world. This particular type of assimilation could be called anticipatory in nature: the articulatory gesture of lowering the soft palate is targeted for the nasal consonant, but it actually begins during the pronunciation of the vowel, in anticipation of the following segment (the nasal at the end of the word/syllable).

Recall Pike’s premise that sounds tend to be modified by their environments.

However, this type of relationship between [a] and [ã] will not necessarily hold true in all languages; for example, recall the case of Shipibo [iša] ‘bird’ vs. [išã] ‘palm oil’, where the nasalized /ã/ is a separate phoneme in this language.

In English there will never be any minimal pairs analogous to the case of Shipibo, contrasting the oral and nasalized varieties of this vowel.

In English, the phoneme /a/ is realized phonetically as [ã] before nasal consonants, and as [a] elsewhere.

This makes for a more natural and systematic prediction. For example, the vowel in the word long is also somewhat nasalized. As a matter of fact, this same type of relationship exists for all of the other vowel phonemes of English as well: these tend to be pronounced with nasalization when the syllable they are in ends with a nasal consonant, and as oral everywhere else.

The implication of these facts is that in English, there is no need to distinguish oral from nasalized vowels in the practical orthography; native speakers will always choose to pronounce the right allophone correctly when reading, since this is a subconscious, automatic reflex that we do without thinking.

In Shipibo, however, this is not the case: since two words can be minimally distinguished by oral vs. nasal vowels, and since this type of situation is very common in this language, they need to write the two types of vowels differently, or else the native speakers would get confused when they are reading. The way this is done in fact is by writing isa vs. isan. Other possible ways to represent nasalization of vowels in practical orthographies is by:

Underlining:  isa

A tilde on top of the vowel symbol:  isã

Or a hook underneath the vowel symbol:  isq:

Illustrate the phenomenon of allophones with the Superman analogy. Also cf. chameleons.

Summary of today’s lesson:

The __________________________ principle

Two or more phonetically similar sounds which are in a relationship of complementary distribution in a specific language can and should be analyzed as automatic phonetic variants (allophones) of the same, basic, underlying phonological unit (phoneme) rather than as distinct, contrastive phonemes.
Your homework for Day 5 consists of the following items:

(1) Read the “phonemics” handout (in this course packet or online).

(2) Fill in the “homework on English minimal pairs” sheet, to turn in. This is on a separate handout, also in the course packet (and available online).

(3) Do the Karuk exercise in Marlett, p. 114.

Directions: analyze the two pairs of sounds [a] vs. [aː] and [i] vs. [iː]. The symbol [ː] after a vowel segment indicates that it is longer than normal. For each of these two pairs of vowels, do they contrast, or are they in complementary distribution? If they contrast, give your evidence — a minimal pair of Karuk forms, transcribed phonetically, with the corresponding English glosses (a translation in English). If they are in complementary distribution, state the relevant environments for each allophone (phonetic variant), and suggest which of the two allophones should be chosen as the underlying (phonemic) form, and why. State a rule in prose (in your own words) predicting how the phoneme is pronounced. Be as concise as possible in stating your rule. Also, state what implications your solution (analysis) has for the orthography (practical writing system) of this language. This problem is for practice only, i.e., it is not to be handed in. A brief answer can be found in Marlett, p. 118.

(4) Seri, found in Marlett, p. 114.

Analyze the pair of sounds [s] vs. [ʃ]. The latter is a voiceless alveopalatal fricative, equivalent to [š] in the Americanist system. Follow the same steps outlined for the Karuk problem above in (3). This problem is to be handed in tomorrow at the beginning of the class.

(5) Italian (on another sheet in this packet and online)

Analyze the pair of sounds [n] vs. [ƞ], where [ƞ] is a velar nasal. Follow the same steps outlined for the Karuk problem above in (3). This exercise is not to be handed in; refer to Marlett pp. 129 and 132-33 for the answer.
Homework on English minimal pairs

For each of the following pairs of English consonant phonemes, try to find two minimal pairs. One pair should show the contrast in word-initial position, and the other one in word-final position, whenever this is possible. Some phonemes do not occur in both positions. Use the first pair as an example. It’s okay to use regular English spelling rather than phonetic transcriptions (for this assignment only, which is to hand in on Day 5).

<table>
<thead>
<tr>
<th>Initial position</th>
<th>Final position</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p b/</td>
<td>pie — by</td>
</tr>
<tr>
<td>/t d/</td>
<td>rope — robe</td>
</tr>
<tr>
<td>/k ɡ/</td>
<td></td>
</tr>
<tr>
<td>/f v/</td>
<td></td>
</tr>
<tr>
<td>/s z/</td>
<td></td>
</tr>
<tr>
<td>/m n/</td>
<td></td>
</tr>
<tr>
<td>/n ɳ/</td>
<td></td>
</tr>
<tr>
<td>/l l/</td>
<td></td>
</tr>
<tr>
<td>/w j/</td>
<td></td>
</tr>
<tr>
<td>/l w/</td>
<td></td>
</tr>
<tr>
<td>/l j/</td>
<td></td>
</tr>
<tr>
<td>/p f/</td>
<td></td>
</tr>
<tr>
<td>/b v/</td>
<td></td>
</tr>
<tr>
<td>/t θ/</td>
<td></td>
</tr>
<tr>
<td>/d ɹ/</td>
<td></td>
</tr>
<tr>
<td>/s ř/ (=ʃ)</td>
<td></td>
</tr>
<tr>
<td>/ʃ ʃ̆/ (=tʃ)</td>
<td></td>
</tr>
<tr>
<td>/ɛ j̆/ (=dʒ)</td>
<td></td>
</tr>
<tr>
<td>/s ə/</td>
<td></td>
</tr>
<tr>
<td>/s ŋ/</td>
<td></td>
</tr>
<tr>
<td>/l n/</td>
<td></td>
</tr>
</tbody>
</table>

Morris Halle and George N. Clements, *Problem Book in Phonology: A Workbook for Courses in*
Homework on Italian nasals

(cf. Marlett page 129 and the reference therein for discussion)

In the following phonetically-transcribed data from Italian, focus on (analyze) the phonological relationship that exists between two nasal consonants: [n] (alveolar) vs. [ŋ] (velar). In these data, a vowel marked with an acute accent above it, such as [á], [ó], etc., indicates primary stress. Other symbols include [j] for a voiced alveopalatal affricate (IPA [dʒ]) and [tʃ] for a voiceless alveolar affricate.

1. [né̞ro̞] ‘black’  
2. [sapò̞ne] ‘soap’  
3. [bjáŋko] ‘white’  
4. [tornáre] ‘return’  
5. [jé̞nte] ‘people’  
6. [nónno] ‘grandfather’  
7. [áŋke] ‘also’  
8. [fíne] ‘end’  
9. [úŋga] ‘(finger)nail’
10. [ónda] ‘wave’  
11. [stáŋko] ‘tired’  
12. [dánt’a] ‘dance’  
13. [fáŋgo] ‘mud’  
14. [líŋwa] ‘tongue, language’  
15. [téŋgo] ‘I have’  
16. [lúŋgo] ‘long’  
17. [kwantúŋkwe] ‘although’
At this point we might be tempted to think that a thorough study of phonetics was sufficient for the description of the sound systems of human languages. A beginning assumption (a linguist would call this the “null hypothesis”) might be that each phonetically distinguishable segment should be represented by a distinct symbol. A straightforward phonetic transcription would then serve as a practical orthography (or writing system) for any language; each sound would have its own letter in the alphabet.

After closer inspection, however, we soon find that we must abandon this position. For one reason, sounds tend to fluctuate. No two repetitions of the “same” utterance are ever completely identical phonetically. If we are using a careful, narrow phonetic transcription, we would have to use a large, cumbersome repertoire of symbols in our writing system. But worse than that, it would be confusing to read. Native speakers of a language tend to ignore insignificant fluctuations in the sounds; they only “hear” or distinguish a limited number of sound units which have a significant function in their language. For this reason a practical orthography should not be strictly phonetic.

A related reason for not limiting our description of the sound system of language to phonetics is the fact that sounds tend to be modified by their environment. Sounds slur into one another. As a result, “one sound” may be realized phonetically in different ways depending on the context in which it appears. In English the “p” sound shows up as [pʰ] (aspirated) at the beginning of a word or of a stressed syllable, and as [p] (unaspirated) in most other places. Because this variation in sound is predictable, we do not need to specify the phonetic details in every case.

This leads us to another major reason for not being satisfied with phonetic transcriptions. We have said that a phonological analysis must account for the patterns and systematic regularities in the sound system of language. We have just noted that the variability in sounds is to some extent predictable, because sounds are modified by their environment. If we transcribed every sound at the phonetic level and left it at that, we would miss significant generalizations about the patterns of particular sounds and their distributions. On another level, the overall sound systems of language have a remarkable tendency toward symmetry. If we ignore these regularities, we cannot adequately describe the sound system of language, let alone explain its relevant features.

In summary, we find that certain differences between sounds serve the important function of distinguishing the meaning of one utterance from another. Other differences between sounds in a language are not significant, and are predictable from the environment. Our first goal in phonological analysis, then, is to identify the significant sounds of a language. These distinctive sound units, called phonemes, are those capable of distinguishing words with different meanings. The various realizations of a phoneme, determined by the environment in which each occurs, are called the allophones of the phoneme. The occurrence of these allophonic variants is predictable
and nondistinctive, whereas the occurrence of the phoneme itself is unpredictable from the phonetic environment. The analytic procedures outlined below are designed to identify the phonemes of a language, and also to describe the distribution of the allophones by the environments in which each is found. The whole of this analytic method is referred to as phonemics.

1. Preliminary procedures to phonemic analysis

Pike (1947) describes a set of procedures that are important first steps for field linguists. In order to identify the phonemes of a language, the data should be collected and arranged as follows:

1. Record the data phonetically in as much reliable detail as possible.
2. Assume the accuracy and completeness of the data.
3. Make a phonetic chart of all sounds encountered.
4. List suspicious pairs of sounds.
5. List nonsuspicious sounds.

We will not comment on the first two of these here, although in actual field work they are of considerable concern. (For problems in this course, it is to be assumed that the phonetic data given are accurate and complete in all aspects relevant to the problem.) The last three steps require further comment, however.

The phonetic chart of step 3 should be organized as we have done in our study of phonetics: consonants arranged by point of articulation and manner of articulation, and vowels arranged by tongue height and tongue backness. The other parameters for describing segments should also be indicated when relevant, e.g., voiced vs. voiceless consonants, or tense and lax vowels; similar segments should be grouped together logically. Each segment found in the data should be clearly identified, i.e., labeled with enough features to distinguish it from all others. It is only as we chart the segments that we can observe the patterns, and pay attention to phonetic similarities and differences between segments. The phonetic chart serves as the working basis for the analysis. Before undertaking further analysis, our beginning assumption is that each sound on the chart will be a separate phoneme in the language.

Is in fact each sound on the phonetic chart a separate phoneme in the language? To determine this, we will examine pairs of sounds, to see if the phonetic differences between them makes a functional (phonemic) difference in the language, or whether the native speaker can ignore the difference (such as in the practical orthography). For certain pairs, such as [k] – [m], the phonetic differences are so great that they would not be predictable, and so we assume that these two segments are separate phonemes. The cases that are not so obvious are the ones that require closer attention; we are thus led to identify suspicious pairs of sounds.

Suspicious pairs of sounds, as mentioned in step 4 above, are sounds which are phonetically similar, making their phonological status suspect. The sounds usually differ in only one, or perhaps two, phonetic features. Because of their similarity, the two members of the pair may prove not to be separate phonemes, but simply allophones of the same phoneme. If this is the case, the phonetic differences between the sounds may be explainable by the environments in
which they occur. The identification of suspicious pairs is based on linguists’ experience with a wide variety of languages. Pairs of sounds which have turned out to be allophonic variants in other languages are suspicious, and should be investigated closely when we work on a new language, because similar patterns of behavior often reoccur. Admittedly, the identification of suspicious pairs is somewhat subjective, and different pairs of sounds will exhibit different degrees of suspiciousness. The following guidelines are not exhaustive, but are indicative of the principal types of phonetically similar segments that deserve attention for phonemic analysis. When working on a new language, the researcher should constantly be alert to identify additional pairs based on his or her observations and preliminary hypotheses about the sound system. Here, then, are types of sound similarities that are notably suspicious:

In general:

1. **Voiced-voiceless** pairs of sounds (like [t] vs. [d]). Such pairs are always suspicious, but are most common among the obstruents.

2. **Modifications.** Any “modified” sound will be suspicious with its plain (unmodified) counterpart. This usually involves an additional feature superimposed on a segment. For example, palatalized or glottalized (ejective) consonants such as [pʰ] or [pʼ] may be considered as modified versions of the nonpalatalized, nonglottalized [p]. Thus we have the suspicious pairs [pʰ] – [p] and [pʼ] – [p]; the pair [pʰ] – [pʼ] may also be considered suspicious, but only secondarily so if our data also include the plain [p]. Among the vowels, nasalization should also be considered a modification, so that [a] and [ã] form a suspicious pair, etc.

For specific groups of segments:

1. Among the **obstruents**, stops, affricates, and fricatives are all suspicious with one another if they have the same or slightly different point of articulation. For this purpose we distinguish three main groupings by point of articulation: (a) sounds made with the lips are labials (bilabial and labiodental); (b) sounds made with the tip or front part of the tongue are called coronals (interdentals, dentals, alveolars (including retroflexed), alveopalatals, and possibly palatals); and (c) sounds made with the back part or body of the tongue are dorsals (palatals, velars, and uvulars). Within any one of these three regions, pairs of obstruents will be suspicious. There are degrees of suspiciousness here; the most similar (and thus most suspicious) pairs should be subjected to further analysis before less similar pairs. Obstruents are not in general suspicious with sonorants.

2. Among the **nasals**. All nasals are suspicious with one another, regardless of point of articulation. This is especially true of those nasals articulated with the tongue, such as [n], [ŋ], [ɲ], etc. The bilabial [m] may not be as suspicious with these others, but there are languages in which even [m] and [n] are allophones of the same phoneme.

3. Among the **liquids**. As in the case of the nasals, all laterals and rhotics (r-sounds) are suspicious with one another, regardless of point of articulation. The fact that the point of articulation is not so relevant for these sounds is undoubtedly due to their sonorant nature,
which is what gives them their characteristic timbre, rather than their point of articulation. In a number of languages, l’s and r’s are allophones of one another, despite differences in their point of articulation.

4. Among the glides. Glides, being in an intermediate position between the vowels and consonants (some people even call them *semivowels* or *semiconsonants*), are suspicious with a couple of different types of segments. First, they are suspicious with their corresponding high vowels – [j] with [i], [w] with [u], etc. But they are also suspicious with voiced fricatives at the same point of articulation. (Remember that the articulation of a glide involves close proximity of the articulators, and that the position cannot be maintained for long without causing audible friction.) So the palatal [j] is suspicious with [j]; the labiovelar [w] is suspicious with [γ] or [γʷ] in the velar position, and also with the bilabial [β].

5. Among the vowels. The range of vowel articulations is spread over a triangle-like area whose extremes are those of the vowels [i], [u], and [a]. These extreme vowels are not normally suspicious with one another, but because of the relatively fluid nature of vowel articulations, a number of types of suspicious pairs can be cited.
   a. Any vowel is suspicious with its close neighbors. Pairs such as [i] – [e], or [ε] – [æ], or [ʌ] – [ə] are suspicious, and worth subjecting to further analysis.
   b. Corresponding tense and lax vowels, such as [i] – [ɪ], [e] – [ε], [u] – [ʊ], [o] – [ɔ], etc. are always suspicious.
   c. All low vowels are suspicious with one another, such as [æ] – [a], [a] – [ɒ], etc., since they are all in one corner of the vowel triangle. All mid vowels may also be suspicious with one another, but only secondarily so. By similar reasoning, all high front vowels are suspicious with one another, as are all high back vowels.
   d. Any pair of vowels that differs in just one feature is suspicious, such as [i] – [y], which differ in values for the feature [round].

Finally, we note that there may be a number of sounds that are not suspicious with any others. We list these, as indicated in step 5 above, as sounds which we shall assume to be separate phonemes, unless we have other reasons to suspect their phonemic status.

2. Phonemic analysis

Now we are ready to proceed with the analysis. Having identified pairs of sounds that are suspiciously similar, we want to find out if the differences between such sounds can make a phonemic distinction, that is, a difference in meaning between words. Classical phonemics uses three important concepts in determining the phonemic status of segments — **contrast**, **complementary distribution**, and **free variation**. As we discuss each of these in turn, notice how important to the analysis is the principle of **distribution**, that is, the positions and environments in which the segments in question can occur.

2.1 Contrast in identical environments (CIE)

Some phonetic differences are very important. Consider the following data from English:
[sip] ‘sip’  [zip] ‘zip’

The slight differences in pronunciation between these two words distinguishes utterances which have different meanings. The differences between the two utterances (phonologically) may be isolated as a difference in the initial segment. The first word has a voiceless sibilant and the second word has a voiced sibilant. Therefore, in English the difference between [s] and [z] is very important since substituting one of these sounds for the other one can change the meaning of a word. In the above pair of words, [s] and [z] contrast in identical environments because the words they are in are otherwise identical phonetically and yet differ in meaning. Words like “sip” and “zip” are called a minimal pair and are obviously prime data for illustrating contrast. Since the sounds [s] and [z] contrast in English, the English alphabet needs to contain a special symbol for each sound; in this case the symbols “s” and “z” are used.

Now examine the following data, also from English, and decide if the sounds [s] and [ʃ] contrast in this language:

[ʃejv] ‘shave’  [ʃel] ‘shell’
[sejv] ‘save’  [sel] ‘sell’

The segments [s] and [ʃ] contrast in identical environments in English and so the spellings of the words will have to distinguish the two sounds. In this case, English spelling normally uses the digraph (sequence of symbols) “sh” to represent [ʃ].

Examine the following data from Chumash, an extinct language formerly spoken in California, and decide if the sounds [s] and [ʃ] contrast:

[ʃax] ‘illegitimate child’  [mes] ‘to cross over’
[meʃ] ‘sack, bag’  [sax] ‘grass (species)’

The sounds [s] and [ʃ] contrast in identical environments in Chumash and would therefore need to be represented by different symbols in the orthography.

2.2 Complementary distribution (CD)

2.2.1 Influence by other segments. Sometimes two similar sounds can never be found to contrast and it seems that the difference between them can be explained as being due to the environments in which they occur. Consider the following data from Japanese:

[sakana] ‘fish’  [ʃigoto] ‘work’
[soko] ‘that place’  [watakuʃi] ‘I, me’
[sunders] ‘living, dwelling’
[arimasen] ‘there is not’

The segments [s] and [ʃ] do not contrast in Japanese. They are always found in different environments. We can use “Y-diagrams” such as those below to show the immediate
environments of these two sounds in the data given above. On each side of the central line, we put the segment(s) that precede and follow the item (sound) under investigation. (The symbol # indicates the boundary – either beginning or end – of a word.)

<table>
<thead>
<tr>
<th>s</th>
<th>akana</th>
<th>#</th>
<th>igoto</th>
<th>#</th>
<th>a</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>oko</td>
<td>wataku</td>
<td>i</td>
<td>#</td>
<td>o</td>
<td>u</td>
</tr>
<tr>
<td>#</td>
<td>ûnde</td>
<td>a</td>
<td>e</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[s] occurs always and only before [i], and [ʃ] never occurs immediately before [i].

When the occurrence of one segment in a certain environment precludes the possibility of another phonetically similar segment occurring in that same environment, these two segments are said to be mutually exclusive. They have a complementary distribution since they never occur in the same environment. They do not contrast.

\[
\begin{array}{c|c}
[ʃ] & \text{complementary distribution} \\
\text{(before [i])} & [s] \\
& \text{(elsewhere)}
\end{array}
\]

When working with a pair of phonetically similar sounds that may be mutually exclusive, it is important to look for differences in their phonological environments that might account for the phonetic differences between the two segments since sounds tend to influence other sounds. For example, in the case of the Japanese sibilants, it is obvious that the feature [high] of the high front vowel [i] carries over to the sibilant that precedes it (in an anticipatory fashion). The feature [high] for sibilants in Japanese is not important for meaning. It does not cause a change of meaning in words in that language. When the differences between any two phonetically similar sounds are not distinctive, i.e., when the segments do not and cannot contrast in a particular language, but instead are mutually exclusive, we consider these sounds to be allophones of one phoneme and therefore represent them with the same letter (or symbol) in the alphabet of that language.

Transcriptions enclosed in slashes /.../ indicate that the data have been phonemicized, i.e. that the allophonic detail is not represented. The word [ʃiɡo] ‘work’ in Japanese would be written phonemically as /siɡo/. The fact that the first segment of this word is pronounced phonetically as alveopalatal is predictable and so it would be superfluous to supply this information in the phonemic transcription. In other words, the feature [alveopalatal] is a nondistinctive feature of the phoneme /s/; it is supplied by the context in which /s/ occurs in Japanese.
To represent the phoneme in its “basic” or “underlying” form, we choose the symbol for the allophone which is least restricted in distribution, and/or least affected (modified) by its environment. In the Japanese example above, we decided that [s] and [ʃ] are allophones of the same phoneme. Of these two, the [ʃ] only occurs in a very specific environment – preceding [i]. The [ʃ] is also influenced directly by the environment: the following [i] is causing the point of articulation of /s/ to become alveopalatal (thus we call this type of process palatalization). The [s], on the other hand, is not restricted in its distribution in this way. It could occur in almost all environments (everywhere but before [i]). It is also the least affected by these surrounding environments. Thus, we choose the symbol [s] to represent the phoneme as a whole, enclosing it now in diagonal slashes – /s/.
Your homework for Day 6 consists of the following three exercises:

1. Sierra Nahuat (p. 129 in the Marlett textbook)

Analyze the pairs of sounds [w] vs. voiceless [w] and [j] vs. its voiceless counterpart. (The voiceless glides are transcribed in the data in the book with a circle underneath. Note that in the Phonetics course one or both of these may be transcribed differently, for example, with the circle above the [j] rather than below it. Note also that the [j] with circle underneath is not intended here to represent a voiced palatal fricative; be careful since these symbols might be confused.) Are these pairs of segments ([w] vs. voiceless [w] and [j] vs. voiceless [j]) in contrast, or in complementary distribution? If they contrast, give your evidence (minimal pairs). If they are in complementary distribution, state which member of the pair should be chosen as the underlying form, and why. State the rule (in prose) which describes how the relevant phonemes are pronounced. In stating your prose rule, generalize as much as possible. Also, discuss the implications of your phonological analysis for the practical orthography of this language. This problem is not to be handed in; the answer is on p. 132.

2. Tetelcingo Nahuatl (p. 130)

Follow the directions for this problem in the book, as well as those given above for Sierra Nahuat. Note that it is possible for a phoneme to have more than two allophones, as the book indicates. The symbol [ɬ] represents a voiceless lateral fricative. The answer to this problem will be collected on Day 6.

3. Guanano (p. 116)

Analyze the series of stops [pʰ p b], [tʰ t d], and [kʰ k ɡ]. Follow the same general procedures as outlined above. You may have worked through part or all of this problem in lab today; if so, you are that much further along. This problem is to be handed in also.
Hints for Homework
Prepared by Brian Paris and Brenda Boerger, spring 2009

The following list includes general things that should be included or addressed in your homework problems, and for which points are deducted if they are not there. It cannot be inclusive since not every problem asks for the same thing. When in doubt, use the following guidelines.

1. Put phonetic data in square brackets:  [ælsəm]
2. Put phonemic data in diagonal slashes:  /tri/
3. Write English glosses in single quotes:  ‘coconut’
4. Always include the English glosses.
5. Always indicate the number of the form being analyzed from the homework sheet:
   Example: Seri, p. 114.    1. [nas] ‘milkweed pod’
6. As the course progresses, continue to provide a prose description of the data in addition to whatever rules you might write; the more we know about what you think is going on the better we can grade it, and we don’t give the benefit of the doubt.
7. Always tell us the implications of your analysis for the orthography: how many graphemes are needed in the writing system to represent the sounds you’ve analyzed? What graphemes might you suggest?
8. Be neat and systematic – lots of squiggly arrows and sideways writing get us lost.
9. When comparing sounds, be sure to state what the relationship is:
   CIE – contrast in identical environments
   CNE – contrast in non-influencing environments = [CAE contrast in analogous environments]
   CD – complementary distribution
   FV – free variation
10. To prove CIE or CNE provide at least one pair of words that show the change in sound causes a change in meaning. CIE is ideal, but CNE is acceptable; be sure to provide as much evidence as possible:  [tas] ‘to build fence’    /t/ ; /d/  CIE
    [das] ‘then, but’
11. To show CD write your observations in an informal rule such as:
   /t/ → [d] / V__V
   [t] / elsewhere
12. Binary features should be in square brackets with plus or minus:   [+voice]
13. Non-binary features should not be in square brackets; they should not use + or -:   palatal

A BIT LATER ON...RULE ORDERING
-Be sure to list all rules at the end of the assignment by placing them in a vertical list, crucial rule orderings at the top, rules that are unaffected at the bottom.
-When ordering rules be sure to connect rules that are crucially ordered with an arc.
-Be sure to prove all crucial rule orderings with two derivations, one with the correct ordering, one without. Do this for all crucially ordered pairs.
-Name all your rules.

MUCH LATER ON...FEATURE GEOMETRY TREES
-When drawing trees make sure the lines go directly above and/or below the node.
-It’s okay to abbreviate the names of the nodes, and the features, too.
### Classical Distinctive Features

**Day 5**

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>pʰ</th>
<th>p'</th>
<th>b</th>
<th>t</th>
<th>d</th>
<th>k</th>
<th>g</th>
<th>q</th>
<th>φ</th>
<th>β</th>
<th>f</th>
<th>v</th>
<th>θ</th>
<th>ó</th>
<th>s</th>
<th>z</th>
<th>ĵ</th>
<th>ʒ</th>
<th>x</th>
<th>y</th>
<th>tʰ</th>
<th>tʃ</th>
<th>dʒ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>syllabic</strong></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>consonantal</strong></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>sonorant</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>continuant</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>delayed release</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>strident</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>distributed</strong></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>lateral</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>anterior</strong></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>coronal</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>nasal</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>voice</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>aspirated</strong></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>glottal</strong></td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>high</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>low</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>back</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>round</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>ATR (tense)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

φ = p  β = b  δ = d  ĵ = š  ʒ = ž  γ = g  tʰ = c = c  tʃ = č = tʃ  dʒ = j = dʒ²  See Marlett, Appendix A for definitions
Classical Distinctive Features
Day 5

<table>
<thead>
<tr>
<th></th>
<th>m</th>
<th>m̃</th>
<th>n̂</th>
<th>ñ</th>
<th>l̃</th>
<th>l̄</th>
<th>r</th>
<th>j</th>
<th>w</th>
<th>?</th>
<th>h</th>
<th>i</th>
<th>í</th>
<th>e</th>
<th>é</th>
<th>æ</th>
<th>œ</th>
<th>a</th>
<th>i/u</th>
<th>u</th>
<th>u̯</th>
<th>o</th>
<th>ó</th>
<th>ē</th>
</tr>
</thead>
<tbody>
<tr>
<td>syllabic</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>consonantal</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>sonorant</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>±</td>
<td>±</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>continuant</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>±</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>delayed release</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>±</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>strident</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>distributed</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>lateral</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>anterior</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>coronal</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>nasal</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>voice</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>aspirated</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>glottal</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>±</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>high</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>low</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>back</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>round</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ATR (tense)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( r = ř \quad j = y \quad i = í \quad u = í \) Note that [i] is phonetically central and [ui] is phonetically back. See Marlett, Appendix A for definitions
Practical premises for phonological analysis

The following is a compilation of practical guidelines which you should find helpful as you do phonological analysis. This list is from the book *Phonemics*, written by Kenneth L. Pike, although the numbering has been changed from the original and footnotes were added for clarification.

1. A phonemic orthography is the easiest one for the native to learn to read and write.
2. Phonemic procedures are based upon universal language characteristics.
3. Borders of major phonological and grammatical units can cause the nonsignificant modifications of sound units.
4. Stress, pitch, and length can affect or be affected by sound segments.
5. Sounds tend to become more like the environments which modify them.
6. Tonal systems, as well as segment relationships, tend to be somewhat symmetrical.
7. Nonsignificant fluctuation of sound should not be written in a phonemic orthography.
8. Each language contains characteristic sequences of sounds.
9. Every language has consonants and vowels.
10. Certain kinds of segments may be vowels in one language but consonants in another, and vice versa.
11. The dichotomy between vowel and consonant is not strictly an articulatory one but is in part based on distributional characteristics.
12. Phonetic syllables are determined by physical and/or acoustic criteria.
13. Phonemic syllables are in part determined by distributional criteria, including potential placement of stress, pitch, and length, and in part by the structural shape of morphemes.

---

1 non-phonemic


© 1947 Kenneth L. Pike. Used by permission from copyright holder (Judith Pike Schram) for use by GIAL. Permission granted on April 7, 2011. (Pike.pdf)
14. A sequence of two segments may in some languages constitute a single phonetically complex phoneme.

15. Occasionally a single segment may constitute a consonant and a vowel simultaneously.

16. Some segments may be nonsignificant transition sounds.

17. Segmental or suprasegmental elements which are predictable are nonphonemic.

18. If two segments are submembers\(^2\) of a single phoneme, the norm\(^3\) of the phoneme is that submember which is least limited in distribution and least modified by its environments.

19. In order to be considered submembers of a single phoneme, two segments must be (a) phonetically similar and (b) mutually exclusive as to the environments in which they occur.

20. Every phonetically distinct segment of a language is a separate phoneme unless it is part of some more inclusive phonemic unit.

21. When two phonemic conclusions each appear to be justifiable by the other premises, and each seems to account for all the available facts of all types, that conclusion is assumed to be correct (a) which is the least complex, and (b) which gives to suspicious data an analysis parallel with analogous nonsuspicious data, and (c) which appears most plausible in terms of alleged slurs into specific environments.

22. Two segments are proved phonemically distinct if they consistently constitute the only difference between two words of different meanings.

23. The native speaker can more easily be taught to recognize and symbolize the difference between two of his phonemes than between two submembers\(^4\) of phonemes.

24. In some languages considerable grammatical analysis, based on phonetic data, is prerequisite to phonemic analysis since spaces and hyphens must be written at certain types of grammatical units, and subphonemic modifications may occur at their borders.

25. A PHONEME is one of the significant units of sound arrived at for a particular language by the analytical procedures developed from the basic premises previously presented.

\(^2\) allophones
\(^3\) underlying form
\(^4\) allophones

CHECK YOURSELF 1

1. True    False  The demonstration of contrast between two sounds establishes them as allophones of the same phoneme.

2. T     F  In a case of complementary distribution, the two sounds occur in mutually exclusive environments.

3. T     F  The words $[p^h\text{ats}]$ ($pots$) and $[p^h\text{ad}]$ ($pod$) constitute a minimal pair in English for the phones $[t]$ vs. $[d]$.

4. T     F  One segment at the phonetic level (one allophone) typically corresponds with several segments at the phonemic level (several phonemes).

5. Fill in the blanks with either the word “phoneme(s)” or “allophone(s)”:

Two sounds are _______________________ of one _______________________ if they occur in complementary distribution.

If two phonetically similar sounds occur in mutually exclusive environments, they are _______________________.

A(n) _______________________ has the potential to create a difference in meaning.

If two phonetically similar sounds contrast in minimal pairs, they are _______________________.


6. Circle the words in which the last sound is a(n) (don’t be confused by the spelling):

fricative: breathe lazy these physics

obstruent: much name wed care

nasal: corn lamb lamp sheep

7. Name two phonetic terms or features that each of the following natural classes of segments share in common:

[b d ɡ] _________________________ _________________________ consonants

[k ɡ x] _________________________ _________________________ consonants

8. When we study the phonological system of a particular language, what are the three main areas in which we are interested?
<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>[kïxšá?]</td>
<td>‘mouth’</td>
<td>[i] = [u], [ʃ] = [ʃ], [W] is a voiceless [w]</td>
</tr>
<tr>
<td>[kuWšúškaʔ]</td>
<td>‘dolphin’</td>
<td>[ʃ] is a voiceless retroflexed alveopalatal fricative</td>
</tr>
<tr>
<td>[nuWpóʃ]</td>
<td>‘snail’</td>
<td>[á] is a vowel bearing primary stress</td>
</tr>
<tr>
<td>[tahpín]</td>
<td>‘house’</td>
<td>[i] is a nasalized vowel</td>
</tr>
<tr>
<td>[ãhämpána]</td>
<td>‘rattlesnake’</td>
<td>[ã] is a vowel bearing secondary stress</td>
</tr>
<tr>
<td>[maht’óti]</td>
<td>‘broom’</td>
<td>[ɾ] = [r] = voiced alveolar trill</td>
</tr>
<tr>
<td>[wiçtáš]</td>
<td>‘shinbone’</td>
<td>[ç] = [tʃ] = [tʃ]</td>
</tr>
<tr>
<td>[piçšín]</td>
<td>‘sleeping mat’</td>
<td></td>
</tr>
<tr>
<td>[βíxtóŋko]</td>
<td>‘forehead’</td>
<td></td>
</tr>
<tr>
<td>[içsáʔ]</td>
<td>‘bird’</td>
<td></td>
</tr>
<tr>
<td>[rïxkín]</td>
<td>‘nose; snout’</td>
<td></td>
</tr>
<tr>
<td>[kahpí]</td>
<td>‘alligator’</td>
<td></td>
</tr>
<tr>
<td>[βuWčǒŋ]</td>
<td>‘wave (of water)’</td>
<td></td>
</tr>
<tr>
<td>[mahcãŋ]</td>
<td>‘evil spirit, demon’</td>
<td></td>
</tr>
<tr>
<td>[nahpúčaʔ]</td>
<td>‘papaya’</td>
<td></td>
</tr>
<tr>
<td>[buWšíʔ]</td>
<td>‘penis’</td>
<td></td>
</tr>
<tr>
<td>[kïxtśóŋ]</td>
<td>‘chigger’</td>
<td></td>
</tr>
<tr>
<td>[šiçtáti]</td>
<td>‘bridge’</td>
<td></td>
</tr>
</tbody>
</table>

Focus on [h], [W], [x], and [ç].

Sources of data:

PHONEMICS – ANALYTICAL PROCEDURES

1. Make phonetic chart of data
2. Identify suspicious pair(s)
3. Identify phonetic differences between segments
4. Examine environments in which each occurs
5. Define differences in environments
6. Make hypothesis as to environment in which each occurs
7. Check data to try to refute hypothesis
8. Modify hypothesis
9. Countereexample(s) found
10. Hypothesis can’t be refuted – no counterexamples

FREE VARIATION

one phoneme

Variation:
- Variation in every environment: unconditioned free variation
- Variation in only some environments: conditioned free variation

Counterexample(s) found:
- Hypothesis can’t be formulated (or modified) in a reasonable way
- Hypothesis can’t be formulated or modified (in a reasonable way)

No possible hypothesis to show predictability of occurrence of either member of the pair

CIE

CAE

CONTRAST – separate phonemes (phonetic differences are not predictable) – no rule can be written

COMPLEMENTARY DISTRIBUTION

one phoneme (phonetic differences are predictable, caused by environment)

Decide on underlying form of the phoneme

Write a rule (obligatory rule) to show predictability of occurrence of allophones
Several websites list the phonemic inventories of multiple languages:

UPSID (UCLA Phonological Segment Inventory Database), 451 languages
http://www.linguistics.ucla.edu/faciliti/sales/software.htm

P-base, 500+ languages
http://aix1.uottawa.ca/~jmielke/pbase/

PHOIBLE (Phonetics Information Base and Lexicon), 1672 languages
http://phoible.org/

A very interesting site for exploring minimal pairs in Thai:
http://www.thai-language.com/minimal-sets

**********

For researching syllable structure in particular (maximal templates, restrictions on consonant clusters), etc., these two sites are very helpful:

World Phonotactics Database, 2000+ languages
http://phonotactics.anu.edu.au/

LAPSyD (Lyon-Albuquerque Phonological Systems Database)
http://www.lapsyd.ddl.ish-lyon.cnrs.fr/lapsyd/

**********

The personal homepage of Bruce Hayes, a very prominent phonologist, has lots of useful links, tools, and applications:
http://www.linguistics.ucla.edu/people/hayes/

**********

At the phonetic level, Chamicuro has three different lateral consonants:

[l]   a voiced alveolar lateral, just like the English [l] in the word *liquid*

dl] a voiced alveolar lateral affricate (a single segment, not really a sequential combination of [d] plus [l]), which sounds similar to the *dl* in an English place name such as *Midlothian*

[l] a voiceless alveolar lateral fricative, similar to the *l* in the English word *play*

Other unusual segments are the following: [§] is a voiceless, retroflexed alveopalatal fricative. Similarly, [ɕ] represents a voiceless, retroflexed alveopalatal affricate. [ʂ] is a voiceless lamino-alveolar fricative (like [s] except articulated with the blade of the tongue rather than the tip or apex). [y] is a voiced palatal glide (IPA [j]). [ɲ] is an (alveo)palatal nasal (IPA [n]). [tʃ] is a voiceless alveolar affricate (IPA [ts]). A colon (:) after a vowel indicates length. An acute accent (´) over a vowel indicates primary stress. A period (.) is used to mark the division between syllables.

**examples of [dl]**

1. [i.ša.ka.tis.ká.dle] ‘abandoned’
2. [ih.pó.dla] ‘mud’
3. [ah.ke.dlóh.ki] ‘heart’
4. [ah.kó.dlo] ‘scorpion’
5. [a.dláʔ.ye] ‘that one’
6. [a.dlij.kwaʔ.ta.kó.či] ‘fever’
7. [a.ñá.dlo] ‘ant’
8. [as.kó.dli] ‘heron’
9. [aš.mu.dleʔ.kó.dlo] ‘club’ (noun)
10. [čoʔ.ka.pé.dli] ‘fishhook’
11. [eʔ.te.šú.dli] ‘cough; cold’
12. [i.ču.dlú.pa] ‘young lady’
13. [ka.še.dleʔ.tá.ka] ‘dead’
14. [ma.dlú.ti] ‘widower’
15. [maʔ.ko.dlé.ti] ‘thief’
16. [pnah.mú.dle] ‘swamp’
17. [pú.ño] ‘lightning bug’
18. [u.dlá:.ti] ‘I laugh’

examples of [l]

19. [ah.tá.kli] ‘flute’
20. [aʔ.lo.t'iʔ.ta] ‘near’
21. [cé.klo] ‘parakeet’
22. [ih.lá.pi] ‘bank of a river’
23. [i.šo.plá.ma] ‘door’
24. [i.táʔ.ła] ‘skirt’
25. [kwéʔ.ła] ‘sparrow hawk’
26. [la.čéʔ.ta] ‘last night’
27. [li.kah.péʔ.ta] ‘yesterday’
28. [lo.kó.ko] ‘grandfather’
29. [ma.si.klo] ‘capybara rodent’
30. [maʔ.lu.šá.na] ‘termite’
31. [me.plo.né.ye] ‘boy’
32. [óš.lo] ‘wild pig’
33. [piš.le] ‘wing’
34. [piš.ło] ‘hummingbird’
35. [plá.wa] ‘long’
36. [so.ko.plé.ye] ‘snake’
37. [syek.púč.le] ‘beer belly’
38. [yiš.ła] ‘island’
39. [ši.po.to.š.kaʔ.šoʔ.ło.či] ‘shotgun’
40. [šam.le.čó.ma] ‘coward’
41. [toʔ.ło.ki] ‘species of palm tree’
42. [túʔ.lu] ‘chest’
43. [klo.čé.to] ‘kidney’
44. [ko.neʔ.lá.wa] ‘younger sister’
45. [lu.šá.na] ‘godfather’
46. [mút.le] ‘daughter’
47. [u.paʔ.lís.ki] ‘I send, order’
48. [ya.pléʔ.ti] ‘lightning’

examples of [l]

49. [yél.na] ‘man; husband’
50. [čé.ti] ‘ripe’
51. [ka.waš.ta.kuž.yá.ka] ‘he who kills’
52. [kiš.ko] ‘three’
53. [i.šőt.ti] ‘he or she throws’
54. [šiš.tó.ki] ‘drum (noun)’
55. [u.čé.t’a.ka.máʔ.ti] ‘I shout’
Instructions for the written homework to turn in on Day 8:

Analyze the three lateral consonants [l], [ł], and [dl]. What type of relationship are they in? That is, do they contrast, or are they in complementary distribution? Give your conclusions and evidence that supports them, as well as any implications for the phonology and/or orthography of the language. State in prose any generalizations you can predict concerning the occurrence of these lateral phones. In other words, state the relevant phonetic environments in which each of these lateral phones appears in Chamicuro, and then try to formalize the rules also, using the type of notation you have been learning about in class. If you are positing just one lateral phoneme, which of its allophones would you choose as the underlying form to represent the phoneme as a whole, and why? For the sake of this problem, you may assume that all other segments are phonemic, although this is oversimplifying the facts somewhat. Also, state the generalization describing the occurrence of stress. In other words, on which syllable of every word does stress always appear? Finally, what can you say about the last sound in every word?

Once you have gone through all of these analytical steps, you should then be able to transcribe (rewrite) Chamicuro words in their phonemic representation (abstract, underlying form). Do this for words # 2, 28, 48, and 53. Assume that all segments (sounds) except for the laterals are phonemic, so don’t change any of them. Furthermore, do not include any syllable boundaries (.) in your phonemic transcriptions. If all of your analysis is correct, your generalizations should correctly account for the two different laterals in the Chamicuro word [śa.çúł.lo], which means ‘young man; youth’. Do they? The last step in this assignment is to make a phonemic inventory (list) of all of the consonant phonemes of Chamicuro. Do not worry about doing this for the vowels. Arrange your consonant inventory into a chart or diagram, placing natural classes of sounds in the same columns and rows (e.g., alveolars, nasals, etc.). In doing this step you should follow the same general type of arrangement as other consonant charts of this kind that you may have seen. For an example, see Appendix D of Marlett on p. 251. However, for this assignment you do not have to put any features/labels on your chart of consonant phonemes.

Sources of these data:


© 2011 Steve Parker. Used with the permission of the author for distribution by GIAL from 2011-2015. (Chamicuro HW.pdf)
Homework for Day 9:
This is listed and explained on one of today’s handouts in the course packet.

Day 8: The big picture
In today’s class we are going to explain how we know which specific pairs of sounds to compare with each other, in order to analyze their phonological relationship. We will also look at a type of phonemic contrast which often exists between phonetically similar segments in near-minimal pairs of words.

Jakobson:
“The dichotomous scale is superimposed by language upon the sound matter.”

- Also known as suspicious pairs or phonetically similar segments (PSS).
- These consist of segments which are very similar, such that they differ phonetically by only 1-3 features (usually).
- We can identify different degrees of phonetic similarity.
- The more similar two sounds are, the greater the likelihood that they might be found to be allophones of the same phoneme, that is, in complementary distribution with each other.
- In order for two sounds to be considered allophones of the same phoneme, they not only need to be in complementary distribution, but they also must be phonetically similar.
- This is why we do not consider [h] and [ŋ] in English to be allophones of the same phoneme, even though they are in complementary distribution.

- Any modification of a basic sound unit is suspicious with that sound, e.g., [p] with [ph], [pw], [pː], etc. And all of these with each other, too.
- Among the obstruents, we can identify three main areas or places of articulation: labial, coronal, and dorsal (define).
- These will almost always be phonemically distinct, so we do not need to compare, for example, [p] and [t].
- This chart is based on the experience of phonologists in a great many of the world’s languages.
- Affricates are suspicious with both sounds which make them up, e.g., [ʃ] with both [t] and [ʃ] ([ʃ]).
- Similarly, any complex segment composed of two basic parts should be compared with both of these individually.
- All nasals are suspicious with each other, so the three-way place distinction among obstruents does not extend to sonorants.
- All liquids are likewise suspicious with each other, regardless of their point of articulation.

- Among the ___________________________, each one is suspicious with all of the other vowel segments which are immediately adjacent to it on the chart.
- Note that in rare cases virtually any vowel may be an allophone of any other one. For example, in English schwa occurs only in unstressed syllables, where it is an allophone of all of the vowel phonemes of the language.
- However, [i], [a], and [u] are almost always phonemically distinct, as are [ɛ] and [o].
- A lengthened vowel such as [oː] is automatically suspicious with every other segment that its short counterpart is suspicious with.

Contrast in Environment (CAE)
- CD and CIE are not the only possible relationships which two sounds can be in; can you think of others?
- Principle: two sounds occur in similar environments, and there’s no way to envision the phonetic difference between the two segments as being conditioned by what (few) differences there are in the environment. For instance, [sːaː] ‘skin’ vs. [ʃaː] ‘to, at, in’ in Chamicuro.
- Contrast this notion with that of CD: the differences in the environment are directly responsible for the phonetic difference between the two segments, i.e., the difference in realization of the two allophones of the phoneme.
- This presupposes that we are always looking for possible processes that could account for such phonetic differences, especially cases of assimilation.
- This concept (CAE) is also known as Contrast in Noninfluencing Environment (CNE) in the textbook.
Which pairs of PSS’s need to be checked out in this language? These should include at least the following ones:

- p : b     t : d     k :
- j : h    m : n :

and secondarily,

- t : s     s : h     b : m     d : n

For the vowels, we would want to check out the following pairs:

- i : e     u : a     e : a

and secondarily,

- i : u
Your homework for Day 9 consists of the following:

1. Read Marlett, chapters 21, 22, 24, and page 136.

2. Review (reread) the “Phonemics” handout, section 1 only.

3. Read the discussion of the Chamicuro problem which you did for today (on another handout).

4. Do the exercise on p. 135 of Marlett, for practice only (not to be collected).

5. Work through the Kuskokwim Eskimo exercise in Marlett, p. 115. Instructions: what type of relationship exists in this language between the segments [k] and [q]? Explain and discuss your answer and its implications. If you observe a phonological process taking place, express the generalization (rule) in prose only. This exercise is to be handed in.

6. Work through the Tlahuitoltepec Mixe problem in Marlett, p. 128. Analyze the relationship between the segments [o] and [ø]. [ø] is a mid, front, rounded vowel and can thus be characterized by the features [–back], [–high], [–low], [+ATR], and [+round]. Follow the same procedure as for the previous exercise (#5 above). Note: the palatal glide [j] has the same features as the vowel [i], except that it is a consonant rather than a vowel. Thus, [j] can be characterized as [+high], ([–low]), [–back], [+ATR], and [–round]. Try to comment on how and why these feature specifications of [j] are important in this set of data. If you observe a phonological rule, you can state it in prose only. Don’t worry about trying to formalize it, unless you really want to make an attempt to do so. This exercise will also be collected on Day 9.
Phonetically Similar Segments

Within the larger circles of segments, specific pairs must be investigated.


Your homework for Day 10 consists of the following three exercises:

1. **Venda (p. 116 in the Marlett textbook)**

   Analyze the pair of suspect consonants [n] vs. [n̪]. What type of relationship are they in? If they contrast, give your evidence. If they are in complementary distribution, discuss which one should be chosen as the underlying form of the phoneme and why, and give a prose rule describing the phonological process which is taking place. Also, discuss the implications of your analysis. This problem is for practice only, **not** to hand in.

2. **Huariapana**

   In a previous class (Day 6) you were given a handout containing some Huariapano data. Analyze the relationship between the four segments [h], [W], [x], and [ç], following the same general directions given above for Venda. This problem **is** to turn it.

3. **Tojolobal (p. 118)**

   Analyze the pairs of sounds indicated in the book. An apostrophe after a consonant indicates that it is an ejective (glottalized). For each pair of segments, follow the same general directions given above. This problem will also be collected. Important note: there appears to be a typo in the data given in the book on p. 118. In word #9 (armadillo), the [p] is supposed to be an ejective: [p’]. At least that is what the solution on p. 120 seems to imply.
Data for suspect pairs exercises

Eskimo (Cape Prince of Wales, Alaska), chart of phonetic consonants:

<table>
<thead>
<tr>
<th>bilabial</th>
<th>labiodental</th>
<th>dental</th>
<th>alveolar</th>
<th>palatalized</th>
<th>palatal</th>
<th>velar</th>
<th>uvular</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>vl stop</td>
<td>p</td>
<td>̃</td>
<td>t</td>
<td>d</td>
<td>ḳ</td>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vd stop</td>
<td>b</td>
<td></td>
<td></td>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vl fric</td>
<td>f</td>
<td>s</td>
<td>s</td>
<td>ç</td>
<td>x</td>
<td>χ</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>vd fric</td>
<td>v</td>
<td>z</td>
<td></td>
<td>j</td>
<td>γ</td>
<td>ḳ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n̄</td>
<td></td>
<td>n̄</td>
<td>n̄</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td></td>
<td></td>
<td></td>
<td>dl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vl lateral</td>
<td></td>
<td>̃</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td></td>
<td>̃</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fore (Papua New Guinea), chart of phonetic vowels:

\[
\begin{align*}
i & \quad u \\
\i & \\
e & \quad e : \quad o \\
\epsilon & \quad a : \\
\Lambda & \\
\end{align*}
\]

([e:], [a:], and [o:] are [+long])

Carib labial consonants (Guatemala):

1. [bani] ‘your thing’ 5. [uβi] ‘meat’
3. [βidi] ‘eight’ 7. [βelu] ‘well’
4. [wibu] ‘hill’ 8. [wani] ‘one thing’

The homework for Day 11 is listed and explained on one of today’s handouts in the course packet.

Day 10: The big picture

In today’s class we are going to define, describe, and illustrate the third major type of phonological relationship which can exist between two phonetically-similar speech sounds: free variation.

Huariapano (Peru)

[josi] ‘old’
[jusi] ‘old’

Free (FV)

I. The principle:
- Two sounds occur in identical environments but do not make a difference in (referential) meaning; cf. the notion of variation as we originally introduced it – non-phonemic differences in sounds that native speakers may not pay attention to.
- Every occurrence of a “sound unit” (phoneme) will be somewhat different phonetically, even if the difference is very slight.
- There is always an acceptable range of pronunciation of each phoneme that still “counts” as an instance of that segment (e.g., scattergram of tokens of a particular vowel on a chart).

II. Pike’s Third Premise for Phonological Analysis: Sounds tend to fluctuate.

A. __________________________ free variation: the variation occurs in any and every environment.

- For example, in Huariapano, /o/ → [u] very freely, as in the word /josi/ ‘old’.
- Some sociolinguists question whether this is ever really possible; there tend to be statistical patterns at least in the occurrence of some allophones. These may be sensitive to conditioning factors that are extralinguistic in nature, such as social class, gender, geographical origin, etc.

B. __________________________ free variation, which is the most common case.
- For example, /l/ → [dl] / V __ V in Chamicuro.
- Here the variation occurs only in certain phonological contexts, usually where you can see a conditioning from something in the environment.
- This type of free variation resembles cases of complementary distribution very closely; in fact, cases of CD often arise historically from an earlier stage in the language when there was conditioned free variation. Later the optionality of the process becomes obligatory, yielding a clear pattern of CD.
• A logical and obvious corollary to these principles is that any two segments which are in a relationship of free variation in a particular language should automatically be counted as a suspicious pair, no matter how many features they may differ in.

II. How to deal with this phenomenon in terms of the phonological analysis
• Since there is a regular pattern of occurrence, we should follow through and write a rule to describe or capture what is going on.

• However, the difference is that in a case of FV, a rule is____________. In other words, regardless of whether or not the rule applies (takes effect), the output (actually pronounced phonetic form) is always acceptable.
• This contrasts with a case of complementary distribution in the sense that in terms of CD, a failure to apply the rule would yield something that simply is never said, as long as it is a case of complete CD.

• Some of the most important suspicious pairs are the following:

  | [t] vs. [tʰ] |
  | [p] vs. [b] |
  | [t] vs. [d] |
  | [k] vs. [ɡ] |
  | [t] vs. [ts] |
  | [s] vs. [ʃ] |
  | [n] vs. [ŋ] |
  | [w] vs. [β] |
  | [i] vs. [ɪ] |
  | [e] vs. [ɛ] |
  | [a] vs. [ə] |

Pajonal Campa (Peru)

tʰ  i  i
p  t  k  e  ə  o
ts  e
b  d  ɡ  a
s  ʃ  h
β
m  n  ŋ
r
w

A. Choose an underlying form for the phoneme. This may be somewhat problematic in the case of unconditioned free variation since neither allophone will be less restricted in distribution. In this situation we would tend to choose as the underlying form that allophone which is more natural universally (in the languages of the world).

B. Write a______________ for it

For unconditioned free variation, the rule would not have any environment. In other words, it occurs in any and all contexts, such as /o/ → [u] in Huariapano.

For conditioned free variation, the rule would be context sensitive: it applies only in a specific environment. In this case it would be no different from a rule written to describe a case of (obligatory) complementary distribution, except that the rule would now be optional. An example of this would be /l/ → [dl] / V __ V in Chamicuro.
AL 4303, Principles of Phonological Analysis
GIAL
Day 10

Your homework for Day 11 consists of the following:

1. read the Phonemics handout, part 2 (in the course packet)
2. read Marlett, chapters 15, 23, and p. 109
3. check yourself with Marlett, pp. 136-37, for practice only. The answers are on p. 137. Most of these questions should be a review for you.
4. Analyze the Corongo Quechua data below (see p. 127), to be handed in. Focus on the pairs of vowels [i] vs. [e] and [u] vs. [o]. What two types of relationships are they in? Discuss and explain your answer, giving prose rules wherever appropriate. How many vowel phonemes are there in this set of data? Do not worry about expressing any rules formally. Note that the place of articulation features corresponding to uvular consonants (such as [χ]) can be uniquely defined as [+back] and [–high]. How might this fact help to explain the two phonological processes you observe operating in this Peruvian language? [χ] is a voiceless uvular fricative. [ɾ] is a voiced alveolar tap or flap. [tʃ] is a voiceless retroflexed alveopalatal affricate.

2. [purin] ‘walks’  9. [χoskin] ‘he gives it all’
3. [χεʃwa] ‘nest’  10. [isχuŋ] ~ [isχon] ‘nine’
4. [ura] ‘below’  11. [janaxi] ~ [janaxe] ‘companion’
5. [uma] ‘head’  12. [roχung] ~ [roχon] ‘it cuts’
7. [ara] ‘corn’


### Pajonal Campa (Peru)

1. [tʰampi] ~ [tʰambip] ‘butterfly’  13. [ˈe̞to] ‘bird species’
3. [aˈkɛndiri] ~ [aˈkɛntiri] ‘we chase’  15. [ˈi̞tɔki] ‘on his head’
4. [ˈʃiŋki] ~ [ˈʃiŋgi] ‘corn’  16. [ˈeiɾo] ‘no, not’
6. [ʃiβtʰɔ] ~ [ʃiβtʰa] ‘vine, rope’  18. [ˈtsiβi] ‘salt’
7. [ˈʃiəwɔ] ~ [ˈʃiwa] ‘fish’  19. [nɔkihoˈtatsi] ‘I doubt’
9. [piˈʃindo] ~ [piˈʃinto] ‘your daughter’  21. [ˈkatstɔri] ‘ant’
10. [tiˈkatɔsi] ‘nothing’  22. [ʃɔwo] ‘rodent species’
11. [oˈɛmbiki] ~ [oˈɛmpiki] ‘dead limb’  23. [noˈmare] ‘my paddle’
12. [kaˈero] ‘termite’

Language Data © 1974 and 1990 SIL. Used by permission from SIL (granted by Gayle Sheehan on July 7, 2011) for use in GIAL’s AL4303 Principles of Phonological Analysis course by up to 100 students each year.

All else © 2011 Steve Parker. Used with the permission of the author for distribution by GIAL from 2011-2015.

## Phonetic Chart of Consonants

<table>
<thead>
<tr>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Interdental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Retroflexed Alveolar</th>
<th>Alveopalatal</th>
<th>Retroflexed Alveopalatal</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricatives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laterals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flaps / Taps</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glides (Approximants)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implosives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clicks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Front</td>
<td>Central</td>
<td>Back</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>------------------</td>
<td>------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>close / tense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>open / lax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>close / tense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>open / lax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>close / tense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>open / lax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## PHONEMIC CHART OF CONSONANTS

<table>
<thead>
<tr>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Interdental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Retroflexed Alveolar</th>
<th>Alveopalatal</th>
<th>Retroflexed Alveopalatal</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stops</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Affricates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nasals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Laterals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Trills</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Flaps / Taps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Glides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(Approximants)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Implosives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clicks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Front</td>
<td>Central</td>
<td>Back</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------------</td>
<td>-----------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unrounded</td>
<td>Unrounded</td>
<td>Unrounded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rounded</td>
<td>Rounded</td>
<td>Rounded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>close / tense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>open / lax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>close / tense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>open / lax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>close / tense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>open / lax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Recall the basic pattern of occurrence of the suspicious pair of sounds \([s]\) and \([ʃ]\) in Japanese: \([ʃ]\) appears always and only before the vowel \([i]\), and \([s]\) occurs everywhere else. Another way in which we might summarize these facts is by writing distribution statements describing in detail the environmental restrictions on the two segments:

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Allophone</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>/s/</td>
<td>[ʃ]</td>
<td>occurs before [i]</td>
</tr>
<tr>
<td></td>
<td>[s]</td>
<td>occurs elsewhere</td>
</tr>
</tbody>
</table>

Alternatively, we could account for these patterns by means of a phonological rule, which would be formalized as follows:

\[
\begin{align*}
/s/ & \rightarrow \text{[alveopalatal]} / \_ \_ V \\
& \quad \quad \quad \quad \quad \quad +\text{high} \\
& \quad \quad \quad \quad \quad \quad -\text{back}
\end{align*}
\]

A formal phonological rule consists of the following basic parts:

<table>
<thead>
<tr>
<th>A</th>
<th>→</th>
<th>B</th>
<th>/</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>input or target or focus</td>
<td>output or structural change</td>
<td>environment or trigger</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“becomes” change made “when” or “in the environment of” surrounding segments, etc. that are relevant in determining the change (a blank line indicates the location of the change)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The input (A) and the environment (C) together are referred to as the *structural description* (SD) of the rule. Here are some examples, along with how they are interpreted:

- \(X \rightarrow Y / \_ \_ Z\) reads “\(X\) becomes \(Y\) when it occurs immediately before \(Z\)”
- \(X \rightarrow Y / Z \_ \_\) reads “\(X\) becomes \(Y\) when it occurs immediately after \(Z\)”
- \(X \rightarrow Y / W \_ \_ Z\) reads “\(X\) becomes \(Y\) when it occurs immediately after \(W\) and immediately before \(Z\)”
An example of an actual rule can be seen below in three different forms: (a) informal, with no distinctive features; (b) more formal; and (c) completely formal (with all distinctive features):

(a) /n/ → [ŋ] / ___ g (The phoneme /n/ becomes [ŋ] when it occurs before [g].)

(b) /n/ → [velar] / ___ C

(c) C → [velar] / ___ C

Of the three formalisms above, (b) and (c) are preferable to (a) because they more clearly (explicitly) show the phonological motivation for this process (change).

Note: when writing formal phonological rules, always check them twice. Don’t leave out the “blank” in the environment, and check carefully to see if it is in the right place (before or after what it is next to, etc.). This is crucial in order for the rule to work correctly.

We will use the solution involving rules because, as will become apparent, rules such as these which supply phonetic detail may interact with and therefore need to be ordered with each other, a situation that is not easily handled by distribution statements.

Let’s take one more example: a uvular [q] in Seri does not contrast with the velar stop [k]. The uvular variant occurs only before or after a uvular fricative:

[stakx] ‘rock oyster’  [qɔaps] ‘harden’
[kokmo:nxk] ‘be zigzagged’  [sliːtyqox] ‘skunk’

When the conditioning environment that triggers the process appears on either side of the affected segment, this is called a “mirror-image” rule. These can be formalized by replacing the “/” introducing the environment with a percent sign (%). The blank showing the location of the change is typically omitted:

k → q % χ

2.2.2 Influence by boundaries. It has been seen that sounds tend to influence other sounds. It is also true that border points of grammatical or phonological units – especially word, syllable, or utterance – may cause the modification of sounds.

Voiceless stops in Texmelucan Zapotec are aspirated word-finally:

[bitʰ] ‘skunk’  [bic:] ‘rabbit’
[citʰ] ‘woodpecker species’  [kun’] ‘turkey’
[slekʰ] ‘sparrow hawk’  [tis:] ‘lame’
The following rule is therefore part of the phonology of Texmelucan Zapotec:

\[ C \rightarrow [+\text{aspirated}] / ___ # \]

Another common conditioning environment is utterance-final position. Consider the following data from Chimalapa Zoque, in which some phonetic details have been omitted:

- [kamaʔM] ‘hard’
- [kamaʔm te kuj] ‘the tree is hard’
- [waniʔN] ‘tiny’
- [waniʔn te kuj] ‘that bean is tiny’
- [hiʔD] ‘a plain’
- [ka hiʔŋ pic wihi] ‘that plain is very pretty’
- [waʔJ] ‘hair’
- [teʔj waʔj pic piti] ‘his hair is very long’

Notice that there is an alternation between voiced and voiceless sonorant consonants. Voiceless sonorant consonants occur only utterance-finally after a glottal stop. The conditioning environment is not the end of the word but the end of the utterance (as well as the preceding glottal stop). The rule might be states as follows, where the symbol “||” refers to the boundary of an utterance:

\[ C \rightarrow [-\text{voice}] / ? ___ || \]

\[ [+\text{sonorant}] \]

2.3 Contrast in analogous environments (CAE)

It is not usually the case that minimal pairs can be readily found for every suspicious pair of sounds to demonstrate that those segments contrast in identical environments. But the members of the pair may not occur in complementary environments, either. There may be no reason to imagine that the sounds are being influenced by surrounding segments or boundaries. In such a case the pair of segments may still be in contrast, even though we have no minimal pair to demonstrate it. What is necessary then is to look for words that are as phonologically similar as possible, with environments in which the differences would probably not account for the differences in the sounds (i.e., non-conditioning environments). We thereby demonstrate that the segments do contrast; several examples of such contrast in analogous environments are desirable.

To be sure that no possible conditioning factor has been overlooked, we try to find examples of this type of contrast in as many different types of environments as possible, e.g., initial, medial, and final position within the word.

Notice the following data from English:
We are concerned here with the segments [f] and [v], which differ only in voicing. They do not contrast in identical environments in the data above (for English, of course, we could also find minimal pairs for these segments, after collecting more data). We can make Y-diagrams to see what environments they occur in:

\[
\begin{array}{c|c|c}
[f] & # & [v] \\
\hline
\text{#} & \text{ijt} & \text{#} \\
\text{æ} & \text{#} & \text{æ} \\
\text{dæ} & \text{æn} & \text{i} \\
\end{array}
\]

The phonological differences between these environments, such as the height of the vowels, are not likely to account for the voicing difference between [f] and [v]. We can make no reasonable hypothesis to explain or predict the difference between these two segments, based on conditioning of the environment. (There are cases, however, where the conditioning environment is somewhat surprising.) We would conclude then that [f] and [v] are separate phonemes in English because they contrast in analogous environments (environments that are almost the same and which would not be a conditioning factor). We might later find minimal pairs to confirm this analysis.

The following table points out some of the most important differences between the principles of contrast and complementary distribution, now that we have discussed both of them:

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Complementary distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>In identical environments (CIE)</td>
<td>Environments are mutually exclusive:</td>
</tr>
<tr>
<td>Minimal pairs</td>
<td>where one sound can occur, the other never will.</td>
</tr>
<tr>
<td>In analogous environments (CAE)</td>
<td>Occurrence is predictable.</td>
</tr>
<tr>
<td>Analogous pairs</td>
<td>The sounds are non-distinctive.</td>
</tr>
<tr>
<td></td>
<td>The sound difference makes no meaning difference, just a pronunciation change (an “accent”).</td>
</tr>
<tr>
<td>Occurrence is not predictable.</td>
<td>Write a rule!</td>
</tr>
<tr>
<td>The sounds are distinctive.</td>
<td>The underlying form is the one which is least predictable, least restricted.</td>
</tr>
<tr>
<td>The sound difference potentially makes a difference in meaning.</td>
<td></td>
</tr>
</tbody>
</table>
Conclusion: the sounds are separate phonemes.

Ex. English /s/ɪp - /z/ɪp
/p/ət - /b/ət

Conclusion: the sounds are allophones of one phoneme – surface (phonetic) variants of the phoneme.

Ex. English [pʰ]ɪl - s[p]ɪl

2.4 Free variation

Sometimes one can notice variations in the way certain words are pronounced. For example, the following words in Nabak might be transcribed in the following ways:

[sali] ~ [sarɪ] ‘owl’
[tʰalak] ~ [tʰarak] ‘naked’
[lum] ‘conch shell’

Since the segments [l] and [ɾ] do not contrast elsewhere in the language, and since they do not distinguish utterances of different meanings here, even though they occur in the same environments, this phenomenon is called free variation between allophones of one phoneme. (It may be argued that there is no such thing as truly “free” variation, and that the alternation between allophones is determined by some broader linguistic or nonlinguistic context (in cases of emphasis or anxiety; depending on the speaker or the audience; etc.). However, this is the term usually used to describe this type of variation.)

We could use the following optional rule to describe this variation:

\[ l \rightarrow r / V \_\_\_ V \]

Speakers may choose to apply this rule, or they may not. (When writing rules to account for free variation, don’t forget to put the word optional in parentheses at the beginning. If this is omitted, the rule is an obligatory one which implies a case of complementary distribution.) As a result, the word for ‘owl’ will be written phonemically as /sali/. In this case, the /l/ was chosen to represent the phoneme, since [l] has a wider distribution than [ɾ]: in [lum] it occurs in initial position, where [ɾ] can apparently not appear in this language.

The segments [l] and [ɾ] are conditioned free variants of the phoneme /l/ – the free variation does not occur everywhere, but only in the specific environment between vowels. In other cases, the variation may occur in all environments; the segments would then be unconditioned free variants, and the rule that reflects such a case would have no environment. A language which had unconditioned free variation between [l] and [ɾ] would be described by a rule such as:

\[ \text{(optional)} \ l \rightarrow r \]

It might be that a variant is observed only in the dialect spoken by a certain group (geographically, socially, etc.). If that were true for the Nabak data, that fact should be indicated:
One final note should be made here. When variation occurs between segments that contrast in other environments, the situation is somewhat different, and we do not have a case of free variation. Some speakers of English pronounce ‘roof’ as [ᵻuf] while others pronounce it as [ᵻʊf]. This is not evidence that [ᵻ] and [ᵻʊ] are allophones of one phoneme since there are many examples like [ᵻut] ‘suit’ and [ᵻʊt] ‘soot’ which establish contrast between these two vowels. Variation between [ᵻ] and [ᵻʊ] in English is limited to a few words and illustrates alternation in the pronunciation of those specific words at the phonemic level. This idiosyncratic behavior of these particular lexical items must be indicated as such in the lexicon, where ‘roof’, for example, may be shown with two phonemic representations: /ᵻuf/ or /ᵻʊf/.

2.5 Summary

The three principles of contrast, complementary distribution, and free variation form the backbone of phonemic analysis. In determining which of these relationships two sounds bear to each other, we pay close attention to the distribution of each one – that is, the environments in which it occurs. Since we define the phoneme as the minimal unit of sound capable of making a difference in meaning, we also pay attention to meaning in our analysis.

The accompanying flow chart sketches out the procedure of phonemic analysis, showing the steps leading to each of the conclusions about pairs of sounds.

3. Transitional sounds

Finally, one must examine whether certain phones contrast with “null”. That is, can the presence of certain phones be completely predicted and not have to be included in the lexical (underlying) representation at all? These phones can include [j], as a transition after a high front vowel or before a palatalized segment; [w], as a transition after a round vowel or before a labialized segment; [ʔ], as a transition between silence and a vowel or between identical vowels; and [ə], as a transition between consonants.

A careful transcription of English reveals a glottal stop before an utterance-initial vowel:

[ʔæt] ‘at’     [ʔɪn] ‘in’
[ʔæp] ‘up’     [ʔawt] ‘out’

Since this glottal stop is completely predictable, it is not considered a phoneme and is not included in lexical representations of English. The following statement is possible:

A glottal stop occurs before a vowel at the beginning of an utterance.

4. Pattern symmetry

Experience in working with many languages has demonstrated clearly that sound systems have a tendency toward symmetry. A related generalization that can be made is that similar sounds tend
to behave similarly. These observations can be quite useful when one begins work in a new language, for they can lead the analyst to important discoveries, and speed the analysis along.

After collecting a limited amount of data from Ocotepec Mixtec, for example, an analyst might construct the following phonetic chart:

\[
\begin{array}{ccccccc}
p & t & k & \acute{q} & \acute{i} \\
b & g & e & o \\
\beta & \acute{\delta} & a \\
m & n & \eta \\
\end{array}
\]

We notice first that there are some “holes” in this chart. Knowing that sound systems have a tendency toward symmetry helps direct further investigation. Maybe gathering more data will fill in some of the gaps in the symmetry. Or, perhaps there are real holes in the pattern. (Every sound system seems to have at least some symmetry and some lack of symmetry, so we must not jump hastily to conclusions here.) For example, we might expect to find a [d] to go with the [b] and [g] to form a voiced stop series. Mixtec does in fact have a [d], as we would soon discover, as in the word [ŋɡandii] ‘sun’. We might also expect to find a [ŋ] as in [ŋɡandii] ‘sun’. But as it turns out, Mixtec does not have [ŋ] – an unexpected gap such as this may prove to be of great interest later. Likewise, a sound that seems quite out of place should be carefully examined. (It would be desirable at this point to also note (1) the frequency of the occurrences of the various segments and (2) the positions in which they occur. A segment that occurred only in word-final position would deserve further investigation, for example.) What other sounds might we predict to occur in Mixtec that are not attested in the chart above?

A complete chart of the phonetic consonants of Mixtec shows the following:

\[
\begin{array}{ccccccccccc}
p & t & t\acute{f} & k & k' \\
b & d & d\acute{ʒ} & g \\
\beta & \acute{\delta} & z & \grave{z} & x \\
m & n & \grave{n} & \eta & \grave{\eta} \\
\end{array}
\]

It can be demonstrated that [p] and [b] are in complementary distribution in Ocotepec Mixtec; similarly, [t] and [d] occur in complementary distribution. Since we know that similar sounds tend to behave similarly, we are now led to expect complementation between the pairs [t\acute{f}]-[d\acute{ʒ}] and [k]-[g]. (As a matter of fact, this does turn out to be true for Ocotepec Mixtec.) Even if we have but scanty data to show anything definite about these pairs, we can make the prediction that they will behave like the other pairs of oral stops, based on the pattern evidence. This sort of reasoning helps us to know what to expect from further data, and provides an intelligent way to proceed with further investigation.
Finally, we point out that the phonological system of a language may have a symmetry that is not true of the raw phonetic inventory. For example, the English obstruent system can be charted phonetically in this way:

```
     p   t   k   ?
  b   d   g
   tʃ
  dʒ
   f θ s ʃ h
  v ŋ z ʒ
```

There are a number of asymmetries in this chart which can be regularized when we consider the phonological system of the language as a whole. By grouping together the labials (bilabials and labiodentals), we allow [p], [b], [f], and [v] to all line up. Then, we fill in a gap in the stop series by treating the affricates as part of that set:

```
        p   t   tʃ k
    b   d   dʒ g
   f θ s ʃ h
  v ŋ z ʒ
```

These adjustments allow a phonological chart of the language to reflect even more symmetry, and this adjusted system may be useful and relevant to further analysis.

5. Perspective

The steps sketched in this chapter are ones which a linguist might take to begin to unravel the phonological system of a language. One seeks to determine (1) the sounds which comprise the lexical (or phonemic) alphabet of the language – the building blocks from which words are made in that language, and (2) the variations in those sounds which form part of the sound patterns of that language. We say that a language has sounds x, y, z, etc., as phonemes, and has a set of rules which specify how those sounds are actually realized phonetically in their various contexts.

We have tried to be careful in our presentation above not to give the impression that these mechanical procedures yield unambiguous results every time. Phonological analysis is not something which can be carried out in such a fashion. As you become more informed of what happens in languages (by reading the Marlett textbook, for instance) you will see the complexities that are involved. As one simple example, we should point out that two superficially contrasting forms (e.g., [p] and [b]) may sometimes result not from two phonemes (/p/ and /b/), but from a distinct sequence of sounds underlingy (/p/ and /mp/, for instance). The superficial contrast between the two sounds simply tells the analyst that something is different between the forms underlingly – but it doesn’t tell you what is different.
Day 13: The big picture

In today’s class we are going to define the concept of the syllable, consider its internal hierarchical structure, and discuss some preliminary principles for dividing phonological words into syllables.

• We have an intuitive, native speaker awareness about the existence of syllables. Syllables have a certain psychological reality, e.g., we know how many syllables there are in an English word. Cf. native-devised syllabaries.

• Position of the syllable in the phonological hierarchy: we will consider it to be the next larger unit after the segment.

• Historically the syllable was defined in phonetic terms as a sequence of sounds whose articulation involves a single chest pulse of pulmonic energy. This is now known not to be true.

• So it is perhaps best to define the syllable in phonological terms as a sequence of one or more adjacent segments which cluster together around a single peak of sonority.

• All syllables have a characteristic internal structure: the _________ is obligatory and in fact can be seen as the defining element in the existence of a syllable.

• In addition to the nucleus, there are possible marginal elements on either side, structured as the onset and the coda. The nucleus or peak + the coda make up the rhyme (rime).

• All speech sounds can be grouped together in terms of their relative sonority, and there is a natural and universal pressure to group the more sonorous segments around the nucleus; cf. the “wave” effect of sonority. Sonority tends to rise in the onset and decrease in the rhyme/coda.

• Every language has some degree of restrictions and constraints on types of syllables and how they may go together within the word.

• The universal syllable type, found in all languages, is CV, such as [ta].

• Most languages also allow other more complex types of syllables.

• The syllable CV is considered to be basic; other patterns can be derived from it by the addition or deletion of one element. Thus the presence of CCCV in a language implies the existence of CCV as well, etc.

• For each language we study, we want to come up with the “maximal syllable template,” summarizing all of the individual syllable types with one all-inclusive formula.

• In addition to the maximal syllable template, we also want to state any observable restrictions on the concatenation of syllables in the word or of segments in the syllable, e.g., the onset is obligatory.

• Huariapano (Peru): [CGVC], where G = glide.

• So permissible syllables would be [a], [ta], [tas], [as], [twa], [twas].

• Some examples of non-occurring syllables are *[pla], *[tamz], etc.

• Ashéninca (Peru): [CVVN], where N = nasal. Onset obligatory except word-initially.

• There is a universal principle called the _______ Onset Condition, which is strictly followed in many languages. The most common exception to this is word-initial syllables.

• So permissible syllables in Ashéninca would be [ta], [tan], [taː], [taːn]. Also, word-initially you can get [a], [aː], [an], and [aːn].

• Some syllable types which do not occur are *[tas], *[twa], *[twaː], *[twaːn], and *[twaːn].
We want to describe any constraints on what segments can be combined within the syllable, e.g., possible consonant clusters. This is known as "phonotactics." We can formulate these predictable patterns in terms of "if-then" conditions. For example, in English an onset cluster consisting of three consonants may only be of limited types:

\[
\text{spr} \quad \text{spl} \quad \text{str} \quad \text{skr} \quad \text{skl} \quad \text{spj} \quad \text{skj} \quad \text{skw}
\]

The first one must be /s/, the second one a voiceless stop, and the third one [+sonorant] and [–nasal].

Spanish: [CCVCC], as in the word *transcribir* ‘to transcribe’.

Initially we might be inclined to posit a maximal syllable template of [CCCVVC] for Chamicuro. But this makes a wrong prediction, since no syllables contain both a long vowel and a coda consonant at the same time.

So our final template for now will be: [CCCVX]σ_{max}

Also, onsets are obligatory except word-initially.

Some patterns of preference, based on frequency:

\[
\begin{align*}
\text{CV} &> \text{CVC} \\
\text{V} &> \text{VC} \\
\text{CCV} &> \text{CCVC} \quad \text{so the lack of a coda consonant is better than the presence of a coda C} \\
\text{CV} &> \text{CV}: \\
\text{V} &> \text{V}: \quad \text{so short vowels are preferred over long vowels}
\end{align*}
\]

Given the basic preference for CV syllables, all else being equal a string VCV will be universally syllabified as V.CV rather than as VC.V

This is known as the “universal syllabification algorithm,” and CV is often called the core syllable type.

A related principle of syllabification is known as “maximize the onset.”

Illustrate with the hypothetical word /tabriolampe/. Initially we might parse this as follows:

\[
[\text{tabriolampe}]
\]

But then based on the notion of the sonority wave effect mentioned above, we may need to. We may onset first principle to avoid violations of the sonority hierarchy. So the final syllabification would be:

\[
[\text{tabriolamp.e}]
\]

In other cases which may be ambiguous, rely on phonotactic restrictions relevant in the language to determine where consonant clusters should be divided: the “Maximal Cluster Approach.”

Cf. Spanish, where #C does not occur, so we syllabify the word /español/ as [es.pañol].
Homework for Day 15:
- Read Marlett, chapters 27-29
- Read Interpretation paper (in the course packet)

Day 14: The big picture
In today’s class we are going to consider how to interpret phones or sequences of phones which are potentially ambiguous in specific languages, and what implications these types of decisions have in terms of the resulting syllable structure.

Review:
What are the different constituent parts of the syllable?
  - onset
  - nucleus
  - coda
  - rhyme
What are:
  - open syllables?
  - closed syllables?
  - heavy syllables?
  - light syllables?

Interpretation
  - Also known as ___________________________________________.
    It can be defined as the process of deciding how certain ambivalent / ambiguous segments and/or sequences should be analyzed as patterning in a specific language.
  - There are two main issues to consider:
    (1) Is the segment x a vowel or a consonant?
    (2) Is the string xy a single phonological unit or a __________________ of two separate segments?
  - Question 1 involves determining whether a certain segment should be linked to a C position or a V position at the top of the tree diagram of its features.
  - It is necessary to deal with this because some phones which sound like a vowel may actually function as a consonant in a particular language, and vice-versa, such as [u] vs. [w].

  • Stops and fricatives are unambiguously considered consonants, and mid and low vocoids are unambiguously considered vowels. This leaves us basically with [h], high vowels, glides, and some types of approximants which may have an ambivalent function from language to language.
  • 1 or 2 segments?
  • Normally, one phonetic segment is straightforwardly attached to one single C or V “slot” in the phonological representation.
  • However, certain combinations of particular phonetic “segments” are consistently treated in some languages as phonological units or wholes which are not separated. For example, in Polish there is a phonemic contrast between the affricate [tʃɨ] and a sequence of stop plus fricative, at the same place of articulation:
    /tʃɨ/ ‘three’ vs. /tʃɨ/ ‘whether’ (/čɨ/)

  • We will identify the latter as a single “segment” (an affricate) for the purposes of our phonological representation of this language, just like the cases which are more clear and obvious.
  • Here are some examples of potentially ambiguous sequences that need to be interpreted:

    | Unit interpretation (one segment) | Cluster interpretation (two segments) |
    |---------------------------------|-------------------------------------|
    | (homorganic) affricates         | stop + fricative                    |
    | aspired consonant               | consonant + h                       |
    | palatalized consonant           | consonant + j                       |
    | labialized consonant            | consonant + w                       |
    | syllabic consonant              | vowel + consonant or consonant + vowel |

  • What helps us to determine how these ambivalent items are to be interpreted phonologically? In other words, what are the criteria we should use in making decisions of this type?
  • The key is always to look at the overall patterns of the language as a whole and interpret them in such a way as to conform to the known regularities.
  • Evidence for the proper interpretation may come from any area of the phonological structure of the language in question.

  (1) __________ structure: in some simple cases, the phonotactic restrictions and/or the maximal syllable template might help us decide one way or the other.
  • For example, take a language with an otherwise regular (________) template of [CVC], i.e., no consonant clusters are allowed within the onset or the coda constituent. In this case a word-initial sequence of [ts] would best be interpreted as a single segment, that is, an affricate. This particular affricate (voiceless alveolar) is then sometimes rewritten (retranscribed) as [c] or [ɕ] after this decision is made, analogously to the alveopalatal affricate [ć].
when one member of a phonetic
yet necessarily

Principles of Phonological Analysis GIAL

•k u
→
•t
ʃ
•t
ł
→
•t
s
•u
→

(3) pattern in the phonemic chart:

• In English, we notice that we have coronal fricatives at the alveolar ([s z]) and alveopalatal ([ʃ ʒ]) points of articulation. We also have the alveolar stops [t d], but no corresponding alveopalatal stops. However, we do have the complexes of [tʃ] and [dʒ], so a nice symmetrical interpretation would consider these “clusters” to really be single phonological units or “stops” at a more abstract level, to round out the inventory of obstruents.

(2) productive of the language:

• For example, suppose there is a process by which a single consonant

(4) distribution: when one member of a phonetic complex never occurs in isolation, it is usually best to interpret it as being part of a single segment rather than a sequence (cluster).

• For example, in Chamicuro there is a lateral affricate comprised phonetically of [dl]. The alveolar stop occurs nowhere else in the language, i.e., never by itself without being followed by [l]. Furthermore, the other two voiced stops [b] and [ɡ] are never heard at all anywhere in the language, either in isolation or in clusters with other consonants. So in this case the [dl] sequence in Chamicuro is interpreted as a single unit (a lateral affricate). Then we can say that in this language there are no voiced obstruents in any phonetic forms. This is further confirmed by the fact that [dl] is an intervocalic allophone of /l/, and /l/ is clearly a single segment, not a sequence.

• This is confirmed by the fact that in English, the voiceless affricate /tʃ/ is allophonically aspirated in syllable-initial position, just like the three normal stops /p t k/ are.

• So maybe there is a gap in the pattern that an ambivalent item or sequence can fill.

Pike’s Fourth Premise for Phonological Analysis: Characteristic sequences of sounds exert structural pressure on the phonemic interpretation of suspicious segments or suspicious sequences of segments.

• Tetelcingo Nahuatl (p. 165)
• Canonical syllable types:
   • CV
   • V
   • CVC
   • VC
• Tentative maximal syllable template: [CVC].
• Onset is not obligatory, at least word-initially.
• Preliminary phonetic charts:

• (Note that these are not phonological rules as we have normally seen them up to this point. Rather, the arrows in the formulas above actually refer to interpretive decisions that we have made.)

• The next step would then be to add these new phonological units to the working phone charts. However, we have not yet necessarily determined that all of these new segments are contrastive phonemes in their own right. The final step then would be to compare these with their suspect counterparts in order to establish what types of phonological relationships they are in (contrast or complementary distribution).

p t k (i) (u) s f l o m n e ø l a

• Now let’s list and then work on analyzing the potentially ambiguous segments and sequences:

i → V → l
i → C / V V (i)
u → C / V V (w)
ts → C → ts (affricate)
tl → C → tl (affricate)
f → C → (affricate)
k → C (?) → k* (labialized stop)
**Interpretation / resegmentation**

### Ambiguous vocoids and contoids

<table>
<thead>
<tr>
<th>VOCOIDS which may function as <em>consonant</em>-like segments</th>
<th>CLUES for interpretation decisions</th>
<th>CONTOIDS which may function as <em>vowel</em>-like segments</th>
</tr>
</thead>
<tbody>
<tr>
<td>high vocoids</td>
<td>occurring adjacent to other vocoids, less stressed, less prominent, less friction, or shorter than neighboring vocoids:</td>
<td>semi-vowels (glides)</td>
</tr>
<tr>
<td>[i], [y], [u], [ɯ]</td>
<td>→ → →</td>
<td>[j], [w]</td>
</tr>
<tr>
<td></td>
<td>stressed or prominent, syllabic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>← ← ←</td>
<td></td>
</tr>
<tr>
<td>voiceless vocoids</td>
<td>occurring in consonant position (borders of syllables); quality is not contrastive</td>
<td>[h]</td>
</tr>
<tr>
<td>[I], [E], [A], [O], [U]</td>
<td>→ → →</td>
<td></td>
</tr>
<tr>
<td></td>
<td>occurring in vowel position (nucleus of syllables); quality is perceptible and contrastive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>← ← ←</td>
<td></td>
</tr>
<tr>
<td>high back unrounded vocoid</td>
<td>occurring in consonant position</td>
<td>lenis voiced velar fricative</td>
</tr>
<tr>
<td>[ɯ]</td>
<td>→ → →</td>
<td>[γ]</td>
</tr>
<tr>
<td></td>
<td>occurring in vowel position; stressed or prominent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>← ← ←</td>
<td></td>
</tr>
<tr>
<td>retroflexed central vocoid</td>
<td>occurring in consonant position</td>
<td>retroflexed</td>
</tr>
<tr>
<td>[ǝ]</td>
<td>→ → →</td>
<td>[ɹ]</td>
</tr>
<tr>
<td></td>
<td>occurring in vowel position</td>
<td></td>
</tr>
</tbody>
</table>
Segments which are generally unambiguous:

contoids when preceded or followed by a vowel (except a high vowel or a mid central vowel, i.e., [ə]):

<table>
<thead>
<tr>
<th>labial</th>
<th>alveolar</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stops</td>
<td>p</td>
<td>t</td>
</tr>
<tr>
<td>voiced stops</td>
<td>b</td>
<td>d</td>
</tr>
<tr>
<td>voiceless fricatives</td>
<td>f</td>
<td>s</td>
</tr>
<tr>
<td>voiced fricatives</td>
<td>v</td>
<td>z</td>
</tr>
<tr>
<td>nasals</td>
<td>m</td>
<td>n (followed by a vowel)</td>
</tr>
<tr>
<td>laterals</td>
<td></td>
<td>l</td>
</tr>
<tr>
<td>other liquids</td>
<td></td>
<td>r (followed by a vowel)</td>
</tr>
</tbody>
</table>

vocoids occurring in isolation, or when preceded by any consonant, or when followed by any consonant except [n] or [r]:

<table>
<thead>
<tr>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mid</td>
<td>e  ø</td>
<td>γ  o</td>
</tr>
<tr>
<td>low</td>
<td>æ  a</td>
<td>ɔ  ë</td>
</tr>
</tbody>
</table>

Sequences which are generally unambiguous:

fricative + stop | [fp] [st] [sp] [xb] [sk] etc.
stop + sharply heterorganic fricative or nasal (does not include [tʃ] or [pf], which are nearly homorganic) | [kf] [pn] [pn] [ks] etc.
nasal + heterorganic stop | [mt] [nk] [np] etc.
nasal + fricative (or vice-versa) | [mf] [nx] [fm] [xn] etc.
fricative + heterorganic fricative | [sx] [fs] etc.
non-high, non-identical vocoid sequences | [ea] [æe] [eo] [oe] [ao] etc.
{stop, fricative, nasal, or lateral} + stressed voiced vocoid, followed by a consonant (or word-final) | [ˈpat] [ˈpek] [ˈsos] [ˈnip] [ˈluf] etc.
consonant + voiced, non-high vocoid (except mid central [ə], [ʌ]) | [pa] [ve] [mæ] [so] [lɛ] [ra] etc.
non-high vocoid (except mid central [ə], [ʌ]) + stop or lateral | [ap] [et] [ok] etc.
| [æl] [ol] [al] etc.
vocoid sequences including a stressed [ə] | [aˈə] ['sə] etc.
## Ambiguous sequences and segments

<table>
<thead>
<tr>
<th>SEQUENCES which may pattern as single segments</th>
<th>Example</th>
<th>SEQUENCES which may pattern as sequences</th>
<th>Example</th>
<th>Complex Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
<td></td>
<td>Example Complex Unit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>any contoid + unstressed high vocoid</td>
<td>[Ci] [Cy] [Cu] [Ci] [Cw]</td>
<td>[C^n] / [C^j]</td>
<td>[k^w] [k^x] [k] [p^j] etc.</td>
<td>labialized or palatalized consonant</td>
</tr>
<tr>
<td>any contoid + [w], [j]</td>
<td>[k^w] [t^j] [k^w] etc.</td>
<td>[p^j] [t^j] [b^j] etc.</td>
<td>[p^j] [t^j] [b^j] etc.</td>
<td>aspirated stop glottal stop</td>
</tr>
<tr>
<td>stop + [h] (or [fi])</td>
<td>[ph] [th] [kh] [bfi]</td>
<td>[k^h] [t^j] [b^j] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>aspirated stop homorganic affricate</td>
</tr>
<tr>
<td>stop + voiceless vocoid</td>
<td>[pE] [tA] [b1] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>aspirated homorganic affricate</td>
</tr>
<tr>
<td>stop + homorganic fricative</td>
<td>[kx] [bβ] [tʃ] [tʃ] [pf]</td>
<td>[t^j] [b^j] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>aspirated homorganic affricate</td>
</tr>
<tr>
<td>stop + homorganic fricative + [h] (or [fi])</td>
<td>[ts] [dz] [tʃh] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>glottalized (ejective) affricate</td>
</tr>
<tr>
<td>stop + homorganic fricative + voiceless vocoid</td>
<td>[tsA] [dzI] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>voiceless stop + glottal stop</td>
</tr>
<tr>
<td>stop + homorganic fricative + glottal</td>
<td>[tʃʔ] [dzʔ] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>[t^j] [b^j] etc.</td>
<td>prenasalized stop or affricate</td>
</tr>
<tr>
<td>voiced stop + mid central vocoid (schwa)</td>
<td>[ba] [da] [ga]</td>
<td>[b^j] [d^j] [g^j]</td>
<td>[b^j] [d^j] [g^j]</td>
<td>stop with vowel off-glide (released stop)</td>
</tr>
<tr>
<td>voiceless stop + glottal stop</td>
<td>[p?] [t?] [k?]</td>
<td>[p^j] [t^j] [k^j]</td>
<td>[p^j] [t^j] [k^j]</td>
<td>glottalized (ejective) stop</td>
</tr>
<tr>
<td>alveolar stop + lateral</td>
<td>[tʃ] [dʃ] [tʃ]</td>
<td>[t^j] [d^j] [t^j]</td>
<td>[t^j] [d^j] [t^j]</td>
<td>lateral affricate</td>
</tr>
<tr>
<td>velar stop + labial stop</td>
<td>[kp] [gb]</td>
<td>[kp] ([k^p]) [gb] ([g^b])</td>
<td>[kp] ([k^p]) [gb] ([g^b])</td>
<td>double stop</td>
</tr>
<tr>
<td>stop or affricate + homorganic nasal</td>
<td>[pm] [bm] [tn] [tn]</td>
<td>[p^m] [b^m] [t^j] [d^j]</td>
<td>[p^m] [b^m] [t^j] [d^j]</td>
<td>post-nasalized stop or affricate</td>
</tr>
<tr>
<td>vocoid + glottal stop (either preceding or following) or sequence of vocoid + glottal stop + like vocoid</td>
<td>[ʔa] [a?] [ʔi] [ʔi]</td>
<td>[t^j] [a^j] ([t^j] [i^j]) ([V^j])</td>
<td>[t^j] [a^j] ([t^j] [i^j]) ([V^j])</td>
<td>pre- or post-glottalized vocoid (checked vocoid)</td>
</tr>
<tr>
<td>vocoid + glottal</td>
<td>[a?] [o?] [e?]</td>
<td>[a] ([a^j] [o] ([o]) [e] etc.</td>
<td>[a] ([a^j] [o] ([o]) [e] etc.</td>
<td>laryngealized vocoid (creaky voice)</td>
</tr>
<tr>
<td>glottal stop + stop</td>
<td>[ʔk] [ʔb] etc.</td>
<td>[t^j] [b^j] ([t^j] [b^j]) etc.</td>
<td>[t^j] [b^j] ([t^j] [b^j]) etc.</td>
<td>stop with glottal onset (implosive) nasal with glottal onset</td>
</tr>
<tr>
<td>glottal stop + nasal</td>
<td>[ʔn] [ʔm] [ʔn] etc.</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>nasal with glottal onset</td>
</tr>
<tr>
<td>glottal stop + lateral</td>
<td>[ʔl]</td>
<td>[t^j] [l^j]</td>
<td>[t^j] [l^j]</td>
<td>lateral with glottal onset</td>
</tr>
<tr>
<td>glottal stop + high, non-syllabic vocoid</td>
<td>[ʔu] [ʔw] [ʔi] [ʔi]</td>
<td>[t^j] [w^j] [ʔi] [ʔi]</td>
<td>[t^j] [w^j] [ʔi] [ʔi]</td>
<td>vocal with glottal onset</td>
</tr>
<tr>
<td>nasal + homorganic stop</td>
<td>[mp] [mb] [nt] [nk]</td>
<td>[t^m] [b^m] [t^j] [k^j]</td>
<td>[t^m] [b^m] [t^j] [k^j]</td>
<td>prenasalized stop</td>
</tr>
<tr>
<td>nasal + homorganic affricate</td>
<td>[nts] [ndz] [ntʃ]</td>
<td>[t^ʃ] [d^z] [t^j] [ʃ]</td>
<td>[t^ʃ] [d^z] [t^j] [ʃ]</td>
<td>nasal + homorganic affricate</td>
</tr>
<tr>
<td>nasal + homorganic affricate + aspiration</td>
<td>[nʃ]</td>
<td>[t^ʃ] [d^z]</td>
<td>[t^ʃ] [d^z]</td>
<td>nasal + homorganic affricate</td>
</tr>
<tr>
<td>nasal + homorganic affricate + glottal stop</td>
<td>[ndzʔ]</td>
<td>[t^ʃ] [d^z]</td>
<td>[t^ʃ] [d^z]</td>
<td>nasal + homorganic affricate</td>
</tr>
<tr>
<td>nasal + mid central vocoid</td>
<td>[mə] [nə]</td>
<td>[ʃ]</td>
<td>[ʃ]</td>
<td>nasal with glottal onset</td>
</tr>
<tr>
<td>stop with glottal onset (implosive) nasal</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>nasal with glottal onset</td>
</tr>
<tr>
<td>nasal with glottal onset</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>nasal with glottal onset</td>
</tr>
<tr>
<td>voiced with glottal onset</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>[t^j] [m^j] [ʔg] etc.</td>
<td>nasal with glottal onset</td>
</tr>
<tr>
<td>nasal + homorganic stop</td>
<td>[mp] [mb] [nt] [nk]</td>
<td>[t^m] [b^m] [t^j] [k^j]</td>
<td>[t^m] [b^m] [t^j] [k^j]</td>
<td>prenasalized stop</td>
</tr>
<tr>
<td>nasal + homorganic affricate</td>
<td>[nts] [ndz] [ntʃ]</td>
<td>[t^ʃ] [d^z] [t^j] [ʃ]</td>
<td>[t^ʃ] [d^z] [t^j] [ʃ]</td>
<td>nasal + homorganic affricate</td>
</tr>
<tr>
<td>nasal + homorganic affricate + aspiration</td>
<td>[nʃ]</td>
<td>[t^ʃ] [d^z]</td>
<td>[t^ʃ] [d^z]</td>
<td>nasal + homorganic affricate</td>
</tr>
<tr>
<td>nasal + homorganic affricate + glottal stop</td>
<td>[ndzʔ]</td>
<td>[t^ʃ] [d^z]</td>
<td>[t^ʃ] [d^z]</td>
<td>nasal + homorganic affricate</td>
</tr>
<tr>
<td>nasal + mid central vocoid</td>
<td>[mə] [nə]</td>
<td>[ʃ]</td>
<td>[ʃ]</td>
<td>nasal with glottal onset</td>
</tr>
<tr>
<td>声母 + 点声</td>
<td>[mʘ] [nʘ]</td>
<td>[ⁿʘ] [ⁿʘ]</td>
<td>偏音化点声</td>
<td></td>
</tr>
<tr>
<td>序列</td>
<td>[HL] [LH] [LHL] etc.</td>
<td>[HL] [LH]</td>
<td>点声序列（音高序列）</td>
<td></td>
</tr>
<tr>
<td>影响音序列，尤其是与[i]或[u]作为上滑音或下滑音影响音，其中影响音不是长音</td>
<td>[iu] [iə] [ei] [ou] [œ] [aœ]</td>
<td>[iʰ] [ʼ] [eʼ] [oʼ] [œ] [aœ]</td>
<td>声母滑音（平滑音）</td>
<td></td>
</tr>
<tr>
<td>影响音序列</td>
<td>[aa] [ss] [tt] [mm]</td>
<td>[aː] [sː] [tː] [mː]</td>
<td>声母长音（二重音）</td>
<td></td>
</tr>
<tr>
<td>影响音 + 基音（通常基音将具有与后续声母相同的 articulation）</td>
<td>[an] [in] [em] [aŋ]</td>
<td></td>
<td>声母化影响音</td>
<td></td>
</tr>
<tr>
<td>影响音 + [ɻ]</td>
<td>[aɻ] [oɻ] etc.</td>
<td>[a] [ʊ] etc.</td>
<td>后退影响音</td>
<td></td>
</tr>
<tr>
<td>齿音或齿龈基音在 [ɻ] 前或后</td>
<td>[ʃ] [ʃ] [ʒ]</td>
<td>[ʃ] [ʃ] [ʒ]</td>
<td>后退基音</td>
<td></td>
</tr>
<tr>
<td>中央非低基音 ([i] [o] [ʌ]) + {声母, 横向, 或 [ɻ]}</td>
<td>[im] [ən] [ʌŋ] [il] [əl] [ʌl] [iɻ] [əɻ] [ʌɻ]</td>
<td>[ŋ]</td>
<td>声母鼻音化</td>
<td></td>
</tr>
<tr>
<td>任何齿音 {闭塞, 附加, 摩擦音, 或声母} 或任何基音基音 + {[ɻ] 或 [i]}</td>
<td>[ti] [dʒi] [tsi] [si] [ni] [ki] [xi]</td>
<td>[tʃ] [dʒ] [ʃ] [ʃ]</td>
<td>声母基音化，基音化，或前基基音</td>
<td></td>
</tr>
</tbody>
</table>

（原始来源：Jim Roberts）
Thus far, we have been analyzing phonetic data in order to identify the phonemes of a language, that is, its distinctive or functional units of sound. We have taken for granted the phonetic transcriptions that we have been given. However, in practical field work, questions often arise which are related to the sound patterns of a language that affect these transcriptions.

Remember that speech in fact involves a continuous stream of acoustic signals that is not neatly broken up into individual segments as we have implied in our phonetic transcription. Sounds merge and blend into one another. Nevertheless, we all perceive our words as sequences of discrete units, or segments. This perception is a result of our internalized knowledge of the patterns of our language. We would all agree, as native speakers of English, that the word *cat* has three segments, and transcribe it phonetically as [kʰæt]. For someone who is not familiar with the structure of English, though, this would not be obvious. He or she might, for example, hear the aspiration of [kʰ] and interpret it as a distinct segment, not realizing that it is supposed to be associated with the release of the [k]. This person would then give [khæt], with four segments, as the phonetic transcription.

A basic assumption of phonological analysis, then, is that all speech can be divided into discrete segments. However, the way this should be done for any particular language may vary, and depends on the patterns of that language. When we set about to analyze the phonology of a language, we must bear in mind that the way we hear the phonetic “facts” may not correspond to the way the native speaker hears and organizes them. Interpretation of the data, then, is a central problem of phonological analysis, and will concern us throughout the semester. The phonemic analysis we have just studied is one type of interpretation of data. Another type of interpretation that we shall consider now is resegmentation.

The example of [kʰæt] above points out another type of interpretation that is involved in phonological analysis, called resegmentation. We will distinguish two types:

1. Certain segments may function in some cases as consonants and in other cases as vowels. The way we interpret these sounds should be consistent with the overall patterns of the language.
2. Similarly, there are certain combinations of sounds that sometimes function as a single segment, or otherwise as a sequence of two (or even more) segments. In the example of *cat* above, the phonetic sequence [kh] is to be interpreted as a single segment, [kʰ]. The basis for this decision is again rooted in the overall patterns of the language.

We will now consider these two kinds of resegmentation in turn. Several types of patterns will be adduced as evidence for decisions in interpretation. When we do field work with a language, we must be constantly aware that our transcriptions really involve hypotheses about the system of
the language; consequently, we may have to reconsider these hypotheses at any stage of the analysis, whenever we find evidence relevant to resegmentation or other issues.

1. Consonant or vowel?

There are two basic, traditional types of segments: consonants and vowels. Some sounds, such as [p], function as consonants in any language, and some sounds, such as [a], function as vowels in any language. (We will call such segments univalent, unambiguous, or canonical sounds.) There are other sounds, however, which function as consonants in one language and vowels in another. Or, in a given language, such segments may function as consonants in one position (environment) and as vowels in another position. These segments, such as high vocoids and voiceless vocoids, we will call ambivalent or ambiguous sounds. An accompanying handout shows those sounds that are identified as having ambivalent function. In the following paragraphs we will examine various types of evidence that enable the analyst to make an interpretation concerning ambivalent segments.

1.1 Evidence from syllable patterns

A common type of evidence that is used to decide between a consonant vs. vowel interpretation is based on the syllable (or CV, i.e. consonant-vowel) patterns found in the language. Notice the following data from Towa (New Mexico):

- [pɔ] 'flowers'
- [tɔ] 'flesh'
- [pe] 'sun'
- [ϕi] 'buckskin'
- [ʔe] 'smoke'
- [tæ] 'where'
- [ϕə] 'hair'
- [ve] 'bought'

If we were to list the CV patterns that we find in our data (the unambiguous ones), we would find that we only have the syllable type CV (one consonant followed by one vowel). On the basis of that pattern, how would the non-syllabic u in the following word best be interpreted?

- [u]i ‘four’

The non-syllabic u would probably be best interpreted as [w], unless other considerations are involved, since there are no attested vowel clusters in our data. We will then (re)transcribe the word ‘four’ as [wi], which then by means of contrast might be established as /wi/ phonemically. We would have a phoneme /w/ in addition to the phoneme /u/ (assuming that Towa has words with [u]). We will not say that there is one phoneme /u/ with an allophone [w] that occurs before vowels because this latter solution would claim that the phonemic representation of ‘four’ is /ui/, which goes against the CV pattern of the language.

Now consider the following data from Cotabato Manobo (Philippines):

- [ʔudan] ‘rain’
- [ʔetaw] ‘person’
- [meʔemis] ‘sweet’
- [kilat] ‘thunder’
- [ʔenai] ‘sand’
- [sabitan] ‘belt’
- [silaʔ] ‘lightning’
- [suqui] ‘soul’
- [melalaʔ] ‘yellow’
Although we do not have syllable divisions indicated here, a number of observations about syllable structure can be made. (Besides, syllable breaks are also very difficult to ascertain, and are themselves open to interpretation.) Since each word begins with exactly one consonant, we infer that a syllable may likewise begin with one (but not more than one) consonant. Similarly, in unambiguous cases words end with one consonant; we infer that syllables may not end in more than one consonant. In the clear cases, then, we have the following syllable patterns:

- CV: sa.bau, me.la.lag, ba.bui
- CVC: me.la.lag, ge.tan, me.?e.mis

Now we turn to the ambiguous cases, involving high vowels or glides. First, we will interpret such segments occurring between nonambiguous consonants as vowels serving as syllable peaks. Other segments are then interpreted so that the words all contain only CV or CVC syllables:

- sabaw -- rather than sabau
- kilaj -- rather than kilai or kjlai or kjlaj
- ?enaj -- rather than ?enai
- ?etaw -- rather than ?etau
- babuj -- rather than babui or babwi or babwj
- sugui -- rather than sugui or sugwi or sugwj or swgui, etc.
- baji -- rather than baii or bajj or baij

Now examine the following words from Towa for another example:

- [i̥] ‘clothes’
- [ɛ̥] ‘axe’

The voiceless vocoid [i̥] could be interpreted as a voiceless vowel (which would then need to be considered with [i] as a suspicious pair) or it could be interpreted as a consonant. Since the only attested pattern found above was CV, it would be most reasonable to interpret it as a voiceless consonant. Because of the dominant CV pattern, we would not try to explain [i̥] as an allophone of /i/ (unless other considerations are involved). The segments [i̥] and [ɛ̥] are a suspicious pair, however, and can be considered to be allophones of one phoneme for which the features of the tongue position are not specified – /h/. Since /h/ is a voiceless glottal fricative, a rule is needed to give the correct specification of these features in a particular word in Towa:

\[
    h \rightarrow V_i / ___ V_i
\]

In other words, /h/ becomes the voiceless counterpart of the following vowel. (The subscript i on the two V’s is used to indicate that these two vowels are identical.)
1.2 Evidence from rules

Another type of evidence for interpreting a segment as either a consonant or as a vowel is based on the way certain rules work. For example, the rule that assigns stress in Seri puts the primary stress on the first vowel of the root. (Morpheme boundaries are indicated in the data below by ‘-’.)

- [k-išni] ‘carry (nom.)’
- [k-ápok-im] ‘put into mouth dry (nom.)’
- [k-aʔ-imex] ‘cause to flow (nom.)’

On the basis of the evidence from that stress rule we would be led to posit the initial segment of each of the following verbs as the consonant /j/ rather than the vowel /i/.

- [k-já:] ‘own (nom.)’ (rather than [k-iá:])
- [k-já:i] ‘be expensive (nom.)’ (rather than [k-iá:i])
- [k-jáił] ‘be deep (nom.)’ (rather than [k-iáił])

If we had posited these stems as -ia:, -ia:i, and -iaił, it would be very difficult to write a general stress rule since the stress rule would have to put stress on the second vowel of these stems.

The decision to consider these as “j”-stems rather than “i”-stems is also supported by other rules in the language. For example, [m] partially assimilates to the point of articulation of a following consonant under certain conditions. It remains [m] before vowels.

- [jo - m - í:p:æ] ‘it wasn’t good’
- [jo - m - áfp] ‘he didn’t arrive’
- [jo - m - pánšR] ‘he didn’t run’
- [jo - n - tópl] ‘it wasn’t stained’
- [jo - ŋ - xiš:i] ‘it wasn’t spicy hot’

The fact that m becomes n before the stems with the ambivalent segment below supports the hypothesis that these are “i”-stems.

- [jo - n - já:i] ‘it wasn’t expensive’

Evidence from rules is generally stronger than evidence from syllable patterns. Therefore this type of evidence for resegmentation is highly valued.

2. One or two (or more) segments?

Certain portions of an utterance are ambivalent in that they may be interpreted as either a single, complex segment or as a sequence of segments. The sound [tʃ] could also be transcribed [tf] – the phonetics of the two are the same but in one case it is implied that there is one sound – a stop
with a strident release – and in the other case it is implied that there are two sounds – a stop followed by a strident alveopalatal fricative. We will call these ambivalent sequences. Univalent sequences are those which do not have a unit interpretation in any language – [pt], [sk], [mg], for example. In the following paragraphs we will examine various types of evidence that enable the analyst to make an interpretation concerning ambivalent sequences.

2.1 Evidence from syllable patterns

The CV patterns of a language are one type of evidence used in interpreting ambivalent sequences. In considering the following data from Orizaba Nahuatl, we first examine data without ambivalent sequences.

[apiʃakak]  'he roasted it'
[kemah]  'yes'
[animahteh]  'hearts'
[afkan]  'now'

The following CV patterns are observed in the words above:

V (as in [apiʃakak] and [animahteh])
VC (as in [afkan])
CV (as in [kemah] and [animahteh])
CVC (as in [apiʃakak] and [animahteh])

It appears that consonant clusters occur only where two syllables are juxtaposed – no word begins or ends with a consonant cluster, so we infer that no syllable may begin or end with a consonant cluster. Now consider the following, additional data in which the ambivalent sequences [ts], [ʃ], and [ł] occur.

[molkatl]  'small ear of corn'
[komitl]  'cooking pot'
[konetl]  'baby'
[sajoltla]  'there are many mosquitoes'
[tlatspanon]  'broom'
[notʃpotʃsin]  'my dear young lady'

If these ambivalent sequences are interpreted as sequences, we will have additional CV patterns in which only the consonants [t], [s], [ʃ], and [ł] may occur – for example, CVCC as in [molkatl]. If these are interpreted as affricates, on the other hand, we will potentially be adding three phonemes to our inventory (if they all contrast, that is) - /tʃ/, /ts/, and /tl/. (In the Americanist alphabet there are independent symbols for these: /č/, /¢/, and /ł/, respectively.) In this case the evidence for a unit interpretation is compelling. The words could then be transcribed and subsequently phonemicized as /molkaʃ/, /nočpočin/, etc.
Ambivalent sequences include stops preceded by homorganic nasal consonants. In Highland Mixtec (Santo Tomas dialect), the syllable pattern is strongly CV. (Tone is omitted from the following data.)

\[
\text{[pipi]} \quad \text{‘baby turkey’} \quad \text{[taka]} \quad \text{‘nest’}
\]

Therefore the ambivalent sequence in the following word is interpreted as a unit.

\[
\text{[tʃaka]} \quad \text{‘fish’} \quad \text{(resegmented as [čaka])}
\]

Likewise, the ambivalent sequences in the following words are interpreted as units – prenasalized stops and affricates.

\[
\text{[ⁿdika]} \quad \text{‘banana’} \\
\text{[ⁿdʃitʃi]} \quad \text{‘it dried’} \\
\text{[ⁿgaⁿdiː]} \quad \text{‘sun’}
\]

This interpretation is supported by the fact that all voiced stops and affricates that occur in this language are prenasalized. (Therefore \([ⁿdʃ]\) could be represented phonemically as /dʒ/, [ⁿd] as /d/, etc.)

We have already seen that a basic syllable pattern of Towa is CV. The long vowels in the following data will be interpreted as long vowels (units), not as sequences of identical vowels, since vowel clusters do not (otherwise) occur in our data.

\[
\text{[φoː]} \quad \text{‘branch’} \quad \text{[peː]} \quad \text{‘heart’} \\
\text{[sæː]} \quad \text{‘messenger’} \quad \text{[pe]} \quad \text{‘sun’}
\]

If these vowels were interpreted as clusters, they would be written (retranscribed) as [φoo], etc.

### 2.2 Evidence from frequency

A less important but possible consideration in interpretation is the frequency of occurrences of the ambivalent sequence. An ambivalent sequence that has a high frequency of occurrence would be especially suspicious and likely to be a unit. (See the Orizaba Nahuatl data above.)

### 2.3 Evidence from limited distribution

If a segment occurs only in an ambivalent sequence, the ambivalent sequence should be interpreted as a unit (unless there is other, stronger evidence to the contrary). For example, [ʃ] occurs only following [t] in Spanish – never elsewhere.

\[
\text{[tfuло]} \quad \text{‘cute’} \quad \text{[mutʃо]} \quad \text{‘a lot’} \\
\text{[tfiko]} \quad \text{‘small’}
\]
There are no words like *[julo], *[jiko], or *[muʃo]. Therefore the ambivalent sequence would be interpreted as [tʃ] ([tʃ]). Spanish uses the digraph ch in its alphabet for this sound.

2.4 Evidence from allophonic distribution

If an ambivalent sequence has a complementary distribution with respect to another phonetically similar segment that is definitely a single segment, it should be interpreted as a unit. In Brazilian Portuguese, for example, [t] and [tʃ] are in complementary distribution; [tʃ] should therefore be interpreted as [tʃ], and is derived from the underlying /t/.

\[
\begin{array}{ll}
/gatu/ & [gatu] \quad \text{‘cat’} \\
/gatiɲu/ & [gatiɲu] \quad \text{‘small cat’}
\end{array}
\]

([ɲ] has already been interpreted in these data.) The rule that generates [tʃ] is given informally below:

\[
t \rightarrow \left[ \text{affricate}_{\text{alveopalatal}} \right] / \_\_\_ i
\]

By similar reasoning, aspirated stops in English are derived from unaspirated stops and therefore should be viewed as units – [pʰ], [tʰ], [kʰ] – and not as sequences of consonant plus [h].

Many (but not all) long consonants and vowels in Seri are derived from short consonants and vowels by the following rules.

\[
\begin{array}{ll}
C & \rightarrow [+\text{long}] / V \_\_\_ V \\
 & [+\text{stress}] \\
V & \rightarrow [+\text{long}] / V C \_\_\_ C \\
 & [+\text{stress}]
\end{array}
\]

\[
\begin{array}{ll}
/ʔimaf/ & \rightarrow [ʔim:a:f] \quad \text{‘my grandmother’} \\
/ʔipaʃ/ & \rightarrow [ʔip:a:f] \quad \text{‘my grandfather’}
\end{array}
\]

Therefore these lengthened consonants and vowels should be interpreted as just that – long segments – and not as sequences of identical segments.

2.5 Evidence from other phonological rules

The strongest evidence for a unit or sequence interpretation is based on how the ambivalent sequence behaves. We will consider several additional types of rules here which provide evidence for resegmentation.

The ambivalent sequence [kw] in Seri is interpreted as [kʷ] because it behaves like a single consonant in every respect. For example, the rule that lengthens consonants:
C → [+long] / V ___ V
   [+stress]
applies to [kʷ]:

/ʔókʷiʔa/ → [ʔókʷ:iʔa] ‘sawn wood’

In Karok /i/ becomes [i] when it precedes a strident consonant.

[ʔip:ih] ‘bone’
[pu:βɪʃ] ‘bag’

The rule could be stated informally as follows:

i → i / ___ C
   [+strident]

If [tʃ] is interpreted as [tʃ̥] ([ʧ̥]), then our rule does not need to be modified in order to handle the following word, since [tʃ̥] is also strident.

[pu:fitʃ] ‘deer’

If it is interpreted as the sequence [t] plus [ʃ], then our rule would have to be changed to read as follows:

i → i / ___ (t) C
   [+strident]

This change involves a seemingly ad hoc modification to the rule. Unless there were overwhelming evidence in favor of the sequence interpretation, the unit interpretation would certainly be preferred in this case.

Some phonological rules may have outputs that require different interpretations than their inputs. In Seri there is a rule that coalesces a consonant with a following glottal stop. The surface phonetic data requires a unit interpretation even though the glottal stop is part of a different morpheme at the phonemic level.

/k-ʔánl/ → [k’ánl] ‘ten (nom.)’
/kʷ-ʔ-á:s/ → [k’wá:s] ‘give to him or her to drink (imperative)’

2.6 Evidence from morpheme boundaries

When the constituents of an ambivalent sequence pertain to separate morphemes, a sequence interpretation is necessary, at least at the phonemic level. Notice the following data from Seri:
<table>
<thead>
<tr>
<th>Nominalized</th>
<th>Interrogative</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>[kpánšR]</td>
<td>[tpánšR]</td>
<td>‘run’</td>
</tr>
<tr>
<td>[kʃímpRa]</td>
<td>[tʃímpRa]</td>
<td>‘be moldy’</td>
</tr>
<tr>
<td>[k:áː:]</td>
<td>[tkáː:]</td>
<td>‘be grown up’</td>
</tr>
<tr>
<td>[ktápw]</td>
<td>[t:ápw]</td>
<td>‘disintegrate’</td>
</tr>
</tbody>
</table>

There are three ambivalent consonant sequences in these data – [t:], [k:], and [tʃ]. The long consonants will be interpreted as the clusters [t] plus [t] and [k] plus [k] respectively, since the morphemes are obviously /t-/, /k-/, and /-kái:/, /-tápẘ/, etc. The ambivalent sequence [tʃ] in [tʃímpRa] will be interpreted as the cluster [t] plus [ʃ] for the same reason.

NOTE: A sequence interpretation in one place does not necessarily mean that every instance requires a sequence interpretation. We have already seen that most long consonants of Seri are allophones of short consonants.

An English example of sequence interpretation because of morpheme boundaries is given below.

```
[howm:ejkɹ̩] is interpreted as [howmmejkɹ̩] ‘homemaker’ (from /howm/ + /mejk/ + /ɹ̩/)
```

2.7 Evidence from word games (ludlings)

Word games that demonstrate native speaker awareness of the structure of words are important evidence for interpretation. Pig Latin is a game in English in which the first consonant(s) of each word is (are) moved to the end of the word and followed by the vowel [ej]. Thus, [pɪɡ] ‘pig’ in Pig Latin comes out [ɪɡpej], and [tɹi] ‘tree’ comes out [jɪtej] or [ɪtej]. The fact that words like [tʃɪp] ‘cheap’ come out only as [iptʃej] and not *[ʃɪtej] gives evidence that [tʃ] must be treated as a unit here, and thus resegmented as [ʃ̊] or [c].
Procedures for syllable structure exercises

NOTE: always follow this for syllable structure homeworks, as well as “Hints for Homework.”

(1) Ignoring “ambiguous” sequences (such as [tʃ]) and “ambivalent” segments (such as high vocoids), construct an inventory of canonical syllable types or patterns:
   e.g. [CV] (include glosses and example numbers)
   [V]
   [CVC]
   [CCVC]

(2) Make note of any syllable positions where restrictions occur, stating where in a syllable particular phones may or may not occur. For example, the only segments which occur in syllable codas are nasal stops.

(3) Make a tentative decision about the ambivalent sequences and segments. Do they link up to a C slot, a pair of C slots, a V slot, a pair of V slots, or a VC slot? If the decision is not easy, make note of the alternative analyses and look for any way(s) in which one analysis is simpler or more explanatory than another. Then rewrite (retranscribe) some of the words with these sounds, in accordance with the interpretations you have just made. For example, use [j] now instead of [i], or use [č] instead of [tʃ], etc.

(4) Draw syllable structure trees for some representative syllables of the language, including some that reflect interpretation decisions about ambiguous segments or sequences.

(5) Make note of any possible restrictions on the distribution of syllables within the phonological word, stating where in a word (initially, medially, finally) the allowable syllables occur, especially if different syllable types occur in different positions. For example, closed syllables occur only word-finally, or heavy syllables occur only in combination with stress, etc.

(6) Characterize the maximal syllable template for the language as a whole, e.g., [CCVC]σ_{max}.

<table>
<thead>
<tr>
<th>Binongan Itneg (Philippines)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ['makmak] ‘rice field’</td>
<td>11. [dim'lon] ‘dew’</td>
</tr>
<tr>
<td>2. ['sagoi] ‘fish net’</td>
<td>12. [be'lat] ‘gravy’</td>
</tr>
<tr>
<td>3. ['ʔabog] ‘to chase’</td>
<td>13. [la'lag] ‘man’</td>
</tr>
<tr>
<td>5. ['sa'got] ‘gift’</td>
<td>15. [ba'baʔi] ‘woman’</td>
</tr>
<tr>
<td>6. [liŋ'siu] ‘to miss one’</td>
<td>16. [tolaɡ] ‘agreement’</td>
</tr>
<tr>
<td>7. ['tanbau] ‘plow part’</td>
<td>17. [polig] ‘to roll’</td>
</tr>
<tr>
<td>8. [sa'gei] ‘woman’s name’</td>
<td>18. [dadan] ‘trail’</td>
</tr>
</tbody>
</table>
Solution and exemplary write-up of the Binongan Itneg exercise

NOTE: always follow this for syllable structure homeworks, as well as “Hints for Homework.”

(1) Inventory of canonical (unambiguous) syllable types, with examples:
- CV ['sa.got] ‘gift’ (# 5)
- CV [be'lat] ‘gravy’ (# 12)
- CVC ['ʔa.bog] ‘to chase’ (# 3)
- CVC [be'lat] ‘gravy’ (# 12)
- CVC ['sa.got] ‘gift’ (# 5)

(2) Restrictiveness of syllable positions:
It appears that the segment [ʔ] occurs only in syllable onsets, never in codas.

(3) Ambivalent sequences and segments:
[i] ~ [j] and [u] ~ [w]

The words containing ambivalent segments will be parsed using the clear cases of syllable types (CV and CVC). If the ambivalent segment fills a V slot, it will be interpreted as [i] or [u]; if the ambiguous segment fills a C slot, it will be interpreted as [j] or [w].

Examples:  (# 2) [sagoj] ‘fish net’  (# 6) [liŋsiw] ‘to miss one’  (# 10) [bolaw] ‘dirt in eye’

Note that after this decision has been made, there are now no occurrences of the vowel [u] in these data.

(4) Representative syllable structure trees:

(5) Syllable distribution restrictions:
Open syllables occur both word-initially (# 2) and word-finally (# 13). Similarly, closed syllables occur both word-initially (# 1) and word-finally (# 12). However, we can now say that every syllable begins with an onset consonant. This is the Obligatory Onset Condition. Also, medial syllables are always light, and word-final stressed syllables are always heavy.

(6) Maximal syllable template:
There is still no need for altering this from the original hypothesis of [CVC].
C, V

These cover symbols refer to the classification of segments as consonants and vowels, respectively. In terms of distinctive features, they correspond most closely to the classes defined by [–syllabic] and [+syllabic] (leaving aside the issue of syllabic consonants). They may not technically be needed in the formulation of every rule, but are often put in by convention for ease in reading the rules. Note that C is not the equivalent of [+consonantal], nor does V stand for [–consonantal].

Subscripts and superscripts

A number subscripted underneath a symbol indicates the minimum number of occurrences of that item which are necessary; similarly, a superscript indicates the maximum number of occurrences:

\[
C^3 \_ \quad \text{means at least one consonant, but not more than three consonants, i.e., C or CC or CCC.}
\]

\[
C^0 \_ \quad \text{this is often used in rules such as stress rules where the number of intervening consonants in a given position is irrelevant; this symbol simply means any number (zero or more) of consonants, with no upper limit.}
\]

\[
V \rightarrow [\text{+stress}] / \_ C^0 V C^0 #
\]

Note that the symbol \( \hat{V} \) can be used for a stressed vowel:

\[
V \rightarrow \hat{V} / \_ C^0 V C^0 #
\]

Null symbol (Ø)

This symbol, which indicates the null set or “zero,” is also used in certain phonological rules. It may appear either in the input or output position of a rule as a placeholder in cases of insertion or deletion of segments:

Insertion: \( \emptyset \rightarrow i / C \_ C # \)

Deletion: \( C \rightarrow \emptyset / V \_ # \)
Parentheses ()

Items placed within parentheses may be optionally present, but are not required to be. The use of parentheses in the environment of a rule allows us to collapse together more than one rule when these are obviously related to each other:

\[
C \rightarrow \left[ \begin{array}{c} +\text{high} \\ -\text{back} \end{array} \right] / \quad (C) \quad \left[ \begin{array}{c} +\text{high} \\ -\text{back} \end{array} \right]
\]

This rule combines the following two rules:

\[
C \rightarrow \left[ \begin{array}{c} +\text{high} \\ -\text{back} \end{array} \right] / \quad C \quad \left[ \begin{array}{c} +\text{high} \\ -\text{back} \end{array} \right]
\]

and

\[
C \rightarrow \left[ \begin{array}{c} +\text{high} \\ -\text{back} \end{array} \right] / \quad V \quad \left[ \begin{array}{c} +\text{high} \\ -\text{back} \end{array} \right]
\]

Note that parentheses can also be used, in some cases, instead of superscripted and subscripted numbers. For example,

\[
(C) = C_0^1 \\
C(C) = C_1^2 \\
C(C)(C) = C_2^3 \\
or \\
C(C(C))
\]

Curly braces {}

This device allows us to indicate an either-or choice between the items enclosed within the brace. These are only to be used when the individual items are completely distinct and cannot be collapsed together by other means:

Correct: \[ C \rightarrow \emptyset / \quad \left\{ \begin{array}{c} C \\ # \end{array} \right\} \quad \] Incorrect: \[ s \rightarrow [+\text{high}] / \quad \left\{ \begin{array}{c} i \\ e \end{array} \right\} \]
(In the incorrect rule above, the [i] and the [e] should be combined together as a natural class with a feature such as [–back].)

**Transformational rules**

Transformational rules in phonology are used to effect more than one change within a single rule. This type of rule is only used when a single process involves an effect on more than one segment. The most common examples are cases of metathesis, whereby two segments are interchanged (inverted in their linear order), and coalescence, whereby two underlying segments combine to become one phonetic segment that typically has characteristics of both of the original segments:

**Metathesis:**

\[
\begin{array}{c}
\text{C} \\
\text{[+son]}
\end{array}
\]

\[
1 \quad 2 \quad \Rightarrow \quad 2 \quad 1
\]

**Coalescence:**

\[
\begin{array}{c}
\text{V} \quad \text{C} \\
\text{[+nas]} \quad \{ \text{C} \} \quad \#
\end{array}
\]

\[
1 \quad 2 \quad 3 \quad \Rightarrow \quad 1 \quad \emptyset \quad 3
\]

In each case, the input and the environment (when necessary) are put together in a linear string. The numbers are placed underneath in order to index each element (segment and/or boundary symbol). As usual, the output of the rule only indicates the relevant changes to the string: any changes to a segment are placed below the number that indexes that segment. A change in the order of the segments (as in metathesis) is indicated by the order of the indexing numbers in the output (to the right of the arrow).

**Variables (α β)**

Consider a language which has a voicing assimilation process, whereby an obstruent takes on the same voicing as the consonant which follows it. This process could be expressed by the following two rules:

\[
\begin{array}{c}
\text{C} \\
\text{[–son]} \quad \text{[+voice]}
\end{array}
\]

\[
\Rightarrow \quad ___ \quad \text{C} \quad \text{[+voice]}
\]

\[
\begin{array}{c}
\text{C} \\
\text{[–son]} \quad \text{[–voice]}
\end{array}
\]

\[
\Rightarrow \quad ___ \quad \text{C} \quad \text{[–voice]}
\]
However, because of the similarity of these two rules, and because we would like to claim that these two cases are really just different instances of the same process, we would like to collapse the two into one rule. We do this by means of variables, symbolized by the Greek letters α, β, γ, etc. Variables of this sort may take on the usual values + or – that are associated with features. Thus we can collapse the two rules above as follows:

\[
C \rightarrow [\alpha \text{ voice}] / ___ C
[\alpha \text{ voice}]
\]

In other words, if the second consonant is [+voice], then the preceding obstruent will also be [+voice]; if the second consonant is [–voice], then the preceding obstruent is also [–voice].

A process of nasal assimilation can also make use of this sort of variable. Consider the case where a nasal consonant takes on the same point of articulation as a following consonant. Using variables with the relevant distinctive features, this could be expressed as:

\[
C \rightarrow \begin{bmatrix}
\alpha \text{ ant} \\
\beta \text{ cor} \\
\gamma \text{ high} \\
\delta \text{ low} \\
\epsilon \text{ back}
\end{bmatrix} / ___
C
\begin{bmatrix}
\alpha \text{ ant} \\
\beta \text{ cor} \\
\gamma \text{ high} \\
\delta \text{ low} \\
\epsilon \text{ back}
\end{bmatrix}
\]

This rule merely says that for the point of articulation features (anterior, coronal, high, low, and back), the nasal will take on the same values as a following consonant has for these same features. The following is a shorthand notation for the same rule:

\[
C \rightarrow [\alpha \text{ place}] / ___ C
[\alpha \text{ place}]
\]

In this rule, the “feature” [place] is simply an abbreviation for the five features that refer to the point of articulation. Keep in mind that technically the rule with [α place] is really not simpler than the previous rule with five features; this notation is handy, however, and is often used to express this generalization in abbreviated fashion.

Another similar case is degemination, which can be expressed by a rule such as the following:

\[
C \rightarrow \emptyset / ___ C
[\alpha \text{ FT}] [\alpha \text{ FT}]
\]
Here the [FT] is an abbreviation for *all* the features used to describe a segment. The variable indicates that all features of the two consonants in the rule must have the same values, i.e., that the two are identical in every way. In such a case, we are told to delete the first one in our rule — in other words, to degeminate.

One caution should be made about using variable notation. Since the Greek letters are used for cross-referencing purposes (to say that the value of one feature is the *same* as something else), each variable must occur at least twice in a rule.

**Boundaries**

Sometimes a rule applies only if the segment occurs in a certain position in the syllable, word, or utterance. In early versions of generative phonology it was not considered proper to make reference to the notion of syllable. This is quite the opposite of most current theories. Reference to the edge of the syllable might be made as follows:

\[
A \rightarrow B / \_ \_ \]_\sigma
\]

where ‘\(\sigma\)’ = syllable and ‘\(]\)’ = right edge.

It has also been argued that phonological rules may make reference to the internal structure of the syllable. The following rule says that A becomes B if it is in the rhyme of the syllable:

\[
A \rightarrow B / \_ \_ \\
\]

The version of generative phonology that dominated the field for many years also explicitly allowed reference to morphological boundaries in rules. These boundaries are symbolized as follows:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>morpheme boundary</td>
</tr>
<tr>
<td>#</td>
<td>single (interior) word boundary</td>
</tr>
<tr>
<td>##</td>
<td>double (exterior) word boundary</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rules such as the following were commonly seen in the literature of the classical period:

\[
A \rightarrow B / \_ \_ + Z \\
C \rightarrow D / X \_ \_ #
\]
The power of such a theory was severely curtailed in some versions of lexical phonology, however. For example, Mohanan (1986) proposed that reference to such boundaries should be considered illicit, and that the only morphological boundary that a phonological rule may refer to is the bracketing itself. In his theory, each morpheme comes from the lexicon with bracketing of its own and the word formation rules provide more brackets, as shown below (using English spelling rather than phonological representations):

\[
[ [ \text{un} ] \text{[happy]} ]_{\text{Adj}} [\text{ness}] ]_{\text{Noun}}
\]

Therefore rules such as the ones given immediately above are not possible in this theory, whereas rules such as the following are allowed (where ‘]’ indicates the end of a morpheme and ‘[’ the beginning of a morpheme):

\[
E \rightarrow F / X \_ \_ ]
\]

\[
G \rightarrow H / \_ \_ ] [ J
\]

The original source for this handout is unknown.
Spanish consonant clusters

Most dialects of Latin American Spanish exhibit the following inventory of consonant phonemes:

\[
\begin{array}{cccc}
p & t & ċ & k \\
b & d & g & \\
f & s & x/h & \\
m & n & ñ & \\
l & r & \\
w & y & \\
\end{array}
\]

The following consonant clusters are attested word-initially (\(r\) is really a flap):

- pr: primero ‘first’
- pl: playa ‘beach’
- tr: tres ‘three’
- kr: krema ‘cream’
- kl: klase ‘class’
- br: braso ‘arm’
- bl: blanko ‘white’
- dr: drama ‘drama’
- gr: gris ‘grey’
- gl: glosa ‘gloss’
- fr: fresa ‘strawberry’
- fl: flako ‘skinny’

What generalizations can you state concerning which consonant clusters are permitted syllable-initially? Show how the following words would be syllabified. Assume that the maximal syllable template is [CCVCC].

1. golondrina ‘swallow (bird)’
2. estreya ‘star’
3. transkribir ‘to transcribe’
4. inspirar ‘to inspire’
5. aktor ‘actor’
6. ablan ‘they speak’
7. akonteser ‘to happen’
8. soporta ‘he/she puts up with’
9. aktitud ‘attitude’
10. respuesta ‘answer’
Isthmus Zapotec (Mexico)

1. [ˈŋaːga] ‘rat’
2. [ˈgi] ‘fire’
3. [ʃneza] ‘correct’
4. [sti] ‘other’
5. [ŋgola] ‘big’
6. [dʒia] ‘griddle’
7. [stʊi] ‘sickness from embarrassment’
8. [na'nuŋi] ‘fine’
9. [ˈruna] ‘obeys’
10. [u'la] ‘pushed (rep.)’
11. [ri'bidʒi] ‘calls’
12. [ri'tʃesa] ‘jumps’
13. [i'la] ‘will push’
14. [u'na] ‘had sight’
15. [ˈranда] ‘able’
16. [ʃkesa'do] ‘fontanel’
17. [ru'tinde] ‘makes fight’

Source of data: Velma B. Pickett.
Homework for Day 18:
read Schane chapter 6

Day 17: The big picture

In today’s class we are going to list and explain the goals of morphophonemic analysis, and then illustrate the procedural steps we can use to analyze morphophonemic data.

• Here’s another similar kind of situation from Michif:
  • [diminikwan] ‘I drink’
  • [tininikwan] ‘I drank’
• In spite of this minimal pair, there is no need to posit a phonemic contrast between /n/ and /d/ in this language.
• As a matter of fact, all of the obstruents in verbal forms in this language are voiceless at the underlying level.
• The phonetic [d] here is the surface realization of an underlying sequence /n - t/, where /n-/ is a prefix (probably a tense or aspect marker). These two consonants fuse or merge together into the intermediate segment [d], which preserves features of both. This is called _____________________________________________________.
• When morphemes are put together in a word, we often find phonological effects that modify the segments near the borders, and sometimes the position of the morpheme boundary itself is blurred. For example, in fast speech the question Did you eat? can be reduced to a single word: [dʒi:t].

• All that contrast necessarily tells us is that the two words have different underlying forms, not necessarily different phonemic contrasts. Phonetically, there will always be a compromise between what you want to say and how you can say it. For example, consider the following pair of words from Hueyapan Nahuatl:
  • [nikoa] ‘I buy it’ vs. [niyoa] ‘I shell it’
  • /nikka/ → [nikoa] (degemination)
  • /nikoa/ → [niyoa] (lenition)
• All that contrast necessarily tells us is that the two words have different underlying forms. But the difference between them phonemically is not always how they differ phonetically.

• Consequently, when we do the basic phonemic analysis of a language (as we have been doing up to this point), it is best to work with forms that are not likely to be morphologically complex. In particular, nouns are usually less likely than verbs to contain more than one morpheme.
• As a preliminary to doing morphophonemic analysis, we start with the assumption that phonologically predictable allomorphy is to be attributed to a single underlying form for each morpheme, plus a set of simple, natural rules which convert these underlying forms to their correct surface (phonetic) representations due to phonological conditioning in the relevant environments.
• In some such cases, these phonological processes might certain distinctions which are otherwise phonemic in the language in question. For example, in English /s/ and /z/ are contrastive phonemes word-initially in forms such as Sue vs. zoo. Nevertheless, the plural suffix on nouns is predictably pronounced as either [-s] or [-z] depending on the voicing of the consonant at the end of the root: cats vs. dogs.

• So there are two specific goals to keep in mind when doing morphophonemic analysis:
  • (1) Try to posit one single and unique underlying form for each morpheme encountered in the data (from which all of its surface allomorphs can be derived).
  • (2) Posit a set of one or more regular, productive, and natural phonological rules for the language as a whole which generate allomorphs can be derived).
• All such phonological processes can be categorized into one of four major types:

  • (1) ________________________: a sound that is present underlyingly does not surface phonetically when the word is actually pronounced. For example, family → fam ɪɭ, compared with familiar. Vowel deletion between two consonants (like in this case) is called syncope.
  • (2) ________________________: a sound that is present underlyingly does not surface phonetically when the word is actually pronounced. For example, comfort [fɒrɪ] vs. comfortable [fɒrəbəl]. This is the least common of the four types of morphophonemic processes.

  • (3) ________________________: a sound that is present underlyingly does not surface phonetically when the word is actually pronounced. For example, comfort [fɒrɪ] vs. comfortable [fɒrəbəl]. This is the least common of the four types of morphophonemic processes.

  • (4) ________________________: a sound that is present underlyingly does not surface phonetically when the word is actually pronounced. For example, comfort [fɒrɪ] vs. comfortable [fɒrəbəl]. This is the least common of the four types of morphophonemic processes.
Cases of coalescence (or merger or fusion) can usually be broken down into two interacting steps: assimilation + deletion, such as Michif 'n - v → [d] above.

Lithuanian (page 57)

1. iljip-u iljib-dava-u ‘get in’
2. djirb-u djirb-dava-u ‘work’
3. suk-u sug-dava-u ‘turn’
4. vjeʒ-u vjeʒ-dava-u ‘transport’
5. injeʃ-u injeʃ-dava-u ‘bring in’

[p] ~ [b] iljip u iljib dava u
djurbu djirb dava u
[k] ~ [g] suk u sug dava u
vjeʒ u vjeʒ dava u
[ʃ] ~ [ʒ] injeʃ u injeʃ dava u

[p k j] ~ [b g ʒ]

C → [+voice] / ___ C

[+voice]
Procedures for morphophonemic analysis

1. Arrange the data into logical paradigms.
2. Make tentative morpheme cuts.
3. Identify those morphemes which exhibit more than one allomorph.
4. Identify the phonological differences between various allomorphs.
5. Group the alternating segments into logical natural classes.
6. Consider the phonological alternations from “both directions.”
7. Analyze the data to see which hypothesis yields the correct predictions.
8. Posit one unique underlying form for each morpheme.
9. Write formal rules which will derive all of the surface allomorphs.
10. Determine in what order the rules must apply to produce the correct results in all cases.
11. Write home to mom about how much you are enjoying linguistics.
Chamicuro morphophonemic data

possessed nouns (the genitive construction)

focus on the prefixes

<table>
<thead>
<tr>
<th>my</th>
<th>his/her</th>
<th>our</th>
<th>your (sing.)</th>
<th>unpossessed</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>üpekáno</td>
<td>ipekáno</td>
<td>âpekáno</td>
<td>pekáno</td>
<td>pekáno</td>
<td>flesh, meat</td>
</tr>
<tr>
<td>üto?sóna</td>
<td>ito?sóna</td>
<td>âto?sóna</td>
<td>to?sóna</td>
<td>to?sóna</td>
<td>shin</td>
</tr>
<tr>
<td>ukúhtu</td>
<td>ikúhtu</td>
<td>akúhtu</td>
<td>kúhtu</td>
<td>kúhtu</td>
<td>foot</td>
</tr>
<tr>
<td>učáno</td>
<td>ičáno</td>
<td>ačáno</td>
<td>čáno</td>
<td>čáno</td>
<td>neck</td>
</tr>
<tr>
<td>úsakódlá</td>
<td>isakódlá</td>
<td>âsakódlá</td>
<td>sakódlá</td>
<td>sakódlá</td>
<td>saliva</td>
</tr>
<tr>
<td>uménu</td>
<td>iménu</td>
<td>aménu</td>
<td>ménu</td>
<td>ménu</td>
<td>tongue</td>
</tr>
<tr>
<td>wawána</td>
<td>yawána</td>
<td>âwawána</td>
<td>pawána</td>
<td>awána</td>
<td>mouth</td>
</tr>
<tr>
<td>wènopíne</td>
<td>yènopíne</td>
<td>awènopíne</td>
<td>pènopíne</td>
<td>enópi</td>
<td>lemon</td>
</tr>
<tr>
<td>wí:la</td>
<td>yí:la</td>
<td>awí:la</td>
<td>pí:la</td>
<td>í:la</td>
<td>blood</td>
</tr>
<tr>
<td>wohpána</td>
<td>yohpána</td>
<td>âwohpána</td>
<td>pohpána</td>
<td>ohpána</td>
<td>liver</td>
</tr>
<tr>
<td>wústawádle</td>
<td>yústawádle</td>
<td>awústawádle</td>
<td>pústawádle</td>
<td>ûstawádli</td>
<td>sleeping mat</td>
</tr>
</tbody>
</table>

sources of data:


CHECK YOURSELF 2

True  False  The demonstration of contrast between two sounds establishes them as allophones of the same phoneme.

T  F  If you cannot demonstrate that the members of a suspicious pair contrast, you should look for complementary distribution.

T  F  When we encounter a case of unconditioned free variation, we should write an optional rule.

T  F  The phonemic inventory of a language may tend toward symmetry even though the phonetic inventory may not.

T  F  The pair of words [windala] and [wiŋgaɾu] constitutes good evidence that the sounds [n] and [ŋ] contrast in analogous environments (assuming that the words mean different things).

T  F  In some languages the syllable nucleus is optional.

Circle the sequences which are clearly ambivalent (ambiguous):

ft  nd  dʒ  pt  oe

ts  mk  kj  ia  th

Which of the following pairs of sounds are clearly suspicious and should therefore be analyzed to determine their phonological relationship?

i / iː  t / n  u / u  l / r

t / tʃ  e / ɛ  d / ɡ  s / ʃ
When dividing a syllable into its smallest component parts, what are the three main divisions?

Draw a syllable structure tree of the word [stalk], including all relevant nodes.

What is the universally preferred syllable type?

The following terms are used in the phonological analysis of suspicious pairs:

(a) Contrast        (b) Complementary Distribution        (c) Free Variation

Indicate which of these terms are described in the following situations by placing the appropriate letter(s) in the blanks (more than one letter may apply to a given situation):

__________ A suspicious pair of sounds may occur in an identical environment.

__________ May be conditioned or unconditioned.

__________ The two sounds in question are distinguished as separate phonemes.

__________ We write a rule to describe the predictable patterns.

__________ Phonetic differences are allophonic, not phonemic.

__________ The environments in which the two sounds occur are mutually exclusive.

__________ The phonetic differences alone can make a minimal difference in the meaning of a word.
AL 4303, Principles of Phonological Analysis
GIAL
Day 19

Your homework for Day 20 consists of the following two exercises (to hand in):

1. Xavante (p. 101 in the Marlett textbook)

What is the underlying representation of the ‘your’ prefix? State (in prose) the generalization which accounts for its distinct allomorphs. Do not worry about formalizing the rule (ignore what the book says at this point). Note that consonants produced with the tip (apex) or front part (blade) of the tongue as primary articulators are known as coronals. This includes the following points of articulation: interdental, dental, alveolar, and alveopalatal (the latter two can also be retroflexed). This natural class of sounds can be uniquely defined as [+coronal].

2. Tetelcingo Nahuatl, p. 165

Follow the instructions on the “Procedures for syllable structure exercises” handout. List the canonical syllable types, maximal syllable template, ambiguous segments and sequences, interpretation decisions (with reasons / justification), etc. Note that the verb *parse* in this context means “to divide into syllables.” The lateral [ɬ] in #2 and 4 is presumably different from the [ɬ] in #11, 13, 14, 17, and 19. Also, continue to follow as well the instructions on the “Hints for Homework” handout for this and all other homeworks, unless specified to the contrary.
Day 20: The Big Picture

In today’s lesson we are going to discuss the theory of formal distinctive features, define many of them, and explain their hierarchical structure in terms of an exhaustive feature geometry tree.

Motivations for using distinctive markers:

1. They are the smallest building blocks of speech sounds – features allow us to describe the characteristics of each sound in terms of their phonetic (articulatory) properties.
2. Features allow us to express the relationships between sounds and give a notion of more or less similar, e.g., [s] and [z] differ only in the feature [± voice].
3. They distinguish contrasts in languages in the most minimal way – our features should be able to capture all the phonemic contrasts in the languages of the world.
4. They describe natural classes in an accessible and revealing way. Thus, sounds which are phonetically similar tend to behave in similar, parallel ways in phonological processes.
5. In writing formal rules, using features points out that many processes do not involve an interchange of one segment for another segment, but really a change in features. (Although of course the change of a feature will have the indirect effect of causing a change in a segment in specific cases.) In other words, using features often allows us to express formal rules and processes in a more general way, which usually reflects what is happening phonetically.
6. Features serve as an evaluation metric in determining the simplicity and naturalness of phonological processes. In general, the more simple and natural a process is, the fewer the number of features will be needed to formalize it.
7. Here are some additional characteristics of the features we will be using:
   - a. They are finite in number (finite).
   - b. They are claimed to be universal, i.e., the same in all languages.
   - c. They are binary (in most cases).
   - d. They are articulatorily defined.

Review the tree structure as we have it so far:

- Note that the arrangement of these features is not random; it is intended to bring out certain interdependencies between the features and nodes. These group hierarchically in bundles and they behave that way in phonological rules. Thus features which are grouped together under a common node will typically undergo a unitary process and/or trigger a process in the same way.
- There have been many different proposals for the feature geometry tree, but the one in the Marlett textbook is the model we will follow in this course.
- Explain, discuss, and illustrate the features, relying on the diagram on p. 230 of the textbook. (Note three additional features.)
## Distinctive Features (with Marlett’s feature geometry tree)

|     | p  | p’ | b  | t  | d  | k  | g  | q  | f  | v  | θ  | δ  | s  | z  | j  | č  | m  | n  | η  | l  | r  | j  | w  | h  | h’ | i  | e  | æ  | ø  | ø’ | ö  | C |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| C or V | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | V | V | V | V | V | V | V | V |
| voice | – | – | – | + | – | – | + | – | + | – | + | – | + | + | + | + | + | + | + | – | – | + | + | + | + | + | + | + | + | + | + |
| spread glottis | – | + | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | – | – | – | – | – | – | – | – | – | – |
| constricted glottis | – | – | + | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | – | – | – | – | – | – | – | – | – | – | – |
| sonorant | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | + | + | + | ± | ± | + | + | + | + | + | + | + |
| consonantal | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | ± | + | + | + | + | + | + | + | + | + |
| continuant | – | – | – | – | – | – | – | – | + | + | + | + | + | + | – | – | – | ± | + | + | + | + | + | + | + | + | + | + | + | + | + |
| nasal | – | – | – | – | – | – | – | – | + | + | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| LABIAL | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • |
| round | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | – | – | – | – | – | – | – | – | – | – | – | – |
| CORONAL | – | – | – | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • |
| lateral | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | – | – | – | – | – | – | – | – | – | – | – |
| anterior | + | + | + | + | + | – | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| strident | – | – | – | – | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| distributed | – | – | + | + | – | + | + | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| DORSAL | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • |
| high | + | + | – | – | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Low | – | – | – | – | – | – | – | – | – | – | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| back | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| TONGUE ROOT | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • |
| atr | – | + | – | – | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| rtr | – | – | – | – | – | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
Feature Geometry
Adapted from “An Introduction to Phonological Analysis” (Marlett 2001, p.230)
Practice with distinctive features and natural classes (in class)

(A)

1. Circle the segments which are [+anterior]: p ð z ʃ y ɬ ɡ i f j
2. Circle the segments which are [+strident]: ɸ ʃ θ s t v γ ɿ h z
3. Circle the segments which are [+back]: m k i ɬ w ɡ dʒ a r y
4. Circle the segments with a Coronal Node: p dʒ d k ɿ h s f e j
5. Circle the segments which are [+sonorant]: n w b r h u v l tʃ β
6. Circle the segments which are [+lateral]: l r ɬ w x ɬ ɿ dl m ʃ
7. Circle the segments which are [+continuant]: f tʃ ʃ tʃ tʃ z o m l pʰ r
8. Circle the segments which are [+round]: i y pʃ o m s ɿ kʷ ø a
9. Circle the segments which are [+low]: e a u h s q i ɚ æ x
10. Circle the segments which are [+back]: i u e o s a ʃ k w q
11. Circle the segments which are [+ATR]: a e ɿ i ɿ o p u ɛ r

(B) Which distinctive features differentiate the sounds in each of the following pairs?

example: p b [p] is [–voice]; [b] is [+voice]

1. i ɿ
2. a ɿ
3. j ɿ
4. u ɿ
5. i y
6. i ɿ
7. ð ɿ
(C) In each consonant and/or vowel system below, some of the segments are circled. Determine whether these constitute a natural class within that system. If they do, give the shared features and/or nodes which uniquely define them as a natural class.
sources:


Homework on distinctive features (for practice only; answers will be posted)

On a separate sheet of paper, draw trees which:

include these sounds but exclude these sounds

eample:

\[ \text{C} \]

\[ \text{Root} \]

\[ \text{Laryngeal} \quad \text{Supralaryngeal} \]

\[ [+\text{voice}] \quad [+\text{continuant}] \]

1. o i e a j w m p f g r

2. t^s m p t j l s č (=ťʃ) a u o e i

3. m n ɲ ŋ ā b s ɣ l (where [ā] is [+nasal])

4. k ɡ x ɣ w p d ʃ j β
Homework on feature geometry trees to hand in on Day 22

\[
\begin{align*}
p & \quad t & \quad t^\prime & \quad k & \quad q \\
b & \quad d & \quad d^3 & \quad g & \quad G \\
\phi & \quad f & \quad s & \quad \int & \quad x \\
\beta & \quad v & \quad z & \quad \gamma \\
m & \quad n & \quad \eta \\
l & \quad j & \quad w \\
\end{align*}
\]

Assuming the consonant inventory given above, draw the appropriate feature trees which uniquely and exhaustively characterize the following natural classes of sounds. There may be more than one set of features which would work for a given set of sounds. Try to use the fewest features and nodes possible for each tree, excluding all other sounds in the inventory.

1. \( n \ \eta \ \lambda \ m \)
2. \( k \ \phi \ x \ \gamma \)
3. \( d \ z \ n \ l \)
4. \( t^\prime \ d^3 \ \int \ \zeta \)

Indicate any ONE feature value or node shared by all members of each of the following sets of phones (natural classes):

Example: \( \tilde{r} \ n \ m \ a \) \ [+sonorant]

1. \( t \ d \ \int \ l \ n \)
2. \( e \ a \ o \ \alpha \)
3. \( i \ g \ w \)
4. \( o \ u \ \alpha \)
5. \( k \ g \ a \ u \)
6. \( p \ b \ f \ m \)
7. \( p \ m \ \tilde{t} \ w \ \tilde{k} \ w \)
8. \( j \ w \ h \ ? \)

Circle the sounds which are [+anterior]: \( p \ \tilde{d} \ z \ \int \ l \ g \ i \ f \ j \)

Circle the sounds which are [+strident]: \( \phi \ \int \ \theta \ s \ t \ \gamma \ v \ l \ h \ z \)

Circle the sounds which are [+back]: \( m \ k \ i \ \partial \ w \ g \ d^3 \ a \ r \ u \)
Homework for Day 23:
• read Marlett, chapters 13 and 34
• read Schane, pp. 84-91 (on reserve)
• read Rule Ordering paper (in the course packet)
• Czech problem (on a separate handout in the course packet, to turn in)
• do the exercises in Marlett, pp. 76-77 (for practice only)

Day 22: The Big Picture

In today’s class we are going to look at how phonological rules interact with each other when more than one process applies to the same underlying form, and the constraints which govern derivations of this sort.

Voicing: p → b / N ___

C → [+voice] / C ___

R  R  L

C  C  R  R

Rule ordering

How should rules be allowed to apply when more than one is needed for the same word?

Two possibilities, among others, have been proposed throughout the years of modern phonological theory:

Hypothesis A: all rules apply directly to the Underlying Representation (UR), to produce the surface form in one step.

The Pastaza Quechua word above is one case in which this would work: /wasi-n-pa/ → [wasimba].

However, this hypothesis does not work for Brazilian Portuguese:

/onde/ → [ondi] ‘where’

Palatalization: d → dz / ___ i

Final Raising: e, o → i, u / ___ #

These facts motivate the following alternative proposal:

Hypothesis B: rules apply in a fixed order (which often must be critically specified or stipulated on a language by language basis).

By convention we always list the rules in a particular set order and then apply them consistently in that fashion, even though in many cases the specific order in which the rules apply doesn’t make any difference in many or most words of the language (e.g., in the Pastaza Quechua example above).

Sometimes, however, the correct fixed order is necessary in deriving the right surface form of some words.

In such cases we will demonstrate that this is necessarily the case, and then make an explicit ordering statement about this case, in addition to stating the overall order that the rules are to be applied in.

Certain aspects of this order might be determined in part by general principles of phonological theory (Universal Grammar or UG).
A ____________ then can be defined as an orderly display of forms showing the successive application of the phonological rules, step by step, from the Underlying Representation to its corresponding Phonetic Representation (PR).

Let’s illustrate this with the word /onde/.

<table>
<thead>
<tr>
<th>(correct order)</th>
<th>(incorrect order)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underlying Representation</td>
<td>/onde/</td>
</tr>
<tr>
<td>Final Raising</td>
<td>ondi</td>
</tr>
<tr>
<td>Palatalization</td>
<td>ondį</td>
</tr>
<tr>
<td>Phonetic Representation</td>
<td>[ondį]</td>
</tr>
</tbody>
</table>

• Always show the UR and the PR separately, within the appropriate diagonal slashes and square phonetic brackets, respectively.
• List the rules one-by-one off to the side, and show the effects of applying each rule in turn.
• Crucial ordering is indicated in the overall list with an arc connecting the names of the two rules that are critically ordered.

• When a crucial ordering is ____________ for some form to be derived correctly, demonstrate this by giving the derivation of an incorrect PR, where the rules apply in the opposite (wrong) order.
• Then show that the order you are positing is sufficient to yield all of the correct surface forms. That is, show that by applying the rules in the right order, you do in fact yield (end up with) the correct results for critical forms.
• These wrong derivations then constitute the necessary evidence for crucial ordering statements between pairs of rules.
• Otherwise, all other things being equal, a solution which requires a critical ordering restriction between two rules is more complicated, less favored, and less simple than one which does not.
• Likewise, all else being equal, a solution which makes no reference to grammatical information (such as morpheme boundaries) is preferable to a solution which requires them in some rules.

• However, since in principle we must allow them (there’s plenty of evidence that things cannot be properly accounted for without such information in some cases), we need not avoid morpheme boundaries entirely. This is especially the case when some overall simplification of the analysis is effected when they are invoked. For example, using morpheme boundaries sometimes allows us to keep rules and processes uncluttered and state them in the simplest way possible. Or they may simplify the inventory of underlying segments, or make these more symmetrical, etc.
• Another case in which this type of critical ordering is necessary is Isthmus Nahuatl:
  • /sikalili/ → [sikalil] ‘put it in it’
  • Sonorant consonants devoice word-finally:  l → ŵ / ___ #
  • Short vowels delete word-finally:  i → Ø / ___ #
• This latter type of process is called ____________________________.

• When a crucial ordering is ____________ for some form to be derived correctly, demonstrate this by giving the derivation of an incorrect PR, where the rules apply in the opposite (wrong) order.
• Then show that the order you are positing is sufficient to yield all of the correct surface forms. That is, show that by applying the rules in the right order, you do in fact yield (end up with) the correct results for critical forms.
• These wrong derivations then constitute the necessary evidence for crucial ordering statements between pairs of rules.
• Otherwise, all other things being equal, a solution which requires a critical ordering restriction between two rules is more complicated, less favored, and less simple than one which does not.
• Likewise, all else being equal, a solution which makes no reference to grammatical information (such as morpheme boundaries) is preferable to a solution which requires them in some rules.

• However, since in principle we must allow them (there’s plenty of evidence that things cannot be properly accounted for without such information in some cases), we need not avoid morpheme boundaries entirely. This is especially the case when some overall simplification of the analysis is effected when they are invoked. For example, using morpheme boundaries sometimes allows us to keep rules and processes uncluttered and state them in the simplest way possible. Or they may simplify the inventory of underlying segments, or make these more symmetrical, etc.
• Another case in which this type of critical ordering is necessary is Isthmus Nahuatl:
  • /sikalili/ → [sikalil] ‘put it in it’
  • Sonorant consonants devoice word-finally:  l → ŵ / ___ #
  • Short vowels delete word-finally:  i → Ø / ___ #
• This latter type of process is called ____________________________.
In Czech, the word which means ‘brain’ (in the accusative case) has the underlying form /mozk/. However, it is actually pronounced as [mozek]. In the dative case, the underlying form of ‘brain’ is /mozk-u/, where /-u/ is the dative case suffix. This word is pronounced phonetically as [mosku]. In order to account for the differences between these underlying forms and their corresponding surface forms, we need to posit two rules. The first rule might be called “e Insertion” or “e Epenthesis”; it says (in prose) that when two consonants occur together at the end of a word, insert the vowel [e] between them. In the type of formalism which was characteristic of classical generative phonology, this rule would be expressed as follows:

(1) e Insertion:  \( \emptyset \rightarrow e / C \_\_\_ C \# \)

or more precisely,  \( \emptyset \rightarrow V / C \_\_\_ C \# \)

\[
\begin{bmatrix}
-\text{back} \\
-\text{high}
\end{bmatrix}
\]

However, according to more recent views of phonology (which we will prefer to follow in this course), this process would be handled somewhat differently, making reference to syllable structure (did you possibly think that we would let you forget what you now know about syllables?). Apparently in Czech, no word (and hence no syllable) may end with two consonants in the final coda (at least among the canonical cases in the data which are available to us), and this restriction would be seen as the motivation for inserting the vowel [e] in certain words.

The other process which can be observed at work in the above forms is a rule which devoices obstruents when they occur before a voiceless consonant:

(2) Devoicing:  \( C \rightarrow [\sim\text{voice}] / \_ \_\_ C \)

\[
\begin{bmatrix}
-\text{sonorant}
\end{bmatrix}
\]

\[
\begin{bmatrix}
-\text{voice}
\end{bmatrix}
\]

In terms of the hierarchical view of feature organization which we discussed recently, this process would be formalized as follows:
Devoicing:

What this rule says is that the Laryngeal Node of any consonant, as well as all of the features it dominates (including [voice]), becomes associated with or attached to the Root Node of a preceding obstruent (as indicated by the feature [–sonorant]). In the coming days you will have more opportunities to become familiar with this newer type of notation. It has several advantages when compared to the style of formal rules which we have written up to this point, and we will explain some of them to you in an upcoming lecture. If we used abbreviations instead of the respective nodes, the rule would be stated as follows:

Devoicing:

Your mission (if you choose to accept it) is to determine in what order the two rules (e Insertion and Devoicing) must be applied in order to correctly derive the two surface forms mentioned above. Assuming that the two rules apply in a specific sequential order, which of these two rules needs to apply first? Demonstrate that your ordering choice is correct by giving two derivations of each of the two words above. In the first derivation, apply the two rules in the correct order and show that they generate the right results (surface forms). Then give a second derivation in which you reverse the order of the two rules, yielding an incorrect phonetic representation in at least one of the cases. Be sure to mark incorrect (ungrammatical) surface forms with an *. In presenting your derivations, follow the procedures outlined in the two reading assignments which you (hopefully) have just completed. This exercise will be collected on Day 23. Please note that, for this homework, you can ignore the other instructions in the book for this exercise.
Some practical tips on writing formal rules using feature geometry trees

On paper, it may appear that feature geometry rules are more complex and cumbersome than rules formalized with the classical notation. In a sense this is only an illusion. There are some very basic and important distinctions between the two models. We will focus on these theoretical issues in an upcoming lecture and explain why feature geometry is a much better approach in the long run. For the moment we highlight here some of the mechanical differences that are helpful to keep in mind when working with feature geometry trees.

In the first place, phonological processes of assimilation are preferably expressed by spreading a feature or node which is already present in the environment triggering the change. Consider a rule whereby an obstruent becomes voiced immediately preceding a voiced consonant:

\[(1) \quad \text{[–sonorant]} \rightarrow \text{[+voice]} \]

In our new model this would be expressed as follows:

\[(2) \quad \text{C} \quad \text{C} \]
\[\text{Root} \quad \text{Root} \]
\[\text{Supralaryngeal} \quad \text{Laryngeal} \quad \text{Laryngeal} \]
\[\text{[–sonorant]} \quad \text{[+voice]} \quad \text{[–sonorant]} \quad \text{[+voice]} \]

Spreading is visually depicted by a dotted association line. This indicates the addition of a new association line in exactly that location. For example, if rule (2) above were broken down into a “before” and an “after” phase, it would look like this:

\[(3) \quad \text{a) before:} \quad \text{b) after:} \]
\[\text{C} \quad \text{C} \]
\[\text{Root} \quad \text{Root} \]
\[\text{Supralaryngeal} \quad \text{Laryngeal} \quad \text{Laryngeal} \]
\[\text{[–sonorant]} \quad \text{[+voice]} \quad \text{[–sonorant]} \quad \text{[+voice]} \]

This makes it obvious why it is simpler to use a dotted association line as in (2), saving us a lot of time and space. Spreading can go from right-to-left (as in (2)) or from left-to-right, but it always goes upwards, never downwards. It may be helpful (although not necessary) to draw in a
pointed arrow at the top of the dotted association line to show the direction of spreading, as follows:

(4) (a) correct: (b) incorrect:

\[
\begin{align*}
\text{C} & \quad \text{Root} \\
\text{Supralaryngeal} & \quad \text{Laryngeal} \\
\quad & \quad [-\text{sonorant}] \\
\quad & \quad [+\text{voice}] \\
\end{align*}
\]

When connecting a new association line via spreading, make sure that the top of it (i.e., the arrow tip, if this is included) comes \textit{underneath} the relevant node in the tree (as in (4a)), not to the side or above the node:

(5) (a) incorrect: (b) incorrect:

\[
\begin{align*}
\text{C} & \quad \text{Root} \\
\text{Supralaryngeal} & \quad \text{Laryngeal} \\
\quad & \quad [-\text{sonorant}] \\
\quad & \quad [+\text{voice}] \\
\end{align*}
\]

Another detail to keep in mind is that the same exact node should dominate the spreading feature in both trees (the target and the trigger of the rule):

(6) (a) correct: (b) incorrect:

\[
\begin{align*}
\text{C} & \quad \text{Root} \\
\text{Supralaryngeal} & \quad \text{Laryngeal} \\
\quad & \quad [-\text{sonorant}] \\
\quad & \quad [+\text{voice}] \\
\end{align*}
\]

At times certain tree diagrams are simplified by leaving out intervening nodes and connecting the spreading feature or node directly to the top of the tree. This is an informal shortcut, sort of like using [$\alpha$ \textit{Place}] instead of [$\pm$\textit{coronal}, [$\pm$\textit{anterior}, [$\pm$\textit{distributed}] etc. in the classical model:
While it is okay to do this in course handouts and similar materials, technically it is an imprecise practice that should be avoided in more serious works such as papers submitted for publication, textbooks, etc. In general you should not take the liberty of doing this in homeworks, tests, etc., unless instructed to the contrary.

Association lines in feature trees are always subject to two important constraints. In the first place, they should never cross over each other. Plan ahead in order to avoid this problem:

Rule (8a) above says that a low vowel becomes mid when the preceding syllable contains a mid vowel.

Secondly, association lines must be straight, never curved. Otherwise it would be possible to undermine and get around the prohibition against crossing lines, in effect.

When two or more binary terminal features are dominated by the same node in the tree, each one should be placed in its own unique matrix (square brackets) with a separate association line, not stacked on top of each other using a single line:
(9)  (a) correct:  (b) incorrect:

\[
\begin{align*}
&V \\
&\quad \text{Root} \\
&\quad \text{Supralaryngeal} \\
&\quad \quad \text{Place} \\
&\quad \quad \quad \text{Dorsal} \\
&\quad \quad \quad \quad [+\text{high}] \quad [-\text{back}] \\
&V \\
&\quad \text{Root} \\
&\quad \text{Supralaryngeal} \\
&\quad \quad \text{Place} \\
&\quad \quad \quad \text{Dorsal} \\
&\quad \quad \quad \quad [+\text{high}] \quad [-\text{back}] \\
\end{align*}
\]

The tree in (9a) describes a segment like [i].

To formalize a deletion rule in the feature geometry model, draw a circle around the timing unit (C or V or X) at the top of the tree and then write in the arrow and null symbol. It normally does not matter which direction the arrow points (up, down, left, right); do it however is most convenient and uncluttered with respect to the rest of your rule:

(10) \( \emptyset \)

\[
\begin{align*}
&\quad \text{C} \\
&\quad \quad \text{Root} \\
&\quad \quad \quad \text{Supralaryngeal} \\
&\quad \quad \quad \quad \text{Place} \\
&\quad \quad \quad \quad \quad \text{Coronal} \\
&\quad \quad \quad \quad \quad \quad [+\text{lateral}] \\
&\quad \text{C} \\
&\quad \quad \text{Root} \\
&\quad \text{Supralaryngeal} \\
&\quad \quad \text{Laryngeal} \\
&\quad \quad \quad [+\text{spread glottis}] \\
&\quad \quad \quad \quad \text{Coronal} \\
&\quad \quad \quad \quad \quad \quad [+\text{lateral}] \\
\end{align*}
\]

Rule (10) says to delete a lateral consonant (like [l]) which is immediately followed by an aspirated consonant (or [h]).

Another way in which the deletion of individual features and/or nodes is sometimes depicted is by delinking. This refers to the deletion of an association line and is graphically indicated by drawing two short, parallel lines in the middle of a regular association line, at right angles to it:
The rule in (11) above says to delink (delete) the association line connecting a syllable-final consonant with its Laryngeal Node. This would have the effect of devoicing an obstruent in that position, for example, assuming that “floating” (unattached) features and nodes are simply eliminated at the end of the derivation since they cannot be phonetically implemented (realized). In such a case there would be a default redundancy rule such as [–sonorant] $\rightarrow$ [–voice] that would fill in the correct surface values of the relevant features. When a node is delinked, all of the other nodes and features which it dominates (implicitly or explicitly) are also detached.

When a feature spreads from one tree to another segment in its environment, you might be worried that this is incompatible with another feature (or the inverse of the same feature) which may already be specified there. To illustrate, let us reconsider rule (8a) above. What if the target which is receiving the feature [–low] (the vowel on the right) is prespecified for [+low], such as an /a/? In such a case we might go ahead and add the feature [+low] in the structural description of the rule, but if we do this then we also need to delink this feature at the same time as the spreading occurs. These two operations can be assumed to apply simultaneously. The following diagram shows a different way in which we might alternatively express the process from (8a):

In (12) it is necessary to delink [+low] since this feature and [–low] are incompatible with each other (mutually exclusive) when realized at the same time. That is, we cannot simultaneously lower our tongue body and not lower it. However, this type of delinking is something that our theory should do for us automatically. We should not need to stipulate delinking in rule after rule in a situation of this kind. So let us assume that whenever a phonological process spreads a feature or node which conflicts with one that is already present, that rule automatically delinks the prespecified feature. Given this, we can formalize rules more simply as in (8a), rather than as in (12).

Nevertheless, a word of caution is in order. In some cases we do want opposite feature values to exist within the same segment. An example of this would be complex segments such as
affricates. In non-linear theories these are typically represented as a stop followed by a fricative, both attached to the same timing element:

(13) informal characterization of an affricate:

\[ \begin{array}{c}
C \\
[-continuant] \quad [+continuant]
\end{array} \]

Notice that in this case the contradictory features are not articulated at the same time, but rather sequentially, one after the other. An analogous example would be double stops such as [kp]. These would be geometrically diagrammed as having both a Labial Node and a Dorsal Node underneath a single Place Node. These two anatomical gestures are not incongruous with each other. So we do not always want to avoid multiple or opposite feature specifications dominated by the same Root Node.

When spreading a feature or a node to an adjacent tree, any intervening nodes that would not otherwise be expressed are automatically filled in along with the appropriate dotted association lines. The universal mechanism which ensures that this is done is called the “Node Generation Convention”:

(14) \[ \begin{array}{c}
V \\
\sigma \\
\text{Root} \\
\text{Supralaryngeal} \\
\text{[+nasal]}
\end{array} \quad \begin{array}{c}
C \\
\text{Root} \\
\text{Supralaryngeal}
\end{array} \]

In (14) above we are making a vowel nasalized when it is followed by a syllable-final nasal consonant. The Supralaryngeal Node of the vowel is automatically added as a consequence of this spreading.

Recall that the Dorsal Node in the feature geometry tree dominates three binary terminal features:

(15) \[ \begin{array}{c}
\text{Dorsal} \\
\text{[high]} \\
\text{[low]} \\
\text{[back]}
\end{array} \]

If some process tells us to spread the Dorsal Node from one segment to another segment, all three of these features go along with the Dorsal Node automatically (for free). This is true regardless of whether they are overtly mentioned in the rule or not. It is a consequence of the domination relationship. The same is also true of all other nodes, such as Place, Coronal, etc. If you analyze a process as involving the spread of just one terminal feature under a node (to the exclusion of other binary features dominated by that same node), you cannot formalize this by
spreading the whole node. Rather, in such a situation you must spread only one single terminal feature, by itself. In other words, when a particular node dominates three features or more, you can spread all of them simultaneously (by spreading that node), or you can spread just one of them, but you cannot spread exactly two of them in the same rule. It is either just one feature, or else all of them, but not some. A central and important claim of feature geometry theory is that any time we encounter a situation requiring us to spread a subset of features under a node, such as two out of three or three out of four, then that is in reality not a single, unitary phonological process. Rather, in such a case we would have to spread the terminal features one-by-one, each with a separate rule. So in general, when formalizing an assimilation process, try to spread the highest node in the tree whenever you can, but be aware that this implies the simultaneous assimilation of all other features and nodes subsumed underneath the higher node.

In formal rules you can refer to nodes by placing their names right where they occur in the tree, or else by drawing in a little circle and writing the name of the node off to the side. But it’s probably best not to mix these two systems in the same rule. And be careful to put the square brackets only around terminal features, not nodes. This is because nodes are different from binary features in the sense that nodes do not have a minus value. In other words, it is impossible to refer to the absence of a node in a formal rule:

\[
\begin{align*}
\text{(16) (a) correct:} & \quad \text{(b) correct:} & \quad \text{(c) incorrect:} \\
C & & C \\
\quad \text{Root} & & \quad \text{Root} \\
\quad \quad \text{Supralaryngeal} & & \quad \quad \text{Supralaryngeal} \quad \text{[Supralaryngeal]} \\
\quad \quad \quad \text{Place} & & \quad \quad \quad \text{Place} \quad \text{[Place]} \\
\quad \quad \quad \quad \text{Coronal} & & \quad \quad \quad \quad \text{Coronal} \quad \text{[Coronal]} \\
\quad \quad \quad \quad \quad \text{[+anterior]} & & \quad \quad \quad \quad \quad \text{[+anterior]} \\
\end{align*}
\]

The trees in (16a) and (16b) above depict an alveolar or dental consonant.

A final matter to point out is that, when we draw a feature tree, the only details which count as a language-particular stipulation (complication) are the binary features. All the rest of the structure comes along for free (automatically), as a consequence of Universal Grammar (our theory). This includes all of the nodes as well as the corresponding association lines that connect them, plus their hierarchical relationships, because these are “built-in,” so to speak. For example, in the trees in (16a) and (16b) above, the only language-specific “cost” is the terminal feature [+anterior]. When looked at in this way, feature geometry mechanisms are not really more cumbersome than the formal devices of the classical model (in a linguistic sense), even though they may look like they are. In order to get further experience in understanding these concepts, part of your homework for Day 23 (to practice only) is to draw in the dotted association lines for the spreading rules in Marlett, pp. 76-77, after reading chapter 13.
Thus far we have established two distinct levels of representation which are relevant in phonological analysis. The surface phonetic representation is that of the raw data that we subject to analysis. The second level is that of the underlying representation, or deep phonological structure. The underlying level is an abstraction which allows us to show that several phonetic variant realizations are to be considered as a single unit. On this level, each phoneme is given a single representation; likewise, each morpheme is given a single representation (for all of its phonologically predictable allomorphs).

These two levels of representation are related by the set of phonological rules, which account for the phonetic realizations of the underlying forms in specific environments. In many of the cases that we have considered, this is a straightforward procedure: a rule will take the underlying structure as input, and produce the phonetic representation as output. But in many cases, several rules may be applicable to the same underlying form. How then are they to be applied?

We shall take the position that the rules are applied sequentially, and in a set order. The first rule will apply to the underlying representation; the second rule will apply to the output of the first one; the next rule will apply to the output of the second; and so on until we arrive at the surface phonetic representation. This procedure will now produce several intermediate levels of representation, as seen in this derivation:

```
Underlying Representation  / # bok' + i:n # /  
Vowel Harmony         ----- 
Vowel Lowering         bok' + e:n 
Vowel Shortening       bok' + en 
Phonetic Representation [bok'en] 
```

These intermediate levels have no independent phonological significance apart from being the input to the next rule.

### 1. Other possible methods of rule application

Let us now examine some other possible ways of applying rules. These may work in a number of cases, but they fall short in significant ways, as we shall see.

#### 1.1 Simultaneous rule application

First, we consider the possibility that all rules apply simultaneously. All rules would then apply directly to the underlying form, so that if the structural description of a rule was not met in the UR, the rule would not apply to the form. In addition, there would be no intermediate stages in the derivation from underlying form to surface form.
This method works well enough if rules do not affect each other in the derivation of a single form. However, consider the case of Brazilian Portuguese, where we find these forms:

(1) /tiagɔ/ → [tʃiagu] ‘that’
     /diferə/ → [dʒifera] ‘differs’
(2) /ɡato/ → [ɡatu] ‘cat’
     /dedo/ → [dedu] ‘finger’
     /tipo/ → [tʃipu] ‘type’
(3) /onde/ → [ondʒi] ‘where’
     /mɔhte/ → [mɔhtʃi] ‘death’

From the forms in (1), we see the need for a rule such as:

(4) Palatalization: C
     [−sonorant
     +coronal]
     → [+delayed release
     +high
     −back]
     / ___ V
     [+high
     −back]

And from (2), we see evidence for another rule:

(5) Vowel Raising: V
     → [+high] / ___ #
     [−low]

Applying these rules simultaneously would work fine to derive the correct forms in (1) and (2). However, in (3), we run into trouble. It is evident from examining these forms that both rules must apply here. The mid vowel is raised to high, and the alveolar stop becomes an alveopalatal affricate. The rules as written would not yield the correct results if applied simultaneously, since only the environment for the vowel raising rule is met in the underlying forms /onde/ and /mɔhte/. The palatalization rule would not be able to apply, because the underlying form does not have a high vowel at all. To correct this situation, while maintaining the position that rules apply simultaneously, we could add a second palatalization rule to take care of these cases:

(6) Palatalization 2: C
     [−sonorant
     +coronal]
     → [+delayed release
     +high
     −back]
     / ___ V #
     [−back
     −low]

Here we allow the palatalization to be triggered by a mid front vowel just in case that vowel is also word-final, and we can now predict the correct results. In the UR /onde/, the structural description of Palatalization 2 and Vowel Raising is met: we have a mid vowel at the end of the word, and an alveolar stop before a word-final mid front vowel. Both rules now apply simultaneously and directly to the underlying form, yielding the correct surface form [ondʒi].

However, the addition of rule (6) has made an unjustified complication: we have recapitulated the environment of the vowel raising rule (the #) within the environment for the second palatalization rule. And the palatalization process really has nothing to do with proximity to the end of the word, and does not in general apply before mid vowels. Writing a new rule like (6) seems to miss the whole point. Palatalization is a process that happens only before high front vowels (no matter
where they come from) anywhere in the word. A simpler and more straightforward solution is achieved if we assume that rules apply sequentially, and in an order. In this way we can get rid of rule (6) altogether, and allow one rule to operate on the output of another rule — not necessarily always to the underlying form. A derivation of /onde/ would then proceed as follows:

UR / # onde # /
Vowel Raising ondi
Palatalization ondʒi
PR [ondʒi]

Sometimes a different sort of problem comes up. Consider the German dialects where the following rules hold:

(7) Devoicing: C → [−voice] / ___ #

(8) Lengthening: V → [+long] / ___ C

Rule (7) is seen at work in (9), and rule (8) is seen in (10):

(9) /rund/ → [runt] ‘round’
/halb/ → [halp] ‘half’
(10) /loben/ → [lo:ben] ‘to praise’
/rozə/ → [ro:zə] ‘rose’

In these dialects, the correct realization of the underlying form /ab/ is [ap]. Notice here that both rules are applicable to the underlying form, but both rules must not be allowed to apply. We must prevent the vowel lengthening rule from applying at all, by applying only the devoicing rule to the underlying form. This rule then destroys the environment for the vowel lengthening rule, which can now no longer apply. This would be expected if rules apply sequentially, but problematical if they must apply simultaneously.

1.2 Free reapplication

Another possibility for rule application is to allow rules to apply to a form whenever their structural description is met. With this procedure, several rules could still apply simultaneously, but one application of the set of rules would not necessarily yield the phonetic output in one step. To the underlying forms, we apply all the rules that can apply. But the changes made in this way may make other rules applicable. At this point all rules would again be checked and applied if possible. This modification would help in some cases, such as the Brazilian Portuguese data given above. To the underlying form /onde/, only the vowel raising rule is applicable. Once the vowel is raised, though, the palatalization rule now becomes applicable. After applying this rule, we get the correct surface form [ondʒi].
With this proposal we are not requiring the rules to apply in any set order, but simply to apply (and reapply) whenever their structural description is met. This method has shortcomings too, however. Consider the following apocope rule in Isthmus Nahuatl, as illustrated here:

(11) Apocope: \[ V \rightarrow \emptyset / V \]
\[ [-\text{long}] \]
\[ \begin{array}{c}
+\text{son} \\
+\text{voice}
\end{array} \] #

(12) /sikalili/ \rightarrow [sikalil] ‘put it in it’
/kitaja/ \rightarrow [kitaj] ‘he already sees it’

Consider also the effects of a second rule, which devoices certain sonorant consonants word-finally.

(13) Devoicing: \[ C \rightarrow [-\text{voice}] / \]
\[ \begin{array}{c}
+\text{son} \\
-\text{nas}
\end{array} #

(14) /tajo:l/ \rightarrow [tajo:̥] ‘shelled corn’
/so:tkuj/ \rightarrow [so:tkuj] ‘elbow’

Note carefully that the devoicing does not affect forms such as those in (12). The rule is not applicable directly to the underlying form, but after the deletion of the final vowel, it becomes applicable. Even so, however, we must prevent the devoicing from occurring. In the Brazilian Portuguese case above, we had the opposite situation. There a rule could not apply until another rule had created the appropriate environment, and we had to make sure that the rule did apply. So it is not enough simply to allow rules to apply whenever their structural description is met. This would work for Brazilian Portuguese, but not for the data of Isthmus Nahuatl. Note again, though, that a sequential and ordered application of the rules would give the correct results for both of these cases. In Isthmus Nahuatl, we would specify that the devoicing rule must apply before the apocope rule. Devoicing would not be applicable to the underlying forms that we are considering, of course, because at this stage the sonorant consonants are not word-final. Then we allow apocope to apply, deleting the word-final vowel. By forcing the devoicing rule to apply first, and thereby not allowing it to apply later, we obtain the desired results.

\[ \begin{array}{l}
\text{UR} / \# \text{kitaja} \# /
\text{Devoicing} ----- \\
\text{Apocope} \quad \text{kitaj} \\
\text{PR} \quad \text{[kitaj]}
\end{array} \]

As we have seen, simultaneous rule application is not adequate to handle all language data. In some cases one rule must apply to the output of another rule. Likewise, any type of free (re)application of rules is insufficient too, since in other cases we must block rules from applying even when the appropriate environment becomes available at some point in the derivation.
2. Ordered rules

As a result of these considerations, we conclude that rules must apply in sequence; the first rule applies directly to the UR, and after that, each rule applies to the output of the preceding rule. In addition, we have found that the sequence in which rules apply is not always free. In some cases we will have to make a specific restriction that one rule must apply before another rule. A rule will apply at only one point in the sequence. If it is ordered at a point where its structural description is not met, it will be applied vacuously (i.e., having no effect). Even if the appropriate environment is created for the rule later on in the derivation, the rule may no longer apply. (Some linguists have found cases where a rule must apply to its own output; in such a situation, all applications of the same rule must come together, and not interspersed throughout the derivation.)

In keeping with this approach, which handles the overwhelming majority of rule interactions, we will list all the phonological rules of a language in one overall sequence. Of course, the order of the rules must be one that predicts the correct results for all the forms of the language. Beyond that, however, certain rules may be crucially ordered with respect to others in the sequence. In this case, we draw an arc between the pair of rules to indicate this necessary and crucial ordering. Rules which are crucially ordered with respect to each other are by convention listed toward the top of the sequence of rules. Rules which do not interact with other rules, on the other hand, are generally ordered last. (Of course, since these rules have no effect on the application of other rules, the same results would always be obtained no matter where they were applied in the overall sequence. By convention, though, they are put last.)

When we state an ordering restriction between two rules, we need to demonstrate that it is needed. (Other things being equal, a solution that does not require rule ordering restrictions would be simpler and preferable to one which does.) To do this, we show that one order produces the correct forms, but that the opposite order of application of the two rules leads to incorrect surface forms. Thus, for each crucial ordering, indicated by an arc in the ordering statement, we show a corresponding wrong derivation.

A phonological statement, then, will not simply consist of a list of the phonological rules of the language. It will also involve an ordering statement of how those rules are to be applied. Here is an example of such an ordering statement.

Vowel Lowering
Apocope
Cluster Simplification
Nonapical Deletion
Nasal Assimilation
Palatalization

2.1 Determining crucial rule ordering

If rules in some cases must be specified as applying in a certain order, how do we determine when such cases will arise? The final test is checking that the correct surface forms are always generated.
But without doing derivations of all the forms in the language, how do we know where to check for cases of crucial rule ordering?

It will be important for us to consider the linear sequence of the input and the environment of each rule, as we consider possible rule interactions. Consider rules (7) and (8) from German, as discussed above. The devoicing rule applies to voiced obstruents followed by a word boundary. The lengthening rule applies to vowels followed by voiced obstruents. Could these two sequences both occur in a single form? Yes, in sequences such as:

\[
\begin{array}{c}
V \\
\text{-son} \\
\text{+voice} \\
C \\
# \\
\end{array}
\]

Investigating forms with this sequence, such as /ab/, will help us decide on the necessity of rule ordering restrictions. Both rules are applicable to the underlying form, but we find that only one rule in fact applies. Here is one case where the order is crucial, because the application of one rule has destroyed the environment for the other. (We sometimes say that the first rule bleeds the second one.)

Consider next the input and environment of the Brazilian Portuguese palatalization rule. It requires a coronal consonant followed by [i]. Since the vowel raising rule produces [i]'s, it will be relevant to investigate what happens to forms which have sequences such as:

\[
\begin{array}{c}
C \\
\text{-son} \\
\text{+cor} \\
V \\
\text{-back} \\
\text{-high} \\
\text{-low} \\
# \\
\end{array}
\]

In forms with this sequence, such as /onde/, we need to find out if the mid vowel is raised to [i] in time for palatalization to occur before the newly formed [i]. When one rule creates the environment for another to apply, as happens here, we say that the first rule feeds the second one.

What forms do we need to look at in Isthmus Nahuatl in order to determine the need for crucial ordering between the apocope and devoicing rules? Here we must look for forms which have both environments (potentially) occurring:

\[
\begin{array}{c}
V \\
\text{+son} \\
\text{–nas} \\
\text{+voice} \\
C \\
# \\
\end{array}
\]

This is exactly the case with /kitaja/, so the derivation of this UR will have to be scrutinized carefully in order to determine if a crucial order is needed.

As a result of crucial rule orderings, rules that look like they should apply to a UR may not actually apply. In other cases, rules may apply to a form even though the environment is not seen in the UR.
Similarly, surface forms may sometimes appear to be counterexamples to the rules: in some cases the rule has applied, although we cannot see the environment for it in the PR; or a rule has not applied, even though the environment for the application of the rule is present in the PR. For example, in the German UR /ab/, the vowel lengthening rule would seem to be applicable, although it does not apply; in the Isthmus Nahuatl PR [kitaj] the final voiced [j] would seem to be a contradiction to the devoicing rule.

The way the rules are written often makes a difference in the necessity of crucial orderings. A general rule, which applies to a large class of segments or in a large number of environments, may interact with other rules that affect similar segments or sequences of segments. Rules that are more specific, on the other hand, may avoid the need for crucial ordering between two rules, because the more specific environments required may not both occur in individual forms. Other things being equal, a solution that does not need crucial rule ordering is preferable to one that does. However, modifying the rules in order to avoid rule ordering often obscures the generality of the phonological processes that the rules are supposed to reflect. A case in point is the palatalization rule in the Brazilian Portuguese example given above. The addition of the more detailed rule (6) avoided the need for crucial ordering, but it obscured the motivation for the palatalization process. (And of course, in this case, we still needed another palatalization rule (4) to handle palatalization in other environments.)

2.2 Derivations

A derivation is a display of forms arranged to show the changes effected by the successive application of phonological rules, starting from the UR and ending with the PR. The notion of a derivation follows from our position that rules apply sequentially, and in a consistent order. The derivation will show each intermediate stage that arises after the application of each rule. Each intermediate stage, then, serves as the input to the next rule.

Let us take an example from Lardil, an Australian language. Consider first the following rules:

(15) Vowel Lowering: \( V \rightarrow [+\text{low}] / \_ \_ \_ \_ \# \)

(16) Apocope: \( V \rightarrow \emptyset / V C_1 V C_1 \_ \_ \_ \_ \# \)

(17) Cluster Simplification: \( C \rightarrow \emptyset / C \_ \_ \_ \_ \# \)

(18) Nonapical Deletion: \( C \rightarrow \emptyset / \_ \_ \_ \_ \# \) 
\([-\text{cor}]\)

The following ordering statement is needed along with these rules:

\[
\text{Vowel Lowering, Apocope, Cluster Simplification, Nonapical Deletion}
\]
Using this order, then, we can show derivations of certain sample forms, as follows:

<table>
<thead>
<tr>
<th>Rule</th>
<th>UR</th>
<th>Vowel Lowering</th>
<th>Apocope</th>
<th>Cluster Simplification</th>
<th>Nonapical Deletion</th>
<th>PR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel Lowering</td>
<td>/ # ŋaluk # /</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td>putuka + n # /</td>
<td>/ # putuka # /</td>
</tr>
<tr>
<td>Apocope</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>putuka + n # /</td>
<td>/ # putuka # /</td>
</tr>
<tr>
<td>Cluster Simplification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>putuka</td>
<td>/ # putuka # /</td>
</tr>
<tr>
<td>Nonapical Deletion</td>
<td></td>
<td>ŋalu</td>
<td></td>
<td></td>
<td>putuka</td>
<td>/ # putuka # /</td>
</tr>
<tr>
<td>PR</td>
<td>[ŋalu]</td>
<td>[putukan]</td>
<td>[putu]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rule</th>
<th>UR</th>
<th>Vowel Lowering</th>
<th>Apocope</th>
<th>Cluster Simplification</th>
<th>Nonapical Deletion</th>
<th>PR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel Lowering</td>
<td>/ # wanka + r # /</td>
<td>-----</td>
<td></td>
<td></td>
<td>[wankan]</td>
<td>[muŋkumuŋku]</td>
</tr>
<tr>
<td>Apocope</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>kantukan</td>
<td>/ # muŋkumuŋku # /</td>
</tr>
<tr>
<td>Cluster Simplification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>muŋkumuŋk</td>
<td>/ # muŋkumuŋk</td>
</tr>
<tr>
<td>Nonapical Deletion</td>
<td></td>
<td>[wankar]</td>
<td></td>
<td></td>
<td>muŋkumuŋ</td>
<td>/ # muŋkumu</td>
</tr>
<tr>
<td>PR</td>
<td>[wankar]</td>
<td>[kantukan]</td>
<td>[muŋkumu]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notice the following things about these derivations:

- The UR of the entire word, placed between diagonal slashes, is shown on the top line. All boundaries are indicated in the underlying form, both word boundaries and morpheme boundaries. No morphemes are added in the course of the derivation; all morphemes are present in the UR. (It is the business of the morphological component of the grammar to account for the make-up of the word, not of the phonological component.)

- Each successive line shows the effect on the form of a single phonological rule, whose name is indicated to the left. Rules for suprasegmental features as well as for segments are included. If a rule does not apply to a form, a dashed line (-----) indicates that there is no effect.

- The only rule that refers to the UR is the first rule whose structural description is met. All other rules apply to the output of the rule which last applied before each of them respectively.

- All rules apply in the same order for every word. All rules must be shown in each derivation, regardless of whether they apply or not. All boundaries are (potentially) carried down through the derivation.

- The PR is the output of the last rule that applied. It is shown on a separate line, without word or morpheme boundaries, and is enclosed in square brackets. (The intermediate levels are not enclosed in slashes or brackets, since these levels have no independent significance.)

We indicated earlier that the need for crucial ordering must be demonstrated with an incorrect derivation. This is done for each pair of rules that are crucially ordered. We choose a form (ideally) in which the application of only those two rules is relevant (and critical). Then, we switch the order of the pair of rules in question, without changing the order of any of the other rules. By applying the rules in this new order, an incorrect surface form will result (for at least some forms). An ungrammatical output is indicated with an asterisk (*). To be complete, a derivation should also be
shown for the same word with the correct order of the rules, showing that the claimed order does indeed yield the correct output.

To show that Apocope must apply before Cluster Simplification:

<table>
<thead>
<tr>
<th>Rule Type</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel Lowering</td>
<td>kantukanta</td>
<td>kantukanta</td>
</tr>
<tr>
<td>Apocope</td>
<td>kantukant</td>
<td>kantukant</td>
</tr>
<tr>
<td>Cluster Simplification</td>
<td>kantukan</td>
<td>-----</td>
</tr>
<tr>
<td>Nonapical Deletion</td>
<td>-----</td>
<td>Nonapical Deletion</td>
</tr>
<tr>
<td>PR</td>
<td>[kantukan]</td>
<td>PR *[kantukant]</td>
</tr>
</tbody>
</table>

When we think about a solution for a set of data, then, we must constantly think about three parts to the solution:

1. the underlying forms,
2. the phonological rules, and
3. the way in which the rules apply.

Each of these represents a distinct hypothesis. We might have the right underlying forms and the right rules, but the wrong order of application. Or we might have the right rules, and the right ordering, but the wrong underlying forms. Or everything might be right except that one phonological rule is overly general or overly restricted. Phonological analysis requires the constant juggling of these factors to come up with the best solution. In general, however, rule ordering is preferred to complication of rules.
Fore (Papua New Guinea)

<table>
<thead>
<tr>
<th></th>
<th>‘my...’</th>
<th>‘your...’</th>
<th>‘his...’</th>
<th>‘our...’</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>kaine</td>
<td>kaiga</td>
<td>kaiwa</td>
<td>kaire</td>
<td>‘clothes’</td>
</tr>
<tr>
<td>2.</td>
<td>aune</td>
<td>auga</td>
<td>auwa</td>
<td>aure</td>
<td>‘eye’</td>
</tr>
<tr>
<td>3.</td>
<td>agene</td>
<td>agega</td>
<td>agewa</td>
<td>agere</td>
<td>‘trap’</td>
</tr>
<tr>
<td>4.</td>
<td>auʔne</td>
<td>auka</td>
<td>auʔwa</td>
<td>aute</td>
<td>‘name’</td>
</tr>
<tr>
<td>5.</td>
<td>koʔne</td>
<td>koka</td>
<td>koʔwa</td>
<td>kote</td>
<td>‘back’</td>
</tr>
<tr>
<td>6.</td>
<td>abeʔne</td>
<td>abeka</td>
<td>abeʔwa</td>
<td>abete</td>
<td>‘navel’</td>
</tr>
</tbody>
</table>

Source of data: Ray and Ruth Nicholson.
Lamba exercise

Your written homework for Day 24 (to be turned in) is the Lamba exercise on page 193 of Marlett. Follow the procedures for morphophonemic analysis from a previous handout and examples in lab. When you consider those morphemes which have more than one allomorph (phonetic variant), you should note that there is an alternation between [s] and [ʃ] in the verb roots which mean ‘hide’, ‘wound’, and ‘plaster’. Posit a rule to account for these forms, stating it in prose only (don’t worry about formalizing it).

You should also note that the ‘applied’ suffix in column three exhibits the two allomorphs [-ila] and [-ela]. In other words, there is a phonological alternation between [i] and [e]. You should also write a rule to account for these forms (hint: what is occurring here is a case of vowel harmony; cf. Marlett, chapter 18). In addition to stating this rule in prose, also try to formalize it using the newer notation which you have been learning (based on Marlett, p. 230). You may find it difficult to formalize this rule on your own, so feel free to get help if you need it. However, do make an attempt at it before seeking help.

Having posited these two rules, we must now ask ourselves if they are crucially ordered (hint: consider the forms [tʃesela] and [kosela]). If the two rules do need to be ordered, demonstrate this by giving two derivations of one of the words just mentioned. In the first derivation, order the rules correctly and show that this produces the correct output (phonetic form). Then give a second derivation in which the order of the two rules is reversed, leading to an incorrect form (which should be marked with an *).

Now answer these questions: what is the underlying representation of the ‘applied’ suffix which occurs in column three? Also, based on just these data, is the sound [ʃ] a distinct phoneme in this language? Note that the phonetic sequence [tʃ] which occurs at the beginning of two roots can be assumed to be an affricate ([č]).

When you turn in your write-up of this problem on Day 24, we will give you a discussion of the solution and analysis.
A SIMPLE NON-RIGOROUS ALGORITHM FOR THE EXTRACTION OF INSIGHTFUL ANALYSES FROM THE PHONOLOGICAL MATERIALS OF HUMAN LANGUAGES (IN 20 EASY STEPS)

1. ORGANIZE FORMS INTO PARADIGMS
   This will let you see alternations and their environments more clearly. It is not much good asking how to account for things until you know what to account for. (This step has been taken for you already in the problems you have been getting. In working on your own data it is a major part of the task. Many times the biggest obstacle to an insight being gotten is disorganization of the relevant data.)

2. MAKE MORPHEME CUTS
   Remember that all morpheme cuts are hypotheses, and you may need to change some of them. Remember that the boundaries are often obscured by phonological rules. Where the boundary is not immediately clear, mark (e.g. with colored pens) which parts clearly belong to one morpheme, which parts clearly belong to its neighbor, and which are questionable. Consider the following possibilities with respect to the questionable material:
   - It may belong totally to the first morpheme, or totally to the second.
   - It may belong to both morphemes, being formed by coalescence of the ending of the first morpheme and the beginning of the next.
   - It may belong to neither morpheme, being either epenthesized or belonging to a third morpheme whose presence you had not suspected.

3. LIST THE MORPHEMES YOU HAVE POSITED
   Take special note of those that alternate. Group those that show similar patterns of alternation together.

4. REARRANGE THE DATA
   Put side by side in the paradigm those forms in which morphemes are undergoing parallel alternations.

5. TAKE THE SIMPLE CASES FIRST
   Pick some pattern of allomorphy where the allomorphs are only slightly different and where it looks clear what the motivating environment is. When you hit difficult cases after doing a bunch of simple rules, go to step #17.

6. CHECK TO SEE IF THE ALTERNATION HAS ALREADY BEEN ACCOUNTED FOR
   Check any rules you have written previously to see if they account (or should be made to account) for the new alternation. For instance, if you have an alternation between voiced and
voiceless consonants and you already have a rule voicing certain consonants, see if the new
alternation is not just a new case of the application of the earlier rule, or if it doesn’t show
that the earlier rule was wrong and you really need to account for both the earlier data and
these new data by a devoicing rule. In either case, make whatever adjustments are necessary
and go back to step #5.

7. WRITE DOWN BOTH POSSIBILITIES
For instance, if you are looking at a pattern in which voiced consonants are alternating with
voiceless ones, then either the voiced ones are devoicing, or else the voiceless ones are
voicing (becoming [+voice]). Write down the two possibilities on opposite sides of a sheet of
paper and draw a line down the middle. Write the hypotheses each one entails (what you get
in steps 8-13) on the appropriate half of the paper. For example:

\[
\begin{align*}
C & \to [-\text{voice}] & & C & \to [+\text{voice}] \\
\text{UR’s:} & \quad \text{kid} & \text{‘goat’} & & \text{UR’s:} & \quad \text{kit} & \text{‘goat’} \\
& \quad \text{bæd} & \text{‘evil’} & & \quad \text{bæt} & \text{‘evil’}
\end{align*}
\]

8. WRITE DOWN THE INPUT FORMS
For each hypothesis, you will be positing that the input form will have what is to the left of
the arrow rather than what is to the right. For instance, in the example given above, the two
rules might be posited as alternative explanations for the alternations in the morphemes
\text{kid–kit} ‘goat’ and \text{bæd–bæt} ‘evil’. The forms ending in \text{d} would be underlying in the
devoicing hypothesis, and the forms ending in \text{t} would be underlying in the voicing
hypothesis. The work sheet would thus look like this:

\[
\begin{align*}
\text{C} & \to [-\text{voice}] & & \text{C} & \to [+\text{voice}] \\
\text{UR’s:} & \quad \text{kid} & \text{‘goat’} & & \text{UR’s:} & \quad \text{kit} & \text{‘goat’} \\
& \quad \text{bæd} & \text{‘evil’} & & \quad \text{bæt} & \text{‘evil’}
\end{align*}
\]

9. DETERMINE THE ENVIRONMENTS
Looking only at the forms which alternate, see what environment would make each one of
the two rules work. Write down the rule with its environment. At this point your work sheet
would look something like this:

\[
\begin{align*}
\text{C} & \to [-\text{voice}] & & \text{C} & \to [+\text{voice}] \\
\text{UR’s:} & \quad \text{kid} & \text{‘goat’} & & \text{UR’s:} & \quad \text{kit} & \text{‘goat’} \\
& \quad \text{bæd} & \text{‘evil’} & & \quad \text{bæt} & \text{‘evil’} \\
\text{C} & \to [-\text{voice}] & / \quad \sigma & & \text{C} & \to [+\text{voice}] & / \quad \text{V} \quad \Sigma \quad \text{V}
\end{align*}
\]

10. ASK, “DOES IT WORK?”
Determine what kind of data would falsify each hypothesis. These will be cases in which
the structural description of the rule is met but in which the change predicted by the rule has
not taken place. This is easiest to check if there are non-alternating forms. For example, if
the morpheme for ‘hat’ always occurs as [hæt] regardless of whether it’s followed by a
vowel or a consonant or a word boundary, this indicates that a voicing rule is not possible
and that the UR’s of ‘goat’ and ‘evil’ end with voiced stops. If there are no non-alternating
forms, check all other sequences that could falsify each rule. For example:

The \text{devoicing} rule above has as its structural description a syllable-final position. The
rule says that in all such situations the consonant should be voiceless. Therefore any voiced consonant in the environment \( \sigma \) would falsify the rule.

The voicing rule has as its structural description a VCV sequence. It predicts that in any such sequence the C should be voiced. Therefore any voiceless consonant in the environment \( V \) would falsify the rule.

Always ask what the rule says should never occur. Whatever that is, try to find it in the data.

11. IF IT DOESN’T WORK, PATCH IT UP

If you succeed in falsifying a hypothesis, see what it takes to make it float. Consider the following possible ways of fixing up an ailing proposal:

a. Modify the rule(s). Broaden or narrow the class of segments to which the rule can apply. Broaden or narrow the class(es) of segments in the environment.

b. Change your morpheme cuts.

c. Posit different underlying forms. Sometimes it may be necessary to break the morphemes into two groups in terms of UR’s. For example, if \( m \) and \( n \) alternate for two different morphemes, with \( m \) occurring before labials and \( n \) occurring before alveolars, this does not mean that the UR for each morpheme must be the same. It may be that the UR for one form contains \( m \) while the UR for the other form contains \( n \).

d. See if it might not be the case that interaction with some other rule(s) is causing the problem.

e. If absolutely necessary, posit nonphonologically-characterized (suppletive) classes of morphemes which are either undergoing/not undergoing the rule, or conditioning/not conditioning it. See if the classes you posit can be characterized syntactically or semantically, or if they are the same classes needed by some other phonological rule. If they cannot be characterized in any such way, they are ad hoc classes. Posit these only as a last resort.

Keep going back to step #10, trying to falsify the rules, until you are pretty sure that they both work. Then go on to step #12.

12. JUDGE THE TWO HYPOTHESES

When both hypotheses finally work, judge them by the following criteria:

a. Simplicity. Ceteris paribus (‘all other things being equal’) the simpler solution is better.

b. Predictiveness. Ceteris paribus the solution that leaves unpredicted what could be predicted is not to be preferred. For instance, if, in some language, \( i \) (and no other vowel) alternates with \( \emptyset \), and epenthesis and deletion solutions work equally well, the epenthesis solution would be preferable because it would explain (predict) why the vowel that alternates was always \( i \), while that would be just an ad hoc fact in the deletion analysis.

c. Naturalness. Ceteris paribus the solution that does what you expect languages to do is preferable. The more you learn about languages’ sound systems the better you will be able to apply this criterion.

d. Adhocity. Ceteris paribus the solution needing less ad hoc machinery (e.g. morphemic classes, new types of formalisms) is preferable.
e. Integration. Ceteris paribus the solution that wreaks less havoc with the rest of your analysis is preferable.
   If it actually makes things easier and/or more natural in the rest of your analysis, latch onto it!
Unfortunately, ceteris are rarely paribus. Often the more natural solution will mess up your analysis of other phenomena, or the simpler solution will be more ad hoc. You will have to weigh the different factors. The above criteria are more or less in order from weakest to strongest. Thus more weight should be given to integration than to simplicity, or to adhocity than to naturalness, etc.

13. MAKE YOUR CHOICE, AND WRITE IT DOWN
If a clear choice can be made, make it. Write down the reasoning behind it. Write down the crucial forms which show this hypothesis to be the correct one. If you cannot make a clear choice, write down the lines of reasoning for and against each hypothesis. In either case, note what predictions each hypothesis makes, so you can check them out against further data. Even after you are 95% certain that one solution is correct and the other wrong, it is wise to keep the rejected one in the corner of your mind. Further data just might show it to be right after all.

14. MAKE SURE THAT THIS IS A SEPARATE PROCESS
Check again to make sure that this new rule is not duplicating the work of a previously posited one. If it is, collapse the two rules into a single generalization.

15. PEEL BACK THE EFFECTS OF THE RULE
Go through the data undoing what was done by the rule you have just posited. For each form affected by the rule, write down the form you are positing as the input to the rule. Use those forms rather than the surface forms as the basis for further analysis. This will make it easier to see what the rules ordered earlier in the derivation are doing.

16. GO BACK TO STEP #5

17. EXPLAIN COMPLEXITIES BY INTERACTION OF RULES
When you have a case of complex allomorphy and/or a small set of forms which are stubbornly exceptional to well-motivated rules, see if you can’t account for the complexity through the interaction of several of the rules you have posited for simpler cases. You’ll be surprised how often you can. This is where the integration business (#12e) will show up best.

18. MAKE SURE IT’S NOT A MISTEAK
If you still can’t explain certain cases, go check with an informant to make sure the problem is not a transcriptional error or a slip of the tongue.

19. KEEP TRACK OF RESIDUE
If you still can’t account for a problem, chuck it into a corner marked RESIDUE and go back to step #5. Periodically review what you have in the Residue section and see if any new rules or analysis have any bearing on it. If you have analyzed everything else and are
left with residue, check to see if only one morpheme is involved. If so, account for the alternation via suppletion. Otherwise, either just leave it (plainly marked RESIDUE), or write a brute-force rule to get rid of it.

20. START OVER AGAIN
When you have accounted for all the alternations in your data, go elicit more data, and start over again at step #1.

Homework for Day 25:
• read Marlett, chapters 30 and 35
• read Schane, chapter 5 (on reserve)
• Dakota Sioux, p. 202 (for practice only)
• Isthmus Zapotec, p. 202 (for practice only)

Day 24: The Big Picture

In today’s class we are going to see how to deal with (analyze) two phonological phenomena which have traditionally been considered to involve more than one segment simultaneously: stress and duration.

Some languages have phonemic (contrastive) oppositions of segmental length. This is more common with vowels than with consonants. Other languages, such as English, have systematic differences in the length of certain segments, but these are phonologically predictable and thus are not phonemic, e.g., beat, beam, bee, bead. 

In the past, such cases were often accounted for with the feature [+long]; now, however, we resort to a difference in timing units in the CV skeleton.

Stress is not always realized phonetically as just greater intensity; it also often correlates with higher pitch, greater length, and greater muscular activity in general. The same is true in sign languages!

It is best to think of stress as a feature that characterizes whole syllables, not just vowels alone.

In some cases, linguists have found it useful to represent long vowels as linked simultaneously with a V and a C slot, and affricates, whose timing might be shorter than what would be normal for two distinct segments, are linked to a single slot, e.g.,

\[
\text{\text{d} \text{zim}}
\]

Long vowels might be alternatively represented as

\[
\text{a} \quad \text{a}
\]

The decision as to how to represent long segments such as these depends on how they behave in the phonological system of the language in question.

Two languages for which the VC interpretation has been proposed are Quechua and Chamicuro.

For example, aspirated stops tend to be more heavily aspirated in stressed syllables than they are in unstressed syllables.

What is the linguistic function of stress?

In most languages that have stress, stress serves a function, that is, each word has one primarily stressed syllable, so you can count the number of words in an utterance by counting the number of primary stresses.

In some languages also, stress serves a delimitating function, that is, if stress always occurs a fixed distance from the word margin (beginning or end), you can always determine exactly where word breaks come in an utterance relative to where you find the stressed syllables.

Stress-timed languages vs. syllable-timed languages

In a syllable-timed language such as French or Spanish, each syllable takes about the same amount of time to say, regardless of whether or not it is stressed.
• This is why such languages sound to speakers of English like they have a machine-gun type of rhythm (vs. Morse code rhythm).
• In stress-timed languages such as English, on the other hand, there is about the same amount of time between stressed syllables, regardless of how many unstressed syllables intervene. The vowels in these unstressed syllables tend to be reduced to schwa.
• However, in these cases, the stressed syllables tend to be spaced relatively evenly throughout, and these determine the overall rhythm of speech.
• Many languages have differing __________________________ of stress.
• For example, English has primary and secondary stresses as well as unstressed syllables: _hortatory_
• Whenever possible, we want to discover and state the predictable generalizations concerning the placement of stress in a particular language.
• Some factors which often govern the predictable placement of stress:

(1) __________________________ of the syllable relative to a word margin:
• Serbo-Croatian: stress the last syllable of the word
• Finnish: stress the first syllable of the word
• Chamicuro: stress the penultimate syllable of the word

(2) __________________________ structure:
• Syllables with long vowels have a tendency to attract stress more than those with short vowels.
• Similarly, closed syllables tend to attract stress more than open ones do.
• The generalization is that __________________________ syllables tend to attract stress more than light ones do. (Define heavy syllables as those which have a branching rhyme structure.)
• Syllables having a shape such as CVVC are sometimes analyzed as “superheavy” because they behave in peculiar ways in some languages.

(3) __________________________ class:
• In some languages, the stress patterns for nouns and verbs are predictably different, e.g., convert, pervert, knockdown, etc.
• Some phonological processes which can be caused by stress are:
  -Diphthongization of vowels which are stressed, e.g., Spanish: _poder_ “to be able” vs. _pwédo_ “I am able”.
  -Reduction of vowels which are unstressed (a type of neutralization)
  -An example would be photograph vs. _photographer_
  -Consonants in stressed syllables have a greater tendency to be aspirated or fortis in pronunciation.
• How do we formalize stress?
• Classical generative phonology used the feature [+stress].

• An example would be the stress rule for Latin:
  V → [+stress] / _____ C0 (( V C1 ) V C0) #
  [-long]
• In the more recent theory of __________________________ phonology, stress is represented as a function of rhythmic patterns in larger suprasegmental structural units such as feet, e.g.,

PrWd
  word level

  s w s w
  foot level

  s w s w
  syllable level

  ex por ta tion
  phoneme level
Evaluating analyses

The following guidelines for choosing between alternative analyses are given in order of importance. The three points under (5) are of equal importance.

1. **THE NULL HYPOTHESIS.** Given two solutions, choose the one that entails identical underlying and surface representations, *unless there is evidence to the contrary* (such as complementary distribution, plausible conditioning factors in the environment, etc.). In other words, don’t propose abstract underlying representations unless there is good reason to.

2. **NONSUPPLETION.** Given two solutions, choose the one that allows a single underlying representation for a given morpheme, with all variants derived by rule.

3. **PREDICTABILITY.** Given two solutions, choose the one that minimizes exceptions. In other words, rules that only apply to certain arbitrary classes of words and not to other phonetically similar sets of words are to be avoided whenever possible.

4. **SIMPLICITY.** Given two solutions, choose the one that is simpler. A solution with fewer rules is (usually) simpler than one with more rules. A rule requiring fewer features and symbols (such as V, C, #) is simpler than a rule with more features and symbols. Usually, the simpler a solution is the more general it is also. This simplicity criterion therefore implies that more general solutions are preferred over less general ones.

5. **PHONOLOGICAL CONDITIONING.** Given two solutions, choose the one that minimizes reference to grammatical information. This includes reference to grammatical categories (noun, verb, etc.) in phonological rules, as well as reference to morpheme (+) boundaries.

   **NATURALNESS.** Given two solutions, choose the one that makes use of phonetically motivated (reasonable) rules. Such rules *usually* (but not always) refer to the same feature(s) in the specified change and in the environment.

   **MARKEDNESS.** Given two solutions, choose the one that allows the least marked underlying inventory of phonemes. (*Marked* means rare or unusual in languages universally.) For example, it is very marked to have voiced stops but no voiceless stops underlyingly. Therefore, a solution having all surface [t]’s derived from underlying /d/’s is more marked than having the [d]’s derived instead from /t/’s.

Cite as:
Day 25: The Big Picture

In today’s class we are going to discuss the concept of redundancy in distinctive feature matrices, and how we deal with it by removing redundant specifications (values) and then filling these in later with either universal or language-particular redundancy rules.

Redundancy rules are those that fill in redundant (predictable, non-distinctive) features making up the intrinsic content of individual segments.

We have said thus far that when we want to define a phoneme, we only put in the minimal amount of features that are necessary to specify that segment as distinct from all the other ones present in the language.

Similarly, in defining natural classes, we use a minimum of features, based on the inventory of phonemes in the language in question.

Redundancy rules give us the necessary apparatus to supply the “phonetic” values for such features which we leave out.

In different languages the same class of sounds might be phonologically defined by a different set and/or number of features because of the overall pattern (inventory) of the system.

For example, to refer to the vowels [u] and [o] in Spanish, it is sufficient to specify [+round].

• For Spanish, [+back] would be redundant.
  • However, since French has some front rounded vowels, we would need to specify both [+round] and [+back] in this language.
  • Because the values of these features are predictable, there is no need to mention them in the basic, underlying definition of a segment (or natural class of segments), so we can simplify our analysis.
  • These redundant values would then be filled in or supplied by rules (since they are predictable).
  • We can distinguish two basic types of redundancy statements:

    • **redundancies**

      are those which would by definition hold true for all languages as a consequence of physiological or cognitive constraints on possible arrangements of articulatory gestures.

      • For example,

• There is a universal mechanism or device known as the Redundancy Rule Ordering Constraint (RROC). What it says is that all redundancy rules are automatically ordered to apply before any other phonological rules that crucially make reference to these redundant values.

• Illustrate this interaction with the following example from Chamicuro:

  • There is a word-final devoicing process which affects unrounded vowels:


    V → [+voice] / [+voice] ___ #

    [–round]

• In the case of the derivations above, the final devoicing rule would be fed by two redundancy rules which supply the feature [–round]:

  [–back] → [–round] and  [–low] → [–round]

• Underlyingly, the front vowels /i/ and /e/ would not be specified for [round] since that feature is redundant in Chamicuro.

• A relationship can be defined as follows:

  Two rules, A and B, are in a feeding relationship if the result of applying one rule, A, is to increase the number of forms to which a following rule, B, may apply.

  In other words, two rules, A and B, are in a feeding relationship if the output of one rule, A, serves as an input to a following rule, B.

  A ______orders are universal and automatic in nature, such as the RROC, which maximizes feeding relationships between redundancy rules and other rules.

  A ______orders, on the other hand, are non-automatic and thus must be stipulated on a language-by-language basis, such as counterfeeding relationships.

---

Homework for Day 26:

• read Schane, chapter 4 (on reserve)
• redundancy matrices on handout (to turn in)
French fricatives

\[ \begin{array}{cccc}
\text{f} & \text{v} & \text{s} & \text{z} \\
\text{anterior} & + & + & + & - & - \\
\text{coronal} & - & - & + & + & + \\
\text{voice} & - & + & - & + & + \\
\end{array} \]

\[ \begin{array}{cccc}
\text{f} & \text{v} & \text{s} & \text{z} \\
\text{anterior} & + & + & + \\
\text{coronal} & - & - & + & + \\
\end{array} \]

\[ \begin{array}{cccc}
\text{f} & \text{v} & \text{s} & \text{z} \\
\text{anterior} & - & - \\
\text{coronal} & + & + \\
\end{array} \]

\[ \begin{array}{cccc}
\text{f} & \text{v} & \text{s} & \text{z} \\
\text{anterior} & + & + & + & - & - \\
\text{voice} & - & + & - & + & + \\
\end{array} \]

\[ \begin{array}{cccc}
\text{f} & \text{v} & \text{s} & \text{z} \\
\text{coronal} & + & + & + \\
\text{voice} & - & + & + \\
\end{array} \]

\[ \begin{array}{cccc}
\text{f} & \text{v} & \text{s} & \text{z} \\
\text{coronal} & - & - \\
\text{voice} & - & + \\
\end{array} \]

\[ \begin{array}{cccc}
\text{f} & \text{v} & \text{s} & \text{z} \\
\text{coronal} & + & + & + \\
\text{voice} & - & + & + \\
\end{array} \]
Day 25: Redundancy

**Principles of Phonological Analysis**

---

### Pronunciation Vowels

- **i**
- **e**
- **u**
- **o**
- **ü**
- **a**

- **high** + – + – – – + –
- **low** – – – – + – –
- **back** – – – – + – +
- **round** – – – – – + +

[+high] → [-low]
[-low] → [+high]
[+low] → [+back]
[-back] → [-low]
[+round] → [-low]
[-round] → [+back]
[+round] → [+back]

---

### Consonants

- **f**
- **v**
- **s**
- **z**

- **anterior** + + + – –
- **coronal** – – + + + +
- **voice** – + – + – +

[-coronal] → [+anterior]

---

### Rumanian Vowels

- **i**
- **e**
- **u**
- **o**
- **ü**
- **a**

- **high** + – + – – – + –
- **low** – – – – + – –
- **back** – – – – + – +
- **round** – – – – – + +

[+high] → [-low]
[-low] → [+high]
[+low] → [+back]
[-back] → [-low]
[+round] → [-low]
[-round] → [+back]
[+round] → [+back]

Illustrative examples for redundancy (to be used in class)


### French fricatives

<table>
<thead>
<tr>
<th></th>
<th>f</th>
<th>v</th>
<th>s</th>
<th>z</th>
<th>ʃ</th>
<th>ʒ</th>
</tr>
</thead>
<tbody>
<tr>
<td>anterior</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>coronal</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>voice</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
</tbody>
</table>

### Rumanian vowels

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>e</th>
<th>u</th>
<th>ɯ</th>
<th>ə</th>
<th>a</th>
<th>u</th>
<th>o</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>back</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>round</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>
### German plosives (stops and affricates)

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>b</th>
<th>t</th>
<th>d</th>
<th>k</th>
<th>g</th>
<th>pʰ</th>
<th>tʰ</th>
</tr>
</thead>
<tbody>
<tr>
<td>delayed release</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>anterior</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>coronal</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>voice</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

### Vulgar Latin vowels

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>e</th>
<th>ε</th>
<th>a</th>
<th>u</th>
<th>o</th>
<th>ɔ</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>low</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>back</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>round</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

### Turkish vowels

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>e</th>
<th>y</th>
<th>œ</th>
<th>i</th>
<th>a</th>
<th>u</th>
<th>o</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>round</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>back</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

### Arara vowels (Brazil)

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>e</th>
<th>a</th>
<th>u</th>
<th>u</th>
<th>o</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>back</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>round</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Latin stress

<table>
<thead>
<tr>
<th>Monosyllabic words</th>
<th>Disyllabic words</th>
<th>Trisyllabic words</th>
</tr>
</thead>
<tbody>
<tr>
<td>árks</td>
<td>fálanks</td>
<td>ádeps</td>
</tr>
<tr>
<td>dúks</td>
<td>árkis</td>
<td>ápeks</td>
</tr>
<tr>
<td>dáps</td>
<td>dúkis</td>
<td>ékwes</td>
</tr>
<tr>
<td>sé:ps</td>
<td>dápis</td>
<td>filikis</td>
</tr>
<tr>
<td>ré:ks</td>
<td>sé:pis</td>
<td>lápis</td>
</tr>
<tr>
<td>úrps</td>
<td>ré:gis</td>
<td>árbor</td>
</tr>
<tr>
<td>pléps</td>
<td>úrbis</td>
<td>púgil</td>
</tr>
<tr>
<td>lí:s</td>
<td>plébis</td>
<td>nó:men</td>
</tr>
<tr>
<td>fráws</td>
<td>lí:tis</td>
<td>kármen</td>
</tr>
<tr>
<td>fróns</td>
<td>fráwdis</td>
<td>lú:men</td>
</tr>
<tr>
<td>déns</td>
<td>fróntis</td>
<td>wé:ris</td>
</tr>
<tr>
<td>sórs</td>
<td>fróndis</td>
<td>sá:lis</td>
</tr>
<tr>
<td>wé:r</td>
<td>déntis</td>
<td>múrmur</td>
</tr>
<tr>
<td>sá:l</td>
<td>sórtis</td>
<td>áktyo:</td>
</tr>
<tr>
<td>fú:r</td>
<td>bú:bo:</td>
<td>gé:rr:</td>
</tr>
<tr>
<td>fár</td>
<td>népo:s</td>
<td>ínku:s</td>
</tr>
<tr>
<td>ás</td>
<td>fú:ris</td>
<td>fárris</td>
</tr>
<tr>
<td>ós</td>
<td>ássis</td>
<td>óssis</td>
</tr>
<tr>
<td>ó:s</td>
<td>ó:ris</td>
<td>fló:ris</td>
</tr>
<tr>
<td>fló:s</td>
<td>mú:ris</td>
<td></td>
</tr>
</tbody>
</table>

### Classical Arabic stress

1. lán  
   ‘not’
2. fáqat  
   ‘only’
3. málikun  
   ‘a king’
4. malíkatun  
   ‘a queen’
5. malikátuhu  
   ‘his queen’
6. maktábun  
   ‘a library’
7. maktabáthu  
   ‘his library’
8. yastaqúbíluhu  
   ‘he receives him (as a guest)’
9. já:wárahu  
   ‘it bordered it’
10. yu:já:wíruhu  
    ‘it borders it’
11. kássarat  
    ‘she smashed’
12. kasárta  
    ‘I broke’
13. kassárthu  
    ‘I smashed it’
14. kasárthu  
    ‘I broke it’
15. sa:fártu  
    ‘I traveled’
16. qábla  
    ‘before’
17. kitá:bun  
    ‘a book’
18. ká:tibun  
    ‘a writer’
19. kassarú:hu  
    ‘they smashed it’
20. kasarná:hu  
    ‘we broke it’
21. sá:fara  
    ‘he traveled’
22. ka:tibi:na  
    ‘writers’

Sources of these data:


Navaho (USA) – Faith Hill

1. [de:h] ‘tea’ 11. [siłtso::z] ‘it is there (a flat, flexible object)’
2. [de:[giɔ] ‘gapped’ 12. [siła] ‘it is there (a slender, flexible obj.)’
3. [dʒo::l] ‘ball’ 13. [be?:eʒo::?] ‘hairbrush’
4. [didzo::l] ‘spherical’ 14. [de::floh] ‘I will lasso it’
5. [be:j] ‘iron’ 15. [de::lo::z] ‘we carry them straight’
6. [le:h] ‘into the soil’ 16. [le::?] ‘in the soil’
8. [na::ba:hi::] ‘warrior’ 18. [jini::? dini:h] ‘my face hurts’
10. [ʔadin] ‘nothing’ 20. [biga::n dini:h] ‘her arm hurts’

Focus on vowel length. [V] = short vowel; [V:] = long vowel; [V::] = extra long vowel. Tone and consonant length have been omitted. Source of data unknown.
Homework for Day 29:
• read Marlett, chapters 36, 37, 38, 39
• read Suprasegmentals paper (in the course packet)
• Isthmus Zapotec, p. 205 (for practice only)
• Mende, p. 205 (for practice only)

Day 28: The Big Picture

In today’s class we are going to see how to deal with tone from a phonological perspective. We will first define what we mean by a tone language, and then we will describe the features and mechanisms we need for analyzing typical tonal phenomena.

• Definition: a language is one in which different lexical morphemes and/or different grammatical categories are distinguished by the relative level of pitch at which the utterances are pronounced.

In other words, a tone language is one in which pitch is used to signal a difference in meaning of some type (lexical or grammatical). That is, tone is contrastive or phonemic.

For example, Nupe (Africa) has three level tones:
• high [bá] ‘to be sour’
• mid [bā] ‘to cut’
• low [bà] ‘to count’

Mandarin Chinese has four tones (tonal melodies):
• /mā/ 5 ‘mother’ (high level)
• /má/ 35 ‘hemp’ (mid rising to high)
• /mà/ 214 ‘horse’ (falling and then rising)
• /mà/ 51 ‘scold’ (high falling to low)

The traditions for transcribing pitch vary. Some use accent diacritics such as acute for high tone, macron for mid, and grave for low.

Others use numbers, especially when more than three pitch levels need to be distinguished.

In SIL and Americanist usage, 1 is used to mark the highest pitch, whereas in other traditions, the number 1 stands for the lowest pitch.

Tone languages compared to stress languages:
• They are different in at least three ways:

1. In a stress language, it is usually predictable which syllable of the word gets stressed, whereas in tone languages, tone is (generally) unpredictable.

2. In a stress language, pitch may be used for special emphasis, showing the speaker’s attitude. In a tone language, on the other hand, pitch is used for showing lexical or grammatical contrasts (tense, person, number, etc.).

A rough and broad generalization which can be made is that most African tonal languages are register whereas most of the tonal languages in Asia are contour.

An interpretive decision which must be made is whether contour tones are indivisible units or sequences.

Early work in generative phonology assumed that some tone languages had underlying contours in the sense that a falling tone, for example, is an indivisible unit specified as [+falling], rather than as a sequence of two tones (HL).

Today, contour tones are generally seen as a or concatenation of two level tones. This can greatly simplify our analysis of tonal processes.

For example, Mandarin has the following assimilation rule:
ùwà ‘personal name’
èkwé ‘personal name’
àdhá ‘personal name’
ùwà cì àkhwá ‘Uwa carries eggs.’
èkwé cì àkhwá ‘Ekwe carries eggs.’
àdhá cì àkhwá ‘Adha carries eggs.’

• In analyzing tone, it often helps to substitute a series of lexical items that you are studying into a constant or sentential context, where all of the other words are held constant, such as in the Igbo examples above.

• The maximum number of level tone registers claimed for any language is five:

• These combinations allow for only a four-way distinction in tone. One way to distinguish between more than four (if in fact that is needed) is to have both [± hi] and [± lo].

• In some tone languages, especially those with only two tone levels, there is a tendency for the absolute pitches of tones throughout an utterance to decline somewhat. This is known as ___________________. As a result, there may be cases in which a phonemic high tone at the end of a long utterance may be lower in absolute pitch than a phonemic low tone earlier in the utterance.

• What counts as far as conveying the meaning is the relative, not absolute, level of pitch, i.e., the fundamental frequency of syllables in relation to other syllables in the same utterance.

• Igbo downdrift:

• Igbo downstep:

• Some tonal processes:

• Spreading: an underlying tone on one syllable attaches to a neighboring vowel or syllable, often resulting in a contour tone at the phonetic level.

• Gwari: Ọkpa/ [ókpá] ‘length’

• /súknù/ [súknû] ‘bone’

• The generalization behind this process is that any tone spreads to the following syllable. The rule can be formalized as follows: ___________________:

  T1 T2
  V C0 V

  L H
  V C0 V

• In some languages with downdrift, it may also be the case that in certain sequences of high tones, the second high tone may be unexpectedly lower than the first. This is called ___________________. Note that downdrift is a phonetic (predictable) phenomenon, whereas downstep is phonemic.

• Often this process can be traced to the presence of an intervening low tone at some earlier point in the history of the language.

• Some tonal processes:

• Spreading: an underlying tone on one syllable attaches to a neighboring vowel or syllable, often resulting in a contour tone at the phonetic level.

• Gwari: Ọkpa/ [ókpá] ‘length’

• /súknù/ [súknû] ‘bone’

• The generalization behind this process is that any tone spreads to the following syllable. The rule can be formalized as follows: ___________________:

  T1 T2
  V C0 V

  L H
  V C0 V

• In terms of feature geometry, we use an interaction between the Mode Node and the Register Node to distinguish between four different levels of tone:

• Upper Register [+hi] High Mode [+Hi] 1 highest tone

• Lower Register [–hi] High Mode [+Hi] 2

• Upper Register [+hi] Low Mode [–Hi] 3

• Lower Register [–hi] Low Mode [–Hi] 4 lowest tone

• Note that changing the feature [+Hi] to [–Hi] under the Mode Node entails a larger drop in absolute pitch than changing the feature [+hi] to [–hi] under the Register Node.

“No matter how sophisticated, liberal, or open minded a field worker thinks he or she is, the tyranny of what you know, innately or through learning, is awesome — it prevents you from seeing the world, in effect. When you study a language in the field, you do so as a big baby. Like a baby, and like a drowning person, you hungrily seize new linguistic data and associate it with what you know. Only, unlike a baby, you are contaminated with one or more native languages, whose grasp upon the mind is fiercely jealous, blocking your view of any other language system. So if you come to understand your field language primarily in terms of your linguistic knowledge, the overwhelming influence of your native language inevitably wins, no matter how good a field worker you are. For the most part perhaps, it does not matter since the greater part of any language is universal grammar, or so we think. But it is a mistake to be too sanguine about this. We cannot easily avoid mistakes.”
The study of suprasegmentals includes all aspects of the sound systems of language that are not confined to individual segments such as vowels and consonants. This includes larger units of segment organization, such as the syllable, and prosodic features of speech, such as pitch, intensity, and length, which can have a domain of realization larger than an individual speech sound.

### 1. Prosodic features of prominence

Features such as stress, pitch, and duration are properties of the speech signal which are better associated with units larger than the segments. They are “overlaid functions” on segments, and can be produced independently of them.

#### 1.1 Length

Segments may differ in their length, or duration. This is a common phonetic phenomenon, in that some segments take longer than others to produce. For example, tense vowels or low vowels are longer than the corresponding lax or higher vowels. Also, there may be allophonic differences in length conditioned by the environment. In English, vowels before voiced consonants are longer than vowels occurring before voiceless consonants.

Some languages, however, use length as a contrastive feature to make phonemic distinctions. In such cases, the relative length or quantity of the segment is the distinguishing feature. Karok is such a language:

| [tas]  | ‘to build a fence’ | [ta:s]  | ‘species of plant’ |
| [pip]  | ‘to sting’         | [pi:p]  | ‘to say’           |
| [xas]  | ‘then, but’        | [xa:s]  | ‘almost’           |

Here we see a contrast in two degrees of length for vowels. Some languages, such as Southern Mixe (Mexico), even make a three-way distinction:


Differences in length are most commonly noted for vowels, but they may occur among consonants too. In Luganda we find contrasts such as:

| [kúlà] | ‘grow up’ | [kːúlà] | ‘treasure’ |
In cases of long segments, the sound system of a language may treat them as sequences of two like segments, so that [k:úlà] might be interpreted as [kkúlà]. Sequences of double vowels or consonants in such analyses are called geminates.

If long segments are not analyzed as geminates, we use the feature [+long] to distinguish two degrees of length. This is appropriate, for example, in describing allophonic differences in length. The rule lengthening English vowels would look something like this:

\[ V \rightarrow [+\text{long}] / ___ C [+\text{voice}] \]

1.2 Stress

The second prosodic feature that we will consider is stress. Stress refers to the relative degree of prominence placed on syllables within an utterance. A stressed syllable is produced with a greater amount of muscular energy than an unstressed syllable. This is usually (although not always or exclusively) manifested by loudness (greater amplitude of the acoustic signal), but this in itself is not entirely accurate. Many languages manifest stress by lengthening the stressed syllable and/or by higher pitch of the voice on the stressed syllable. Stress may also be manifested in increased muscular activity involved in articulatory movements. Thus, in English, a stressed syllable is often pronounced louder and with a higher pitch; the vowel of the syllable is usually longer, too. The greater muscular activity involved is also realized in features such as the aspiration of voiceless stops at the beginning of stressed syllables.

Although stress is actually realized on the syllable as a whole, it is most evident in the vowel, which has greater intrinsic prominence anyway, because of its resonant qualities. For this reason, linguists assign the feature [stress] to the vowel of the syllable in particular. (To make specific reference to the syllable in descriptions, another notation will be necessary, and this will be presented later.)

Languages sometimes exhibit different degrees of stress. For example, in English, three levels of stress can be distinguished within individual words:

mándatòry or "manda'tory or 'manda,tory

The acute accent on the vowel is used to indicate primary stress, the grave accent is used for secondary stress, and no mark is used for unstressed syllables. Another system of notation (the IPA) uses a raised vertical stroke before the syllable for primary stress, and a lowered vertical stroke before the syllable for secondary stress.

Stress is used with differing linguistic functions. For one thing, languages characterized by stress (called stress languages) use it with a culminative function, marking one (and only one) syllable in each word as having greatest prominence. In these languages, the number of (phonological) words in any utterance can be determined simply by counting the number of primary stresses within it. Beyond this, however, the placement of stress within the word may be fixed, regular, and predictable (hence, nonphonemic); or it may be non-predictable and contrastive (hence,
phonemic). In Russian, stress is partially phonemic; it must be marked on each lexical item separately, since the position of the stress cannot normally be predicted.

<table>
<thead>
<tr>
<th>phonetic representation</th>
<th>lexical (phonemic) representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[múka]</td>
<td>‘torment’ /múka/</td>
</tr>
<tr>
<td>[muká]</td>
<td>‘flour’ /muká/</td>
</tr>
<tr>
<td>[zámok]</td>
<td>‘castle’ /zámok/</td>
</tr>
<tr>
<td>[zamók]</td>
<td>‘lock’ /zamók/</td>
</tr>
</tbody>
</table>

In Chimalapa Zoque, on the other hand, primary stress occurs regularly on the penultimate syllable of the word. (There are some exceptions to this generalization in the form of loan words, which usually retain stress on the same syllable as in the original source language.)

| [ʔá:wat]        | ‘louse’          |
| [hó:ho]         | ‘palm tree’      |
| [hu:kú:ti]      | ‘fire’           |
| [minsukkéʔtpa]  | ‘they come again’|

Therefore stress can be assigned by a rule:

\[ V \rightarrow [+\text{stress}] / \_ \_ \_ \_ \ C_0 \ \ V \ \ C_0 \ \ # \]

The subscripts on the consonants indicate that the number of consonants is immaterial – no consonants or numerous consonants may appear in that position; the rule is interested in the vowels.

Stress placement may sometimes depend on syllable structure or syllable weight. The classic example is Latin, where stress was assigned to the penultimate vowel unless that vowel was a short vowel followed by only one consonant, in which case the antepenultimate syllable was stressed. The following examples (from Hyman, 1975:206) illustrate this fact:

| refer:cit       | (he/she) remade, repaired, rebuilt’ |
| reféctus        | ‘having been remade, repaired, rebuilt’ |
| réficit         | ‘(he/she) remakes, repairs, rebuilds’ |

The stress assignment rule, which is still reflected to some degree in Romance languages and English, could be formalized as follows:

\[ V \rightarrow [+\text{stress}] / \_ \_ \_ \_ \ C_0 \ ( V \ \ C ) \ \ V \ \ C_0 \ \ # \]

\[ [–\text{long}] \]

The parenthesis notation enables us to collapse two rules, the larger expansion applying first. In other words, if the penultimate vowel is short and followed by only one consonant, the antepenult will be stressed; otherwise (if the penultimate vowel is long and/or is followed by 2 or more consonants) the penult will have the stress.
There are two kinds of stress rules that are observed cross-linguistically. The first kind places stress on a certain syllable of the root and the stress stays on that syllable throughout the entire paradigm. Stress rules in many languages work this way:

Seri: Stress the first syllable of the root.
Isthmus Zapotec: Stress the first syllable of the root.
Texmelucan Zapotec: Stress the last syllable of the root.

The following data show this for Seri (with the roots underlined):

[tít] ‘Is it dull?’
[tá:npχ] ‘Did he go home?’
[intaʔá:s] ‘Did you dissolve it?’
[intaʔá:kox] ‘Did you make it large?’

No matter what prefixes or suffixes are added, the stress never strays from this position in such languages.

In other languages, the stress rule refers to the edge of the word (either the beginning or the end). Therefore stress does not always appear on the same syllable of the root, depending on what affixes are added. In fact, it may not appear on the root at all.

Chimalapa Zoque: Stress the penultimate syllable of the word.
Finnish: Stress the first syllable of the word.
Serbo-Croatian: Stress the last syllable of the word.

Other types of grammatical information may also be necessary. In English, stress sometimes depends on whether the morpheme is used as a noun, adjective, or verb.

pérfect (adjective) vs. perféct (verb)
cónvert (noun) vs. convért (verb)

In Spanish, the stress pattern for nouns and adjectives is quite different from the one found for verbs. An excellent discussion of these facts can be found in Harris (1983).

Of course, most languages have exceptions to any general rule of stress placement. One way to handle these is by including the stress of these exceptional words in the underlying representation. Current theories of stress have proposed other means of accounting for irregular stress which are more constrained than direct marking, but we cannot go into these here. While such examples are
often cited to illustrate that stress is “phonemic” (i.e. not completely predictable), they should not deter the analyst from looking for the regular stress patterns in the language.

Current work in metrical phonology accounts for stress by labeling syllables as strong (S) or weak (W), or by other devices which do not make use of the feature [stress]:

```
  S
 / \ /
/   \ /
 W   S
  \  \  \\
 S W S W W * * * *
  \  \  \  \  \\
 metricality metricality
```

Each of these notations indicates that some of the syllables are more prominent than others.

1.3 Pitch

Unless a speaker of a language speaks in a monotone, there will be variations in the wave frequency of the sounds he or she makes. Changes in frequency are perceived as changes in pitch, which is often used for linguistic purposes. Since a man will generally speak at a lower pitch than a woman, it is not absolute pitch that is significant, but the relative pitch of one syllable in comparison with the syllables around it.

1.3.1 Tone

Pitch variations that can affect the meaning of a word are called tones. A language that uses pitch for lexical and/or grammatical purposes is called a tone language. A great number of the world’s languages fall into this category. Nupe (Nigeria) is a tone language, as can be seen from the following words:

- bá [˥] high tone (H) ‘to be sour’
- bā [˦] mid tone (M) ‘to cut’
- bà [˩] low tone (L) ‘to count’

Tone here is being used for a lexical purpose, to contrast the meanings of different words. The three separate levels of pitch are significant, contrastive, and hence phonemic.
Many tone languages also use tone to make changes in grammatical meaning. Notice that in Godié (Ivory Coast) a difference in pitch may indicate a change in aspect or person.

\[
\begin{align*}
\sigma^2 \text{ li}^2 \text{ su}^2 \text{ ka} & \quad \text{‘he’s eating rice’} \\
\sigma^2 \text{ li}^3 \text{ su}^2 \text{ ka} & \quad \text{‘he ate rice’} \\
\tilde{\alpha}^2 \text{ mu} & \quad \text{‘I’m going’} \\
\tilde{\alpha}^3 \text{ mu} & \quad \text{‘you’re going’}
\end{align*}
\]

Note: Tone is indicated here with superscripted integers after each syllable. In this marking system, which may be used in tone systems with four or five lexical tones, the lowest integer (here, 1) represents the highest tone.

Similarly, in Lushai (NE India) tone is used to mark causativity:

\[
\begin{align*}
\text{núj} & \quad \text{‘laugh’} & \text{núj} & \quad \text{‘laugh at’} \\
\text{hèr} & \quad \text{‘be turning’} & \text{hèr} & \quad \text{‘turn something’} \\
\text{ʔàaw} & \quad \text{‘shout’} & \text{ʔàaw} & \quad \text{‘call to’}
\end{align*}
\]

### 1.3.1.1 Register tone languages

Most tone languages of Africa and many of the ones in Mexico are considered register tone languages because the tones that occur in them are best analyzed as level tones phonemically, even though glides may sometimes occur phonetically. All of the languages mentioned in the last section are register tone languages. In Chiquihuitlan Mazatec (Mexico), another register tone language, there is contrast between four tones:

\[
\begin{align*}
\text{tʃha}^1 & \quad \text{‘I talk’} \\
\text{tʃha}^2 & \quad \text{‘difficult’} \\
\text{tʃha}^3 & \quad \text{‘his hand’} \\
\text{tʃha}^4 & \quad \text{‘he talks’}
\end{align*}
\]

Tones 1, 2, and 3 are level in this position, but according to Jamieson (1977:110), tone 4 (1ow) is a “low falling tone when it is the only tone on a syllable which is utterance final.” Phonetically, therefore, the tone on the word for ‘he talks’ is \([ \ddot{o} ]\). This glide, however, is not phonemically significant, so Chiquihuitlan Mazatec can be considered a register tone language.

Other apparent glides in register tone languages are due to two syllables with different tones which are juxtaposed. If there is no consonant between the tone-bearing segments, the glide may be very noticeable. The following data are from Ocotepec Mixtec (Mexico). The phonetic contours in this language are best analyzed as a sequence of level tones.

\[
\begin{align*}
\text{ku:} & \quad \text{[ɨ] ‘is’ derives from /kúu/ (High-Mid)} \\
& \quad \text{(cf. [ʒiʒi] ‘badger’, which has the same tone pattern for the whole word.)}
\end{align*}
\]

\[
\begin{align*}
\text{sa:} & \quad \text{[ɨ] ‘bird’ derives from /sāâ/ (Mid-Low)} \\
& \quad \text{(cf. [bīkō] ‘cloud’, which has the same tone pattern for the whole word.)}
\end{align*}
\]
It is not necessary that the vowel be long in order for there to be more than one tone associated with it. The following data are also from Ocotepec Mixtec. The continuative morpheme can be analyzed as a prefix consisting solely of a high tone that subsequently must be associated with the verb stem. This association results in a glide when the following tone is a low tone:

\[
\text{Continuative} + \text{Stem} \quad ' + /kākū/ \rightarrow [kākū] \quad [˥˦] \quad 'escape' \\
' + /tiβi/ \rightarrow [tiβi] \quad [˥˥˧] \quad 'become inoperative'
\]

1.3.1.2 Contour tone languages. If some of the phonemic tones of a tone language are characterized principally by glides, the language is called a contour tone language. Here the glide itself (as a whole) is as significant as the levels at which it begins or ends. Thai is one such language:

<table>
<thead>
<tr>
<th>Tone</th>
<th>Pronunciation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>low level</td>
<td>/nāa/ [˩]</td>
<td>‘a nickname’</td>
</tr>
<tr>
<td>mid falling</td>
<td>/nàa/ [˩˧]</td>
<td>‘rice paddy’</td>
</tr>
<tr>
<td>high rising</td>
<td>/náa/ [˦˥]</td>
<td>‘younger maternal uncle or aunt’</td>
</tr>
<tr>
<td>high falling</td>
<td>/nàa/ [˥˧]</td>
<td>‘face’</td>
</tr>
<tr>
<td>low falling-rising</td>
<td>/nàa/ [˨˩˦]</td>
<td>‘thick’</td>
</tr>
</tbody>
</table>

Note: Much of this variation may be phonetic detail; thus, Fromkin and Rodman (1978:87) give the five tones as [˩] low tone, [˧] mid tone, [˥] high tone, [˥˨] falling tone, and [˩˦] rising tone, respectively.

Mandarin Chinese is also a contour tone language, as seen from these contrasts:

<table>
<thead>
<tr>
<th>Tone</th>
<th>Pronunciation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>high level</td>
<td>/ma/ [˥]</td>
<td>‘mama’</td>
</tr>
<tr>
<td>rising</td>
<td>/má/ [˦]</td>
<td>‘numb’</td>
</tr>
<tr>
<td>falling</td>
<td>/mà/ [˥˩]</td>
<td>‘scold’</td>
</tr>
<tr>
<td>falling-rising</td>
<td>/mà/ [˨˩˦]</td>
<td>‘horse’</td>
</tr>
</tbody>
</table>

As the discussion of glides in Ocotepec Mixtec showed, surface contour tones can be represented by a sequence of levels. Using such an analysis, it can be claimed that even in contour tone languages all tones are level at the lexical level. The contour tones would then have to be represented (phonemically) with a sequence of levels on a single syllable. For example, the Thai word for ‘thick’ with a low falling-rising tone could be written /naa215/. The assumption that all tones are represented as level tones in lexical forms will prove useful in tone analysis.

1.3.1.3 Restricted tone languages. All the tone languages discussed above are nonrestricted tone languages. In each one, tones are assigned to each syllable on a relatively free basis. The following data from Igbo (Nigeria) show that it too is a nonrestricted tone language:

<table>
<thead>
<tr>
<th>Word</th>
<th>Pronunciation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>àkwà</td>
<td>[___]</td>
<td>‘bed’</td>
</tr>
<tr>
<td>àkwá</td>
<td>[___]</td>
<td>‘egg’</td>
</tr>
</tbody>
</table>
ákwá [¯¯ ___] ‘cloth’
ákwá [¯__ ___] ‘crying’

Many tone languages, however, have restricted distribution of tones. Hyman and Schuh (1974:82) state: “There are probably many types of restricted tone languages. These all have in common the property that the various tones are not freely assigned to syllables as in the Igbo and Nupe cases.”

In some restricted tone languages, tone patterns are assigned to morphemes or words as a whole, rather than to individual syllables. The following data are from Engenni (Nigeria). Although Engenni has phonemic H and L, as well as phonetic HL, only a small percentage of the theoretically possible tone patterns are actually found.

<table>
<thead>
<tr>
<th>Basic Pattern (Melody)</th>
<th>2 syllables</th>
<th>3 syllables</th>
<th>4 syllables</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>ọkò ‘canoe’</td>
<td>èsènì ‘fish’</td>
<td>èbhùfùmù ‘ash’</td>
</tr>
<tr>
<td>H</td>
<td>ígbó ‘fish net’</td>
<td>òlǐló ‘bottle’</td>
<td>óprépó ‘pig’</td>
</tr>
<tr>
<td>HL</td>
<td>ômù ‘house’</td>
<td>èsàrù ‘mosquito’</td>
<td>ãvùrâmù ‘woman’</td>
</tr>
<tr>
<td>HH</td>
<td>ìlò ‘lie’</td>
<td>ávúvò ‘bat’</td>
<td>ákùkùrò ‘cassava’</td>
</tr>
<tr>
<td>LLHL</td>
<td>umè ‘faith’</td>
<td>èkàsì ‘chair’</td>
<td>ègàdìgà ‘fish trap’</td>
</tr>
</tbody>
</table>

Notice that there are only five basic tone patterns. Thomas (1974:17) describes the phonetic form as follows: “Each phonological word bears a set sequence of tones. The sequence is manifested with differing numbers of syllables. Low tone is less stable than high tone and is the first to drop out when the pattern is compressed.”

Other restricted tone languages limit the number and position of tones within a word or morpheme. These are called pitch-accent languages. Pitch-accent languages are like the tone languages discussed above in that tone may be contrastive. At the same time they are like stress, or accent, languages in that only one syllable per word is marked for the contrasting tone(s) in the pitch-accent system. The pitch on the remaining syllables is then predictable from this “accent”.

Karok (U.S.) has three contrasting tones which co-occur with stress – H, F (falling), and L. The pitch on the unstressed syllables is predictable and depends on proximity to stress, etc.

Japanese is another example of a pitch-accent language. Only one syllable in each word is contrastingly marked for the “accent”. This accent is a signal indicating that that syllable has high tone. The following rules specify the pitch on the remaining syllables:

(1) All the syllables preceding the accented syllable are H, and all syllables following it are L.
(2) The first syllable of a word is L if it is unstressed.

The following data illustrate these facts:
/mádo/ → [médô] ‘window’
/hájaku/ → [hájakû] ‘quickly’
/sajonára/ → [sâjónárà] ‘good-bye’
/anáta/ → [ànátà] ‘you’
/tabénakattaroo/ → [tábénákâtáròò] ‘probably did not eat’
/okakekudasaimáse/ → [ökákékúdásimáse] ‘please sit down’
/kanemotí desjoo née/ → [kànémóti désjòò nèè] ‘He must be rich, don’t you think?’

Note that the tone rules have the effect of keeping all of the H’s contiguous within a word – there are no HLH sequences.

1.3.1.4 Downdrift. Pitch levels in some tone languages are fairly constant throughout an utterance. In such languages, a H at the beginning of a phrase will have about the same absolute pitch level as a H at the end of the phrase. In other tone languages, when two H’s are separated by a L, the second H is slightly lower in pitch than the first. This phenomenon is called downdrift. The following example from Twi (Ghana) illustrates the effects of downdrift:

k ò f í  h w è h w é á d ù à n é k à k r á m à n ’ à d á m f ò b á

‘Kofi searches for a little food for his friend’s child.’

Many (although not all) tone languages exhibiting downdrift are like Twi in that L’s as well as H’s undergo lowering. This may lead to situations like that above in which a H at the end of a phrase is actually on a lower pitch level than a L at the beginning of the phrase. Once again, then, it is relative rather than the absolute pitch levels that preserve the lexical tone contrasts.

1.3.1.5 Factors that affect tone

Tone is susceptible to modification from its environment, just like segments. As a result, tones may exhibit allophonic (“allotonic”) and other types of predictable variation. The most common factor that affects the tone of a syllable is the tone of the syllables in its environment. A number of other factors are also relevant sometimes; a more comprehensive list of these factors is given in an accompanying handout.

Tonal alternations that are predictable will be accounted for by phonological rules. For the present, we will content ourselves with representing tone with features that attribute tone to individual segments (typically vowels, the segments on which tone is most noticeably realized). We would thus have rules such as the following:
When the individual segments are irrelevant to the operation of the rule, we often indicate only the tones in a less formal type of rule:

\[
\begin{align*}
V & \rightarrow \text{[low tone]} / \_ (C) \_ \text{[low tone]} \\
\text{[high tone]} & \_ \_ \_ \_ \\
\end{align*}
\]

Here we have a high tone becoming low when it precedes a low tone in the following syllable. Later we will consider another way of representing tone which appeals specifically to the suprasegmental nature of tone: tone will be indicated on a level distinct from that of the segments, and will then be associated with the segments by various means.

### 1.3.2 Intonation

Pitch differences are also used in language to mark off larger phonological and grammatical units (phrases, clauses), and to distinguish between different types of such units. Here the variations in pitch do not affect the lexical meaning of utterances. This phenomenon, which is found in all languages, is called intonation.

English uses a falling pitch sequence at the end of (neutral) declarative phrases:

- John is from Los Angeles.

A rising pitch sequence often indicates a yes-no question:

- Is John from Los Angeles?

If the final syllable is stressed, the entire rising pitch sequence is mapped onto that syllable, resulting in a gliding pitch sequence:

- Is he here?
Stress and other factors interact with intonation in complicated ways. For example, a focus stress on ‘John’ will (among other things) cause the final rising pitch sequence to begin at a higher level:

Is John from Los Angeles?

The interaction between stress and pitch can also be used to disambiguate sentences like the following:

What did you put in my drink, Jane?

As Fromkin and Rodman (1978:88) note, with the first intonational pattern, “the questioner is asking what Jane put in the drink.” With the second, “the questioner is asking whether someone put Jane in the drink.”

Ladefoged (1975:225) states that “all languages use pitch differences to mark the boundaries of syntactic units. In nearly all languages the completion of a grammatical unit such as a normal sentence is signaled by a falling pitch.” This is probably due to the dynamics of airstream mechanisms. There is, however, great variation in how a given language uses intonation to mark different types of sentences. In Seri, for example, while declaratives are marked by a falling pitch sequence that begins at a mid level, interrogatives are marked by a falling pitch sequence that begins at a high level:

ításití ‘Did he fool him?’  ításití ‘He fooled him.’

tí:m ‘Is he sleeping?’  tí:m ‘He’s sleeping.’
Just as stress in English interacts with intonation, tone in many languages also interacts with intonation. The fact that a language makes use of tone for lexical and/or grammatical purposes does not mean that it does not have intonation patterns superimposed upon the basic tone patterns. For one thing, downdrift seems to be comparable to the neutral falling intonation discussed above in relation to stress languages. There have even been claims that downdrift can be found in all tone languages to some extent. As will be shown later, however, in some ways downdrift is different from other intonation contours in that it interacts both with segmental phonological rules as well as with other tone rules.

Let us consider examples of the interaction of intonation with tone. The following data from Vietnamese show that it has lexical (phonemic) tone (it is, in fact, a contour tone language). The tone markings are those used in Vietnamese orthography.

má [˧˥] ‘cheek’
ma [˥˧] ‘ghost’
mả [˨˩˧] ‘tomb’
mã [˧˩] ‘rice seedling’
mà [˨˩˩] ‘but’
mă [˦˥˥] ‘horse’

In addition to this lexical tone, however, Thompson (1965:17) claims that:

“Final syllables accompanied by tones beginning and ending lower than usual signal the end of sentences which are produced with certainty.... If, on the other hand, the final syllable has its tone beginning and ending higher than usual, the sentence expresses some kind of uncertainty or indefiniteness.... Finally, there are syllables both final and medial, which are accompanied by a tone contour that covers far more of the voice range than usual. This intonation signals a variety of exclamatory colorings, but also often signifies the lack of concern on the part of the speaker for the outcome of his utterance.”

A similar interaction between tone and intonation can be seen in the following data from Angal Henang, South (Papua New Guinea):

nípí ípisí ‘They came.’

nípí ípisí pé ‘Did they come?’

nípí ípisí ‘You two came.’

nípí ípisí pé ‘Did you two come?’
A comparison of the declarative sentences demonstrates that tone is phonemic in Angal Heneng, South, as it alone distinguishes the pronouns. In addition to the tone patterns, a comparison of the declarative and interrogative sentences shows more general intonation patterns, also. (The pitch contours given for the declarative sentences are not completely accurate in that they do not show the effects of downdrift.) Notice that in the interrogative forms it is the relative rather than the absolute pitches which are significant. Thus, while a L may be slightly higher than a preceding L due to intonation, a H is significantly higher than a preceding L.
**Mandarin Chinese**

The Mandarin data below illustrate the operation of rules that modify tone. Each expression consists of two or more morphemes. In column A these morphemes are given with the tones they have when each morpheme occurs by itself (in isolation). These sequences can be considered the Underlying Representations, and are somewhat abstract. The phonetic representations of these expressions are given in column B. Formulate the rules needed to account for the data and determine how they must be ordered. Give the derivation of the last utterance (# 11). The diacritics have their standard interpretations: [á] = high tone; [â] = low tone; [ã] = rising tone; and [â] = falling tone.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>lào mà</td>
<td>lào mà</td>
<td>‘old horse’</td>
</tr>
<tr>
<td>yí shië shî</td>
<td>yí shië shî</td>
<td>‘department of medicine’</td>
</tr>
<tr>
<td>hăn shû biào</td>
<td>hăn shû biào</td>
<td>‘household thermometer’</td>
</tr>
<tr>
<td>dòng nân féng</td>
<td>dòng nân féng</td>
<td>‘southeast wind’</td>
</tr>
<tr>
<td>wŏ yĕ yòu</td>
<td>wŏ yĕ yòu</td>
<td>‘I also have’</td>
</tr>
<tr>
<td>lào mà</td>
<td>lào mà</td>
<td>‘old jute’</td>
</tr>
<tr>
<td>hào jiù</td>
<td>hào jiù</td>
<td>‘good wine’</td>
</tr>
<tr>
<td>mài mà</td>
<td>mài mà</td>
<td>‘buy a horse’</td>
</tr>
<tr>
<td>mài lào mà</td>
<td>mài lào mà</td>
<td>‘buy an old horse’</td>
</tr>
<tr>
<td>siăng mái lào mà</td>
<td>siăng mái lào mà</td>
<td>‘wish to buy an old horse’</td>
</tr>
<tr>
<td>lào lì mái hào jiù</td>
<td>lào lì mái hào jiù</td>
<td>‘old Li buys good wine’</td>
</tr>
</tbody>
</table>
Suena (Papua New Guinea)

Acute accent (á) indicates high pitch; grave accent (à) is low pitch; mid pitch is unmarked.

1. [bóà] ‘shelter’ 11. [buà] ‘species of tree’
2. [atù] ‘shovel’ 12. [dai] ‘too’
5. [métà] ‘nephew’ 15. [ógù] ‘charcoal’
7. [aménà] ‘parrot’ 17. [símai] ‘to decorate’
9. [tápinoyà] ‘he finds’ 19. [ziamunoyà] ‘he will curse’
10. [gapinoyà] ‘he pulls up’

Adapted from:

Mixteco of Santo Tomás Ocotepec (Mexico) (restricted)

1. [ 'tu³tu⁵ ] ‘paper’
2. [ 'bi³ko³ ] ‘fiesta’
3. [ 'ka₄zu⁵ ] ‘will burn’
4. [ 'zu³u³ ] ‘straw mat’
5. [ 'ti³ka² ] ‘basket’
6. [ 'bi³ko⁵ ] ‘cloud’
7. [ 'zo⁵so² ] ‘metate’
8. [ 'zu³u⁵ ] ‘stone’
9. [ 'ši⁴ni⁵ ] ‘head’
10. [ 'ka³zu³ ] ‘will cough’
11. [ 'ki³ni⁵ ] ‘pig’
12. [ 'xi¹ka¹ ] ‘far’
13. [ 'ka¹zu³ ] ‘is coughing’
14. [ 'ku⁵nu³ ] ‘meat’
15. [ 'zu⁴u⁵ ] ‘ditch’
16. [ 'ka¹zu⁵ ] ‘is burning’
17. [ 'ba⁴tu⁵ ] ‘belt’
18. [ 'na⁴ma⁵ ] ‘soap’
19. [ 'sa³a¹ ] ‘thus’
20. [ 'xi¹ka³ ] ‘is walking’
21. [ 'ku¹ni⁵ ] ‘wants’
22. [ 'zu¹tu¹ ] ‘carrying rope’
23. [ 'ki⁵si³ ] ‘cooking pot’
24. [ 'sa³a⁵ ] ‘bird’
25. [ 'ža⁵ka² ] ‘garbage’
26. [ 'ku³u³ ] ‘will be’
27. [ 'zi¹ži³ ] ‘badger’
28. [ 'le³xo³ ] ‘rabbit’
29. [ 'ža¹a¹ ] ‘tongue’
30. [ 'ti³ka³ ] ‘grasshopper’
31. [ 'ku¹u³ ] ‘is’
32. [ 'su⁵xa³ ] ‘pitch’
33. [ 'lu³li¹ ] ‘little’
34. [ 'ko³ko¹ ] ‘will swallow’
35. [ 'xi¹te⁵ ] ‘wide’
36. [ ‘ni³i⁵ ] ‘salt’

Assume that 1 is the highest phonetic tone and 5 is the lowest phonetic tone.

Shona tone problem

For Day 30 you can work on the following tone problem from an African language. This is for practice only (not to be handed in). On Day 30 we will give you a discussion of the solution / analysis to this exercise.

Vowels marked with an acute accent (e.g., [á]) are indicating high tone. Unmarked vowels bear low tone. Give the Underlying Representations of the following forms in terms of their tonal specifications, as well as any rule(s) you need to account for the phonetic tonal patterns.

‘he bought .... ’

1. zvumá ákaténgá zvumá ‘beads’

2. badzá ákaténgá badzá ‘hoe’

3. tsamba ákaténgá tsámba ‘latter’

4. cingwa ákaténgá cíngwa ‘bread’

5. cirongó ákaténgá círongó ‘pot’
As we have pointed out, tones undergo changes just like segments do. Tonal alternations may be allophonic and/or morphophonemic, and we will consider both types of effects here.

1. Allophonics

Just as there are procedures for determining contrastive and noncontrastive units with the analysis of segmental units, so with tone. Consider the following data from Villa Alta Zapotec (Mexico):

\[
\begin{align*}
[\text{za}^1\text{i}^2\text{n}] & \quad \text{‘that lard’} \\
[\text{za}^4\text{i}^2\text{n}] & \quad \text{‘that bean’}
\end{align*}
\]

We see that the words \([\text{za}^1]\) ‘lard’ and \([\text{za}^4]\) ‘bean’ above contrast in identical environments; therefore we know that we have at least two phonemic tones. If we find another phrase like:

\[
[\text{za}^2\text{i}^2\text{n}] \quad \text{‘that day’}
\]

we have evidence that there is yet another phonemic tone in Zapotec. This hypothesis is supported by the following data:

\[
\begin{align*}
[\text{de}^4\text{zan}^1\text{j}] & \quad \text{‘a lot of bamboo’} \\
[\text{de}^4\text{zan}^1\text{j}^2] & \quad \text{‘many steam baths’} \\
[\text{de}^4\text{zan}^1\text{j}^4] & \quad \text{‘a lot of weapons’}
\end{align*}
\]

The use of a phrase such as ‘that _____’ or ‘many _____’ is a good way to begin the analysis of tone in a language because it allows you to control the data better.

“In order to control free variation in the general height of pitch and so to observe pitch contrast in the most advantageous circumstances, as the relative contrast from syllable to syllable (contrasts which constitute the basis of the system) words are studied in context rather than in isolation. The contexts studied are chosen from the data available to the investigator. They are so selected that within a single sentence or single phrase one of the words may be withdrawn and another put in its place. Then, in turn, this word is withdrawn and replaced by a third substitution item, and so on, until all the words are studied which can legitimately be found in that place in the particular sentence chosen. Such a sentence is called a FRAME. The replaceable words may be called a SUBSTITUTION LIST. The utilizing of frames for pitch analysis is an extremely important tool for the tone analyst.” (Pike 1947:107)
Now note the following data from Zapotec:

- [la⁴quꜟr³]  ‘leaf’
- [gʷa²ʒɯʔ³]  ‘a fat corncake’
- [zna³]  ‘mother’
- [lo²ʒɯʔ²]  ‘tongue’
- [jerw⁴beʔ³]  ‘mist’
- [ba’duʃ³]  ‘tumpline’
- [zi¹soʔ³]  ‘your water’

It is apparent that tone 3 has a very restricted distribution, contrasting with tone 2 at least (compare ‘tongue’ and ‘a fat corncake’), but not with tones 1 or 4 in these data. If we had more data, we would look for evidence to see if either tones 1 or 4 occur in the environment in which tone 3 occurs. If we were to find tone 3 contrasting with tone 1, but in complementary distribution with tone 4, then tone 3 would be considered a variant of tone 4. We might then write a rule as follows:


The above rule further illustrates the influence that segments and boundaries may have on tones. It is also the case that tones influence other tones. In Cheyenne, which has phonemic H and L, L is raised to M when it precedes a word-final H:

L → M / ___ H #

2. Morphophonemics

In most (if not all) tone languages, it is necessary to posit rules of tone perturbation (or tone sandhi) to account for morphophonemic alternations that may occur. The following data are from Lushai (NE India), which has H, M (mid = ā), R (rising = ē), and F (falling = ê) tone patterns:

- [ʔāká]  ‘he goes’
- [ʔālē]  ‘he buys’
- [ʔābēj]  ‘it sticks’
- [ʔābāl]  ‘it’s dirty’
- [ʔāpēn]  ‘he looks’

When we determine what the morpheme for ‘third person singular subject’ is, we see that we have two allomorphs: [ʔā] and [ʔá]. Examining their environments, we find that they occur as follows:

- [ʔā] occurs before H and R
- [ʔá] occurs before M and F

It is impossible to tell from the data above which is the underlying form of the prefix, but once a decision is made, a tone perturbation rule would have to be written to describe the morphophonemic change that occurs when the prefix is attached to certain stems.
2.1 Natural tone rules

Hyman and Schuh (1974) discuss a number of tone rules which they consider “natural” in that they are found in many African tone languages. A brief discussion of a few of these rules should illustrate the ways in which tone perturbation can be expected to behave.

The first process, spreading, can be illustrated by the following derivations from Gwari (Nigeria):

\[
\begin{align*}
/ \text{òkpá} / & \rightarrow \ [ \text{òkpá} ] \ ‘length’ \\
/ \text{súkù} / & \rightarrow \ [ \text{súkù} ] \ ‘bone’ \\
/ \text{ōzà} / & \rightarrow \ [ \text{ōzà} ] \ ‘person’
\end{align*}
\]

In all of these derivations the tone of the first syllable has “spread” to the second. The second syllable is then realized with a contour tone. This is an assimilatory process in that the second syllable is becoming more similar to the first syllable.

A related process is absorption. Its effects are illustrated by Hyman and Schuh with the following hypothetical derivations:

\[
\begin{align*}
/ \text{ābá} / & \rightarrow \ [ \text{ābá} ] \\
/ \text{ābà} / & \rightarrow \ [ \text{ābà} ]
\end{align*}
\]

This process is similar to spreading in that the tone of the first syllable moves to the right, where it is absorbed by the following tone. Notice that in rules like this our claim that all contour tones are lexically sequences of level tones provides a simpler statement of the process – that is, only the final tone of the first syllable spreads to the right, not the entire contour. This can be symbolized as follows:

\[
\begin{align*}
\text{LH-H} & \rightarrow \text{L-H} \quad \text{and} \quad \text{HL-L} & \rightarrow \text{H-L}
\end{align*}
\]

A third process is simplification. As was mentioned above, Gwari contains a rule creating contour tones. The following data show, however, that these contour tones are simplified preceding a high tone:

\[
\begin{align*}
/ \text{ōjē} / \ ‘cloth’ & + / \text{bmjá lō} / \ ‘is good’ \rightarrow \text{ōjē bmjá lō} \rightarrow \ [ \text{ōjē bmjá lō} ] \\
/ \text{ōzà} / \ ‘person’ & + / \text{bmjá lō} / \ ‘is good’ \rightarrow \text{ōzà bmjá lō} \rightarrow \ [ \text{ōzà bmjá lō} ]
\end{align*}
\]

Simplification is similar to absorption in that it levels contour tones. It is different, however, in that the following tone does not have to be identical to the final tone of the contour, and the contour is sometimes realized as different from either of the component levels (e.g. M from HL or LH).

Another example of simplification is found in Fe Fe (Cameroon). Here contour tones are simplified to nonlow tones before low tones:
One last process found in many languages is downstep. Downstep occurs in languages with downdrift. As was mentioned earlier, in a language with downdrift a H following a L (or a sequence of L’s) will be slightly lower in pitch than a H preceding the L. In such languages, if the low-tone segment is deleted, the second H will still be slightly lower in pitch than the first. This can be seen in the following derivation from Twi:

\[
\text{/ mé } + \text{ ɔ̀ bó / } \rightarrow \text{[ mé'bó ] ‘my stone’}
\]

In this instance, then, downdrift interacts with segmental deletions. In some languages the L which originally conditioned downdrift can no longer be justified in the lexical forms. Downstep in these languages must simply be posited in the lexical forms for the morphemes in question.

Finally, it should be noted that tone perturbation rules, like other phonological rules, can interact with each other. This can be seen in the following derivation from Ngizim (Nigeria):

<table>
<thead>
<tr>
<th>Underlying Representation</th>
<th>/ ná bâkó tlûwáí / ‘I roasted the meat.’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreading</td>
<td>ná bâkó tlûwáí</td>
</tr>
<tr>
<td>Downstep</td>
<td>ná bâkó tlû’wáí</td>
</tr>
<tr>
<td>Simplification</td>
<td>ná bâkó tlû’wáí</td>
</tr>
<tr>
<td>Phonetic Representation</td>
<td>[ ná bâkó tlû’wáí ]</td>
</tr>
</tbody>
</table>

2.2 Boundary and consonantal influence on tone

In the processes discussed above, tone sandhi patterns have been affected by other tones. Both boundaries and consonants can also affect tone sandhi patterns. In many languages, for example, there is a correlation between high tone and voiceless consonants on the one hand, and low tone and voiced consonants on the other. Consider the following data from Nupe (Nigeria):

<table>
<thead>
<tr>
<th>[ pá ] ‘peel’</th>
<th>[ èpá ] ‘is peeling’</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ bá ] ‘be sour’</td>
<td>[ èbá ] ‘is sour’</td>
</tr>
<tr>
<td>[ wá ] ‘want’</td>
<td>[ èwá ] ‘wants’</td>
</tr>
</tbody>
</table>

The L from the prefix seems to be spreading in ‘is sour’ and ‘wants’, creating a LH on the second syllable. In ‘is peeling’, however, the spreading is blocked by the voiceless nonsonorant ([p]). That is, the L of the prefix could not “get through” the voiceless nonsonorant.

A similar phenomenon can be seen in the following data from Ewe (Ghana):

<table>
<thead>
<tr>
<th>[ dà ] ‘snake’</th>
<th>[ dà lá ] ‘the snake’</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ ŋà ] ‘pig’</td>
<td>[ ŋà lá ] ‘the pig’</td>
</tr>
</tbody>
</table>
Here an underlying L-H is realized as M-H unless the L is preceded by a voiced nonsonorant. In other words, the voiced nonsonorants prevent raising of a L.

Notice that in both of these instances it is the consonant that affects the tone of the vowel, and not vice versa. Hyman and Schuh (1974:108) make this claim as follows: “...the interference of segmentals with tone processes is unidimensional: consonants affect tone, but tone does not affect consonants. Thus, there seems to be no example where a low tone causes a voiceless consonant to become voiced.” (emphasis by H&S)

2.3 Morpheme classes

It has been found that tone perturbation may be idiosyncratic in at least two ways. There may be an ad hoc class of words that cause the tone of a following word to perturb, while another class of words with the same tones does not effect the perturbation. Also, there may be an ad hoc class of words whose tones are never perturbed.

Texmelucan Zapotec morphemes (Speck 1978) have been analyzed as belonging to one of two classes – the class that causes certain perturbations (morphemes in this class are marked [+P] and the class that do not [-P]). The classes are completely ad hoc (unpredictable). One perturbation rule is as follows: a L becomes HM following a tone that is a perturber:

\[
\begin{align*}
L & \rightarrow \ H M / V C_0 \text{ ___} \\
[-P] & \quad [+P]
\end{align*}
\]

The effects of this rule are illustrated on the word /gòp/ ‘armadillo’:

\[
\begin{align*}
[-P] & \quad [+P] \\
[tàp \, gòp] & \quad \text{‘four armadillos’} \\
[bìk\‘l \, gòp] & \quad \text{‘brother armadillo’} \\
[k\‘jùp \, gòp] & \quad \text{‘two armadillos’} \\
[bìlj\‘l \, gòp] & \quad \text{‘sister armadillo’}
\end{align*}
\]

The following data illustrate the fact that in Texmelucan Zapotec, [+P] morphemes cannot be perturbed by other perturbers:

\[
\begin{align*}
[+P] & \quad [-P] \\
[zù \, tàp \, gòp] & \quad \text{‘be four armadillos’} \\
[+P] & \quad [+P] \\
[zù k\‘jùp \, gòp] & \quad \text{‘be two armadillos’}
\end{align*}
\]

2.4 Grammatical conditioning

Speck (1978) has given evidence of perturbation in Texmelucan Zapotec that is not conditioned by phonological factors. The four phonemic tone patterns of the language show up on adjectives in the context ‘he is ___’:
'lame' 'skinny' 'hot' 'strong'
'he is ___' [ tîs jù ] [ rìt jù ] [ zīɡ jù ] [ fèrt jù ]

When an adjective is used in a noun phrase, however, the tone of the adjective always shows up as a H.

\[ V \rightarrow H / ___ \ ]_{Adj} \ ]_{NP} \]

'___ person' [ mbék tîs ] [ mbék rìt ] [ mbék zīɡ ] [ mbék fèrt ]
'___ armadillo' [ göp tîs ] [ göp rìt ] [ göp zīɡ ] [ göp fèrt ]
'___ fox' [ bēz tîs ] [ bēz rìt ] [ bēz zīɡ ] [ bēz fèrt ]
'___ cat' [ bîj tîs ] [ bîj rìt ] [ bîj zīɡ ] [ bîj fèrt ]

2.5 Interaction of morphophonemic and allophonic processes

Because of tone sandhi, several frames should be tried out when first working with tone in a language. It is entirely possible that the environment of one frame would cause, say, only three of five phonemic tones to show up. In Cheyenne, for example, H becomes L word-finally:

\[ H \rightarrow L / ___ \ ]\]

Thus, the contrast between H and L is only seen when the verb or noun has a suffix that separates the stem-final syllable from the end of the word. (There is still indirect evidence as to the lexical status of the final tone. As was mentioned in section 1, Cheyenne also contains a rule raising L to M before word-final H. This rule applies before the lowering rule. Thus, word-final L-H becomes M-L, the M “marking” the L as being derived from a H. Furthermore, Cheyenne does not seem to contain word-final H-H syllable patterns.)

In Texmelucan Zapotec, a HF becomes a MR when the syllable on which it occurs is unstressed:

\[ HF \rightarrow MR / \underline{[–stress]} \]

In this environment, then, the contrast between lexical HF and MR is neutralized.

2.6 Restrictions on tone patterns

As was pointed out in section 1, restrictions on the distribution of a given tone may lead to the discovery of tone allophones. In addition to looking for restrictions on the distribution of individual tones, it is necessary to also look for restrictions on the distribution of tone patterns as a whole. Many tone languages, as we have noted, restrict the permitted types of tone patterns. Only by discovering the distribution of tone patterns, as well as of individual tones in a language, can all of the generalizations be accounted for.
Day 30: The Big Picture

In today’s class we are going to discuss the notion of sonority. First we will look at attempts to define sonority in physical terms. Then we will formally define the Sonority Sequencing Principle and see what work it does for us in phonological analyses.

A typical consonant inventory from a hypothetical language:

<table>
<thead>
<tr>
<th>Consonant</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>t</td>
</tr>
<tr>
<td>b</td>
<td>d</td>
</tr>
<tr>
<td>f</td>
<td>s</td>
</tr>
<tr>
<td>m</td>
<td>n</td>
</tr>
<tr>
<td>l</td>
<td>r</td>
</tr>
<tr>
<td>w</td>
<td>j</td>
</tr>
</tbody>
</table>

- Discuss the typical definitions of sonority.
- Some of the earlier attempts focused on the notion of [continuant] of the vocal tract / degree of “stricture.”
- However, this runs into problems when we consider fricatives vs. nasals with respect to the feature [continuant].

Homework for Day 31:
Check Yourself 3 (for practice only)

- The classical model defined the feature [sonorant] in terms of (the possibility for) “spontaneous voicing.”
- However, this too is problematic because they also considered [?] and [h] to be [+sonorant].
- Perhaps the best definition which has been given in terms is the inverse of the degree of build up of air pressure in the mouth behind the point of constriction.
- However, the division between voiced fricatives and nasals is somewhat arbitrary; we need more concrete criteria in making a distinction such as this.
- Other definitions have been proposed along lines:
  - loudness
  - intensity
  - energy
  - as well as in terms as inherent perceptual prominence.

Now what significance does sonority have for phonological analyses?

- The Sonority Sequencing Principle (__________________________) can be defined in three analogous ways:
  1. Within any given syllable, there can be one, and only one, peak of sonority.
  2. Every peak of sonority must correspond to a distinct syllable.
  3. Within the onsets of syllables, sonority must rise, and within the codas of syllables, sonority must fall or descend.

Work through the Serbo-Croatian problem, on p. 173.

- Focus on the alternation between [a] and Ø.
- When we have an alternation between a certain vowel and Ø, it is usually going to be a case of epenthesis.
- The alternative would be in the environment VC___CV, such as in the word family.
- However, since syncope leads to more complex syllable structure (closed syllables), it is less natural than epenthesis, since the latter normally leads to more simple and preferred syllable structures.
Day 30: Sonority

- So here we have a case of ____________, and for this language the default vowel is [a] (comment on the application of Redundancy Rules at this point).
- Stray or “unlicensed” consonants are dealt with in potentially one of four ways:
  1. They might become phonetically ____________________________, such as in the word prism (give the corresponding grid).
  2. Stray Epenthesis — either before or after the stray consonant:
  3. Stray ____________________: automatically ordered last in the list of rules; a universal mechanism rather than being language-specific; a default or “mopping up” type of operation.
  4. _____________________________________________________: in Zoque (Mexico), /j-kama/ → [kjama] “his cornfield”
- So the SSP is a universal option with the language-specific parameters on/off.
- The SSP can be applied differently to onsets and codas, and/or it can be violated only by certain segments, the most frequent one being /s/.
- As an example, consider the sonority grid of the English word straps.
- Some native speakers of other languages might perceive this word as containing two or perhaps three syllables.
- For reasons such as these, some phonologists classify /s/ as being more sonorous than all of the voiced fricatives.
- As Latin evolved into Spanish, the SSP became activated:
  - skola > escuela “school”
  - speculu > espejo “mirror”
  - Stephanus > Esteban “Stephen”

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: *** *

---

Vowel: *
Glide: *
Liquid: **
Nasal: ***
Fricative: ****
Stop: **

C' → V C
C' → C V

---

Vowel: *
Glide: *
Liquid: *
Nasal: *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * * *
Stop: *** *

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

C' → V C
C' → C V

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **

---

Vowel: * *
Glide: * *
Liquid: * *
Nasal: * *
Fricative: * *
Stop: **
<table>
<thead>
<tr>
<th>Category</th>
<th>Symbol 1</th>
<th>Symbol 2</th>
<th>Symbol 3</th>
<th>Symbol 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Glide</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Voiced Fricative</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Voiced Stop</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Voiceless Fricative</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Voiceless Stop</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Example: 
- * * * * * * *orphin
- * * * * * * *orphine

With: a m

B: i ċ a
Universal hierarchy of relative sonority

- low vowels: 17
- mid peripheral vowels (not ə): 16
- high peripheral vowels (not i): 15
- mid interior vowels (ə): 14
- high interior vowels (i): 13
- glides: 12
- rhotic approximants (ɹ): 11
- flaps: 10
- laterals: 9
- trills: 8
- nasals: 7
- voiced fricatives: 6
- voiced affricates: 5
- voiced stops: 4
- voiceless fricatives: 3
- voiceless affricates: 2
- voiceless stops: 1

Source:

AL 4303, Principles of Phonological Analysis
GIAL
Day 30

CHECK YOURSELF 3

True   False  1. At times a phonological rule may insert a morpheme which was not present in the underlying representation of a word.

True   False  2. When two phonological rules are applied in different orders, different surface (phonetic) forms sometimes result.

True   False  3. The following pair of words from a hypothetical language demonstrates a good contrast between high pitch and mid pitch: [xútá] ‘glass’ vs. [xútā] ‘knife’, (where [⇑] bears high tone, [⇑] bears low tone, and [ā] bears mid tone).

4. Among the following segments, circle those which are [+consonantal]:

    h  r  a  ṣ  m  w  k

5. Among the following segments, circle those which are [+sonorant]:

    o  k  r  z  d  w  ŋ

6. Name three segments which are [+strident]:  ______  ______  ______

7. Name the three terminal, binary features dominated by the Dorsal Node in Marlett’s feature tree:

    (a)       (b)       (c)

8. Different surface (phonetic) forms of the same underlying morpheme are called

    ________________________________.
9. Matching: Place the appropriate letter in the blank to the left of the corresponding phonological representation:

1. \( \hat{V} \) 
   (a) \( \hat{V} \)
   (b) \( H \)

2. \( \hat{V} \)

3. \( \hat{V} \)
   (c) \( L \)
   (d) \( L \)

4. \( \hat{V} \)

10. Formalize this rule in terms of feature geometry: an /e/ becomes an [i] and an /o/ becomes a [u] when immediately preceded by one of the following sounds: [k], [g], [x], [ŋ], [j], or [w]. (However, the vowel /a/ does not change at all in this environment.)
Western Apache (USA) – Faith Hill

1. \( [na^6 da^6 ri^6 d^2 i^5 d^5] \) ‘we will work’
2. \( [?i^6 d^6 n^6 d^4 l^a d^2] \) ‘it will shine’
3. \( [dah^6 si^6 da^4] \) ‘he will sit (on something)’
4. \( [?a^1 ko^1 hi^6 ga^1] \) ‘he will come there’
5. \( [ha^6 na^6 hwi^6 ni^1] \) ‘he will deliver us’
6. \( [da^6 hi^6 t^6 o^a] \) ‘we will number’
7. \( [bi^6 yi^6 na^1 ri^1 di^6 n^6 di^1 n^2] \) ‘it will be light within it’
8. \( [bi^6 go^6 ye^1] \) ‘he will worry’
9. \( [na^1 da^6 ri^6 ge^1] \) ‘they will reap’
10. \( [da^6 n^6 ne^z^5] \) ‘they will be tall’
11. \( [na^6 t^6 i^1 da^6 k^h c^6] \) ‘they will think about it’
12. \( [nas^6 zi^h^3] \) ‘I will heal him’
13. \( [ha^6 ni^6 he^6 yo^d^5] \) ‘he will drive them out’
14. \( [na^6 na^1 go^6 de^6 t^a h^6] \) ‘you will be forgiven’
15. \( [ha^6 ri^6 d^2 o^t^6] \) ‘it will spill out’
16. \( [na^6 da^6 go^6 n^6 h^6 k^h a^d^5] \) ‘they will fight’
17. \( [da^6 d^6 n^6 t^a^1] \) ‘you will send them’
18. \( [da^6 ho^6 ze^4] \) ‘they will be called’
19. \( [da^6 no^6 hwi^6 si^1] \) ‘they will name you’
20. \( [na^1 gol^6 3^o^4] \) ‘she will sweep’
21. \( [ya^6 na^1 i^6 d^6 n^6 t^a^5 o^s^6] \) ‘he will give it to him (a flat, flexible object)’

Write a rule to predict stress placement, and then analyze tone. Indicate what the phonemic tones of Apache are, along with any rule(s) needed to account for surface variants (allotones). Also, is vowel length contrastive? 1 = highest phonetic pitch; 6 = lowest phonetic pitch.


© 1975 SIL. Used by permission from SIL (granted by Gayle Sheehan on July 7, 2011) for use in GIAL’s AL4303 Principles of Phonological Analysis course by up to 100 students each year.

(Apache data.pdf)
A list of some of the factors which affect tone

1. The vowel of the syllable:
   a. The vowel /i/ (and sometimes /u/) may cause higher allotones.
   b. Nasal vowels may cause lower allotones.
   c. Breathy vowels may cause lower allotones.
   d. Long vowels may cause lower allotones of a high tone.

2. Type of syllable:
   a. A high tone may upglide or be raised in a closed syllable.
   b. A high tone may downglide in an open syllable.
   c. A high tone may be lowered in a syllable with a long vowel.

3. Adjacent consonants:
   a. Higher allotones may occur preceding a glottal stop.
   b. Higher allotones may occur following a voiceless consonant.
   c. Lower allotones may occur adjacent to a voiced obstruent.

4. Adjacent tones:
   a. A high tone may be raised preceding a low tone.
   b. A high tone may be lowered or upglide following a low tone.
   c. A low tone may downglide following a high tone.
   d. A low tone may have higher allotones between mid and/or high tones.
   e. An upglided tone in an underlying form may become a level tone in a phonetic form.

5. Position within the word or phrase:
   a. A low tone often downglides in a final (prepause) syllable. (High tones seldom
downglide in this environment.)
   b. A sequence of high tones may downdrift.
   c. A sequence of low tones may downdrift.
   d. Within the phrase nucleus, a high tone may have a raised allotone, and a low tone may
   have a lowered allotone or a heavy downglide.
   e. In pre-nuclear position within a word (that is, when in a pre-stressed syllable), a high
tone may be lowered, but a low tone may be raised.

Cite as:
**Artificial Language “J”**

1. sháh ‘to eat’
2. grùb ‘be surrounded’
3. nát ‘spirits’
4. krum ‘to clear space’
5. ğàb ‘to be hot’
6. got ‘to drive’
7. khrau ‘to surround’
8. shi ‘he’
9. ŋañ ‘wanton’
10. ŋáʔ ‘to be loose’
11. ráp ‘to cross’
12. ŋàd ‘to swing’

á = high pitch
à = low pitch
a = mid pitch (unmarked)

Data adapted by Steve Parker, loosely based on an actual language of Southeast Asia.
Day 34: The Big Picture

In today’s class we are going to discuss two different types of unusual characteristics of underlying forms which we sometimes need to resort to: suppletion (irregularity), and abstractness.

- Defined as non-phonologically explainable allomorphy.
- The short, common words are very often suppletive, e.g., in English we have am, is, are, was, were.
- Words which are not quite so frequent tend to get regularized due to analogy, e.g., the plural of cow in Old English was kine.
- At times we simply have to resort to arbitrary (ad hoc) grammatical classes, but we only want to do this as a last resort.
- For example, in English we have a regular phonological rule for plurals of nouns, but we also get ox vs. oxen, brother vs. brethren, foot vs. feet, goose vs. geese, deer vs. deer, sheep vs. sheep, etc.
- Sometimes suppletion arises due to historical factors, e.g., umlaut in German and Old English accounts for foot vs. feet.

- The second person possessive prefix in Chamicuro is suppletive:

<table>
<thead>
<tr>
<th>unpossessed noun</th>
<th>your noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>awahko</td>
<td>p-awahko</td>
</tr>
<tr>
<td>i:la</td>
<td>p-i:la</td>
</tr>
<tr>
<td>ohpana</td>
<td>p-ohpana</td>
</tr>
</tbody>
</table>

- Compare these forms with the following ones:

<table>
<thead>
<tr>
<th>absi</th>
<th>py-absi</th>
<th>‘tooth’</th>
</tr>
</thead>
<tbody>
<tr>
<td>ohki</td>
<td>py-ohki</td>
<td>‘eye’</td>
</tr>
<tr>
<td>elo</td>
<td>py-elo</td>
<td>‘egg’</td>
</tr>
</tbody>
</table>

- So in Chamicuro this morpheme is suppletive; it has the two underlying forms /p-/ and /pi-/, and there is no way to predict which one any particular noun will take.

- Possessive suffixes in Chamicuro:

<table>
<thead>
<tr>
<th>unpossessed noun</th>
<th>my noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>čihpana</td>
<td>u-čihpana-ne</td>
</tr>
<tr>
<td>kinili</td>
<td>u-kinili-ne</td>
</tr>
</tbody>
</table>

- Compare these forms with the following ones:

| čēpeka | u-čēpeka-7e | ‘flower’ |
| čēšana | u-čēšana-7e | ‘deer’ |

- In one related language, the choice between /-ne/ and /-te/ is predictable based on semantic factors, and in a different related language, the choice is phonological: you get one suffix whenever the noun root has one or two syllables, and the other suffix whenever the noun root is three syllables or more long.

- We can define abstractness as positing an underlying representation which is different from its corresponding phonetic representation(s) in some significant way.
- For example, let’s consider Yawelmani:
- The surface vowel system shows some systematic gaps (no long high vowels):

<table>
<thead>
<tr>
<th>i</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td>a</td>
<td>a:</td>
</tr>
</tbody>
</table>

- In this language there is a rule of absolute neutralization which has the effect of converting all underlying long high vowels into mid vowels:

| /i:/ → [e:] |
| /e:/ → [o:] |

- So this is an abstract solution since we have posited underlying segments which never appear as such in surface forms.

- Such types of analyses should be considered highly dubious unless they are very well motivated.
- Another example of abstractness comes from Czech (p. 174):
- The word for ‘brain’ in three different cases is:

| mozek | mošk-u | mošk-ent |

- We can derive both of these allomorphs from the UR /mozk/ via Stray Epenthesis in the first case and Devoicing in the last two.
- This UR is abstract since the underlying form of the morpheme never surfaces in exactly that way; it is always modified in one way or the other.

Now let’s consider the status of the segment [ŋ] in English:

| rum | run | rung |

- Apparently all the nasals contrast in identical environments.
• However, there is a limited or ________________ distribution of [ŋ] since it cannot occur syllable-initially, while [m] and [n] can.
  * * *  
• If [ŋ] is phonemic, there is no apparent explanation for this otherwise ad hoc fact.
• Now let’s look at the comparative suffix:
  big     bigger
  fun     funner
  cold    colder
  dark    darker
• where the suffix is /-r/.
• But cf.
  long    longer
  strong  stronger
• Here we have an alternation between [ŋ] and Ø.
• So either [ŋ] is part of the comparative suffix, which is obviously not true, or else it is being inserted, which makes little sense, or else it is part of the UR’s of long and strong.
• So now we can posit the underlying form /ŋg/, and arrive at [ŋ] through a two-step process of coalescence.
• First the /n/ assimilates by a very general process which is almost universal, and then the /g/ deletes in certain cases.
• These two rules obviously need to be crucially ordered.
• This type of analysis is confirmed by the spoonerism springtime for Hitler > sprigtime for Hiniter.
• This is abstract since we have only one segment on the surface (usually) but two underlyingly.

• Summarize the three different types of abstractness which we have considered today:

  (1) A ____________________________ which never surfaces unchanged in any environment(s).

  (2) A ____________________________ which never surfaces unchanged in any environment(s).

  (3) A ____________________________ of two adjacent segments which never surface unchanged in any environment(s).

• Note the chart in Marlett, p. 61.

### Palauan

<table>
<thead>
<tr>
<th>Present</th>
<th>Future Participle (innovative)</th>
<th>Future Participle (conservative)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. mədənəb</td>
<td>dəŋəbáll</td>
<td>dəŋobl</td>
<td>‘cover’</td>
</tr>
<tr>
<td>2. mətəʔəb</td>
<td>təʔəbáll</td>
<td>təʔibl</td>
<td>‘pull out’</td>
</tr>
<tr>
<td>3. mənətəm</td>
<td>nətəmáll</td>
<td>nətəml</td>
<td>‘lick’</td>
</tr>
<tr>
<td>4. mətəbək</td>
<td>təbəkáll</td>
<td>təbákl</td>
<td>‘patch’</td>
</tr>
<tr>
<td>5. məʔərəm</td>
<td>ʔərəmáll</td>
<td>ʔərəml</td>
<td>‘taste’</td>
</tr>
<tr>
<td>6. məsəsəb</td>
<td>səsəbáll</td>
<td>səsəbl</td>
<td>‘burn’</td>
</tr>
</tbody>
</table>

1. Give the Underlying Representation of each morpheme (insofar as you can determine them).

2. You should note that there is a very obvious and consistent relationship between stress and the vowel [ə]. State this as a phonological rule, in prose only (don’t worry about formalizing it).

3. In this set of data, the location of stress is totally predictable. State the generalization in prose and also try to formalize the rule using the classical notation. That is, your rule should start off like this:

\[ V \rightarrow [+\text{stress}] / . . . . . \]

and then you should fill in the relevant environment(s) in terms of consonants, vowels, and the appropriate word boundary (#).

4. Demonstrate the crucial order of your two rules (Stress Assignment and the rule relating to [ə]) with a correct and an incorrect derivation of the same word.
## Day 36: Orthography Development

### Homework for Day 37:
- Read Marlett, Appendix B
- Read article by Amy Bauernschmidt (in the course packet)
- Katukina exercise, on handout (to turn in on Day 38)

### Orthography Development
- Issues surrounding the development and implementation of orthographies are very important in most SIL field programs.
- However, to be quite frank, the great majority of non-missionary linguists do not deal with orthographies much at all and do not even care about them.
- Some of the factors in orthography decisions are purely phonological in nature:
  - In a phoneme-based orthography, we try to have just one symbol (grapheme) for each contrastive phoneme.
  - In general we do not need to represent sub-phonemic (allophonic) variants.
  - If we do, such an orthography is said to be “over-differentiated”.
  - Representing allophonic features usually only serves to confuse the reader, e.g., in English we do not write aspiration on voiceless stops.

- In some cases of phonemic contrasts which have a low “functional load” (such as stress, tone, θ/ vs. ð in English), we can get by with an “under-differentiated” orthography.
  - For example, in English we do not mark the contrastive stress in minimal pairs such as convert (noun vs. verb) and permit (noun vs. verb).

- In some cases we will need to use (such as ch in English), but in general we should avoid these as much as possible, and stick with a straight 1:1 relationship.
  - In a morphophonemic-based orthography, we have one consistent representation for each morpheme, at least in those cases where its different surface allomorphs are phonologically predictable.
  - An example would be the past tense suffix in English, which we represent with d even when it is actually pronounced as [-t], e.g., baked. Also, s for the plural marker.
  - Share the anecdote about my shorthand class in high school.

- In many cases we will also need to take into account sociolinguistic factors:
  - Dialect variation – in many cases it would be beneficial if we could cover several dialects with a single standard orthography.
  - However, there will often be factors which must be taken into account too.
  - For example, just suggest that we revise the English spelling system and you will see that a whole can of worms is opened up.
  - As another example, discuss the situation of the Quechua orthographies in Peru.
  - In most countries another issue will be of reading skills from the mother tongue into the national language or a trade language.
  - Comment on the use of qui and que vs. ca, co, cu in Spanish-speaking countries.

- There are three basic types of writing systems:
  1. Alphabets are basically phoneme-based
  2. Syllabaries are based on one symbol per syllable, e.g., Japanese
  3. Logographic systems, such as Mandarin Chinese, have one symbol per morpheme.
  - While these can be very difficult to learn, they allow for very fast reading and they also have the advantage that they can be used by speakers of different dialects or even different languages which use the same system.
  - Mathematical symbols are logographic in a certain sense:
    \[ 2 + 2 = 4 \]
  - For long vowels, one way to represent them is with a geminate notation, such as ‘aa’. In many varieties of Quechua a dieresis (ä) on top of the vowel is used.

---

**Day 36: The Big Picture**

In today’s class we are going to discuss some of the typical factors which often must be taken into account when developing a practical orthography for a previously unwritten language.

- In some cases of phonemic contrasts which have a low “functional load” (such as stress, tone, /θ/ vs. /ð/ in English), we can get by with an “under-differentiated” orthography.
- In some cases we will need to use (such as ch in English), but in general we should avoid these as much as possible, and stick with a straight 1:1 relationship.
- In a morphophonemic-based orthography, we have one consistent representation for each morpheme, at least in those cases where its different surface allomorphs are phonologically predictable.
- An example would be the past tense suffix in English, which we represent with d even when it is actually pronounced as [-t], e.g., baked. Also, s for the plural marker.
- Share the anecdote about my shorthand class in high school.

- Factors can also be important to take into account:
  - Native speaker intuition and reaction, which sometimes can also help in our phonological analysis.
  - As an example, share the anecdote about the word jto in Chimúcaro.
  - Because of factors such as these, field workers often do orthography testing before deciding on the final system to use.
  - In general, phoneme-based orthographies are easiest for beginning readers to learn to use, but they are less efficient in the long run when compared with morphophonemic orthographies, cf. divine (/divain/) vs. divinity in English.
  - Conversely, morphophonemic orthographies are harder to learn initially, but they can be more efficient for faster reading for experienced readers, who read for meaning rather than sound.
  - In many field situations another very important factor is what type of hardware and software are available.

- In many cases we will also need to take into account sociolinguistic factors:
  - Dialect variation – in many cases it would be beneficial if we could cover several dialects with a single standard orthography.
  - However, there will often be factors which must be taken into account too.
  - For example, just suggest that we revise the English spelling system and you will see that a whole can of worms is opened up.
  - As another example, discuss the situation of the Quechua orthographies in Peru.
  - In most countries another issue will be of reading skills from the mother tongue into the national language or a trade language.
  - Comment on the use of qui and que vs. ca, co, cu in Spanish-speaking countries.

- There are three basic types of writing systems:
  1. Alphabets are basically phoneme-based
  2. Syllabaries are based on one symbol per syllable, e.g., Japanese
  3. Logographic systems, such as Mandarin Chinese, have one symbol per morpheme.
  - While these can be very difficult to learn, they allow for very fast reading and they also have the advantage that they can be used by speakers of different dialects or even different languages which use the same system.
  - Mathematical symbols are logographic in a certain sense:
    \[ 2 + 2 = 4 \]
  - For long vowels, one way to represent them is with a geminate notation, such as ‘aa’. In many varieties of Quechua a dieresis (ä) on top of the vowel is used.
Chamicuro inventory of phonemic consonants and graphemes:

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>p t k ṭ</td>
<td>i u ii uu</td>
</tr>
<tr>
<td>tʰ č čʰ</td>
<td>e o ee oo</td>
</tr>
<tr>
<td>s š šʰ h</td>
<td>a aa</td>
</tr>
<tr>
<td>m n ū n</td>
<td></td>
</tr>
<tr>
<td>l x λ ly</td>
<td></td>
</tr>
<tr>
<td>w y</td>
<td></td>
</tr>
</tbody>
</table>

Vowels: u’ti unacha’ti chamekalo ukumpale musta kana pewa.

Ticuna is another native language of the Peruvian jungle. It has been analyzed as having five phonemic level tones. Of these, only the two highest tones are marked in the practical orthography, both with an acute accent on top of the basic vowel symbol (V). The three lowest tones are unmarked. The language also has contrastive nasalization and laryngealization (glottalization) possible for all vowels, in combination with any of the five phonemic tones. Nasalization is written with a tilde on top of the vowel (ã), and laryngealization with underlining (a̱). There is also a high back unrounded vowel written as ü. So if this phoneme is nasalized, glottalized, and also pronounced with the highest tone, it would have three diacritics stacked on top of each other, all above the basic symbol, and one underneath.

Iñapari is yet another Amazonian language of Peru that I have studied. It is a pitch-accent language with two tones, high and low. The high tone is marked with an acute accent, while the low tone is unmarked. All vowels can also be contrastively nasalized. Nasalization in this case is written with the French cedilla diacritic underneath the vowel symbol, so that it does not conflict with the high tone accent:

ã

The word /anaw/ means ‘canoe’. It ends with an /a/ that bears high tone and contrastive nasalization. I tried writing this word orthographically as anawá, without indicating the nasal vowel, and my language helper, who had about a fourth grade education in Spanish, told me, “It’s missing the letter n.”
Data and instructions for Katukina problem

<table>
<thead>
<tr>
<th>‘purpose’</th>
<th>‘having’</th>
<th>‘... –ing’</th>
<th>‘past’</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. kai</td>
<td>kaʃo</td>
<td>-----</td>
<td>kaβai</td>
<td>‘go’</td>
</tr>
<tr>
<td>2. aʃani</td>
<td>aʃaʒo</td>
<td>aʃaŋɡi</td>
<td>-----</td>
<td>‘fish’</td>
</tr>
<tr>
<td>3. βanai</td>
<td>βanaʃo</td>
<td>βanakĩ</td>
<td>-----</td>
<td>‘plant’</td>
</tr>
<tr>
<td>4. nii</td>
<td>niʃo</td>
<td>-----</td>
<td>-----</td>
<td>‘hunt’</td>
</tr>
<tr>
<td>5. pii</td>
<td>piʃo</td>
<td>pikĩ</td>
<td>piβai</td>
<td>‘eat’</td>
</tr>
<tr>
<td>6. isini</td>
<td>isiʒo</td>
<td>-----</td>
<td>-----</td>
<td>‘hurt’</td>
</tr>
<tr>
<td>7. tikoni</td>
<td>tikoʒo</td>
<td>tikoŋɡi</td>
<td>tikoβai</td>
<td>‘shoot’</td>
</tr>
</tbody>
</table>

Source of data unknown; adapted by Steve Parker.

Your written homework to turn in on Day 38 consists of formalizing the rules for these Katukina data. Although there is not a lot of data, there are some complicated phonological processes taking place, so let us walk you through the solution step by step. On Day 38 we will hand out a written discussion of the solution. Feel free to work together on this assignment and/or come to the staff for help. Read all of these instructions before you start; this may save you a lot of time and hassle later on (see point # 7 below).

1. Go across the rows and down the columns and make tentative morpheme cuts. As you do so, you should notice several phonological alternations. Make a note of these somewhere on your paper.

2. The best solution for this problem consists of grouping together all the alternations into four processes, which means that you should come up with four rules:

   (a) Perhaps the easiest alternation to analyze is the one between [n] and [ŋ]. This can be accounted for by a Nasal Assimilation rule of the type we have seen many times. State this rule both in prose and formally (using the Feature Geometry type of representation).

   (b) There is also a voicing alternation between the pairs of segments [ʃ] and [ʒ] and [k] and [ɡ]. State the generalization in prose in the broadest terms possible and then formalize it with the minimal amount of tree structure.
(c) Some of the vowels alternate between oral and nasal. Assume that the oral ones are underlying and give prose and formal rules which nasalize them in the appropriate environment (hint: the environment for vowel nasalization crucially relies on syllable structure).

(d) Finally, you need a rule which deletes some underlying nasal consonants in certain environments. Be careful to consider all the forms when analyzing what the relevant environment is.

3. In coming up with these four rules, you might have realized that it is possible to predict all surface occurrences of vowel nasalization. In other words, we can hypothesize that underlingly there are no nasalized vowel phonemes, only oral ones. By relying on the Vowel Nasalization rule and the Nasal Deletion rule, we can even predict the nasalization of the vowels in the ‘having’ and ‘...-ing’ suffixes, getting a “free-ride” (so to speak) on the rules which we have already seen are needed anyway. This would constitute an abstract solution in one of the senses we talked about in lecture recently. This is also a solution of economy since we can then minimize the underlying inventory of vowel phonemes. What we are saying is that the underlying representation of the ‘having’ and ‘...-ing’ suffixes has the shape /-CVC/. Your job is to determine what exactly those consonants and vowels are. In order to convert these underlying suffixes into their corresponding phonetic representations, you will need to modify the Nasal Deletion rule which you posited earlier in the sense that it now needs to apply in two different environments rather than just one. It would be nice if we could express these two environments in one general way, but it simply cannot be done with the theory we have taught you this far. In cases like these we just have to grit our teeth and rely on a “curly brace” notation:

\[
\{ X \} \quad \{- Y \}
\]

4. You should have realized by now that some of your rules need to be crucially ordered. Follow the usual procedures for demonstrating which orders are necessary. A key word to consider would be [aʃaŋgi] ‘fishing’. If you have done your analysis correctly up to this point, you should find that all four of your rules apply in one way or another to this form.

5. Give a list of your rules, by name, arranging them vertically according to the order in which they apply. Connect each pair of ordered rules with an arc.

6. Give the underlying representation of each morpheme. There are a total of eleven different morphemes in these data: seven verb roots and four suffixes.

7. All that you are required to hand in on Day 38 are the formal expressions of the four rules (in feature geometry notation). Finally, smile at the Phonology staff all day.
“Brute Force” Rules

The derivation of ‘Moses’ from ‘Middletown’

I. The naive solution

Given: /mídltawn/ (after stress placement)
(1) ídltawn – Deletion m
(2) ózɨs – Insertion mózɨs

II. A sophisticated solution

Given: /mídltawn/ (after stress placement)
(1) Lateral Deletion mídtawn
(2) Stop Assimilation míttawn
(3) Degemination mítawn
(4) Spirantization mísawn
(5) Rounding míson
(6) Vowel Harmony móson
(7) Syncope móswn
(8) Glide Deletion mósn
(9) Dental Assimilation móss
(10) Schwa - Epenthesis mósis
(11) Intervocalic Voicing mózɨs

Source:

The ideal orthography

by Amy Bauernschmidt

© 1980 Summer Institute of Linguistics, Inc. Copyright held by SIL.

Originally published as:

Reprinted by permission from Nova Lit 5.3 (Dec. 1977). See also: Bauernschmidt 1980
(in Bibliography (Literacy))

Table of Contents

0. Introduction
1. Factors to be considered when forming an orthography
2. General principles in the choice of symbols
3. A few reminders about symbols
4. Recommended symbols for orthographies in Mexico
5. Stages in orthography approval
6. Files and goals

[Topics: orthography, symbolization]

0. Introduction

The best orthography is not necessarily the one which uses one symbol for every element which is in phonemic contrast. The best orthography uses the number of symbols necessary for the native speaker to encode and decode his language with the least amount of frustration. It takes into consideration a balance between natural or deliberate redundancy of symbolization on the one hand, and built-in ambiguity through underdifferentiation on the other. Above all, it should facilitate accurate and efficient communication of information in written form.

But efficiency is not the only requirement for acceptable orthographies. In fact, although the phonological basis of our orthographies is basic, it may be greatly outweighed by pressures from the various cultures involved and the attitudes of the speakers of the language.

One test of a good orthography is that the speaker of the language be able to learn to read and write it. Or to put it another way, the orthography must be teachable. But, Smalley (1963:23) has summarized an alternative viewpoint as follows: “It is not what is easiest to learn, but what people want to learn and use which ultimately determines orthographies.” We need to be open and very flexible in our attitudes as we seek to formulate orthographies.

1. Factors to be considered when forming an orthography

Along with the completed phonological statement, the orthography is one of the major products, of the first months of study, in an indigenous language. This can be an exciting time of
discovering phonemic contrasts, phones in complementation, and other phonological rules. The formation of an orthography to represent the deeper structures which are discovered is almost a creative activity. It is easy to become emotionally involved in the “creation.” Therefore, it is good at all times to keep in mind that there are many factors which influence (and rightly so) the orthography decisions. By the time all have been considered, we seldom find that there are great ranges of choices for the symbols to be used. We should not allow ourselves to become so personally involved in the alphabet making that we are unwilling to accept the suggestions and pressures that come to us from the various outside sources.

1.1. Linguistic factors

An orthography should be based primarily on a good phonological analysis of the phonemes in the language. Many of the orthographic problems which we encounter are due to the lack of a thorough understanding of the phonological system of the language. One should also understand well the suprasegmental and higher phonological or grammatical factors in a language which dictate the rules for word breaks, the use of punctuation and hyphenation, and so forth.

Therefore, a complete phonological analysis of the language is basic and indispensable. It is considered a prerequisite to receiving approval for a working orthography.

1.2. Psycholinguistic factors

1. The magic of written language.

It will be good to remind ourselves that written language will become an integral part of a peoples’ culture. This should be another one of our goals. When a new alphabet has been formulated and presented to the native speaker, there comes into his experience a very powerful tool which can communicate just as he does. It can reflect his speech, thoughts, and tell his folk tales. It may be a threat to him. It is certainly important that he feel he and his community have had a part in its formulation. Beware that it not be considered the “outsider’s alphabet” and, therefore, rejected. It needs to be accepted as a new vehicle for communication or there may be little cooperation in collecting cultural data in the language and, perhaps, even less interest in reading published materials; the content of which comes from the outside. (Consider this a plug for native-authored materials near the beginning of the literacy program.)

2. Native speaker reaction.

Another consideration is the native speaker’s attitude toward the structure of his own language. Does he view a certain sound in his language as being closely related to a similar sound in Spanish? Does he accept the Spanish symbol to represent this sound? Or does he feel that the sound in his language is different enough to merit a distinct symbolization or a modification? Does he have a feel for what constitutes the vowel nucleus; for how tone, stress, or length might be indicated? Note: To establish genuine “native reaction,” one should record the reaction of more than one native! An orthography decision based on “my language helper says …” does not have a broad enough base for decision.
3. Optimal inventory of symbols.

For practical reasons, one must attempt to assess the relative functional load of certain phonemes or suprasegmental phonemes in the language. It is usually not necessary for every feature, which is phonologically contrastive, to be symbolized. A decision to underdifferentiate contrasts must be made carefully and on the basis of testing rather than on hunches or personal prejudice. In the early stages of the literacy program, it is safer to symbolize more of the contrasts and err on the side of redundancy than to underdifferentiate and later have to add more symbols. It is a less painful experience for all concerned (including the linguist) to subtract rather than add symbols.

4. Overuse of symbols.

When choosing symbols, it is also necessary to anticipate the teaching and learning problems that might arise with the multiple use of certain symbols. If there are several digraphs, try to avoid using the same symbol in more than one of these, for example, \((x, xh, dx)\), \((h, ch, ski)\), and \((s, ts, sh)\). In these sets of three, the \(x\), \(h\), and \(s\) are used to represent different values. It is sufficient to use one single letter and one digraph. The second digraph may add too much to the memory load to be useful.

1.3. Sociolinguistic factors

1. Symbol value.

The orthography must be acceptable, above all, to the very ones for whom it is prepared. It must retain those aspects or characteristics which they consider to be most important. If they want it to look like Spanish then, within the limits dictated by the phonological, psycholinguistic, and practical aspects, it should be made to look like Spanish. Are they conscious of their culture and their language having intrinsic value? Then there will be more freedom to use non-Spanish appearing symbols or combinations of them.

2. Adjustments for dialects.

Dialect variations should be taken into consideration. The orthography must be formed with the desire to make it acceptable over as wide a dialect area as possible. It is important to acquire an understanding of the phonological and allophonic variations in closely-related dialects for which this orthography will be responsible. Before the orthography receives established status, perhaps just a few symbol adjustments will make it possible to include the variations found in these other dialects.

3. Unity of language families.

Certainly, it would be ideal if it were possible to use uniform symbolization for all those sounds which are held in common throughout a language family. This should be one of our goals. It has been done the most successfully in the Mixtec family.
4. Prestige, numbers, and so forth.

It is expedient to take into consideration which dialect(s) may enjoy the greatest prestige, which may be the largest numerically, or which may be the most widely understood. One or more of these factors may influence in a parallel fashion which orthography will be most acceptable to the most people.

5. Established alphabets.

SIL workers in new allocations should give consideration to symbols already in use in related nearby dialects. This especially applies to the more recent alphabets as they more accurately reflect the direction that government agencies are going as well as the direction that SIL policies are taking.


Before the orthography becomes established, it should be acceptable to the government agencies with which we have to do.

7. Transfer value.

The orthography should not in any way hinder the native speaker from making the transfer from his own set of orthography symbols to the Spanish orthography. That an orthography should be deliberately formulated to facilitate maximum transfer to the national language is a debatable premise. There may be more important or conflicting goals in some of the other aspects of our own programs.

1.4. Practical factors

In addition to all the other pressures which greatly influence the decisions one makes in regard to symbols, there are some very practical factors to be considered.

1. Ideally, the symbols chosen should be available on the standard Spanish typewriter keyboard, printing presses, or typesetters. The first mentioned is very important as more and more indigenous people learn to write their language and some obtain their own typewriters. We need to realize that it is unfair to all concerned when we use diacritics, tone marks, and other symbols which only we have access to on our personal typewriters or in our own publications department.

2. The symbol should be legible even with diacritics. It should not be so encumbered with extra lines that it is difficult to read. There are limits to how much information one small letter should bear. For example ʰ is too much!

3. The symbol should be easily written in cursive script.
2. General principles in the choice of symbols

In general, it is desirable for the vernacular alphabet in Mexico to look as much like Spanish as possible. An unmodified Spanish symbol is preferred if the sound in the indigenous language is similar to the Spanish sound, or if it is made in the same general area of articulation.

On the other hand, it is not without precedent to use a Spanish symbol to represent a sound in the indigenous language which is different from the sound in Spanish. Obviously, certain letters from the Roman alphabet are used for a great variety of sounds in languages around the world. Each language must be considered unique, and this uniqueness should be partially reflected in the orthographic system. In many instances, especially with languages which have a large inventory of phonemes, it is impossible to construct an alphabet that “looks like Spanish.” Unfortunately, the ultimate decisions in alphabet formulation are still influenced by the pressures which come from the many directions as mentioned above.

3. A few reminders about symbols

1. Diacritics not available on the typewriter should be avoided.
2. The use of Greek letters, upside down letters, or numbers (for segmental phonemes) are not permitted.
3. Punctuation marks should not be used to symbolize phonemes. These are to be used for punctuation (except for apostrophe for glottal stop).
4. The use of strange-looking symbols or diacritics is discouraged.

4. Recommended symbols for orthographies in Mexico

There are three areas of choice for symbolizing those “extra” phonemes. Those mentioned here in parentheses are not recommended at this time for orthographies in Mexico.

1. Use the unencumbered symbol, that is, use any symbol from the Spanish alphabet which has not yet been designated to represent a phoneme in the alphabet being formed, and which is reasonably phonetically similar to the sound under consideration.
2. Use diacritics through the letter, below the line of type, or above the line of type.
3. Use digraphs.

4.1. Acceptable symbols for vowels

A. Diacritics
   1. Bar or slash ɨ, u, ø
   2. Underline ɨ̄, ʉ̄, ø̄
   3. Dieresis ï, ë, ä, ö, ü
   4. Macron ɨ̍, ʉ̍, ø̍, ü̍
B. Digraphs: ae, ea, oe, oa, iu (perhaps others)
C. Comments
   1. The bar or slash written on the letter is probably preferable to the other diacritics. They are preferred by some people in the government agencies with which SIL works, but ø is messy if
accent or tone marks must also be written over the vowel, and ɨ in cursive writing can be confused with t.
2. Underlining is preferable to dieresis if accent or tone are to be written over the vowel, but it can be missed by the reader, especially if it is not placed carefully in handwritten material, and it must be close enough to the bottom of the vowel in typed material so that it does not interfere with letters on the line below.
3. Dieresis is messy if accent or tone marks must also be written over the vowel, or if the vowel is capitalized. It is not very legible on the i. The ɨ should be used only, as in Spanish, to indicate labialization of a consonant.
4. The i with any modification probably needs to be treated as a front vowel, that is, written with qu: quɨ, quɨ, quɨ.
5. Digraphs have the advantage that they look more like Spanish and are easily written, but they may prove to be confusing if there are phonemic vowel clusters in the language, and they can be confused with Spanish vowel sequences.
D. Suggested choices for specific phonemes:
\(/\ören/\) or \(/\ören/\) use u, iu, u, i, u, ī
\(/\ören/\) use o, e, ē, ē, o
\(/\ören/\) use e, ē, ē, ae, ā (e is not acceptable)
\(/\ören/\) use e, ē, ē, a, ae, ea
\(/\ören/\) use a, ā or other modified vowel symbol, according to which Spanish symbol most closely represents the sound as judged by bilingual speakers (schwa and caret are not acceptable).
\(/\ören/\) use u, u, i, ī
\(/\ören/\) use o, ao, ō, o, a, ā, a, oa

4.2. Modification of vowels
A. Nasalization may be written: Vn, Ṽ, V̂, Ŷ (Ṽ, V̂, and V̂ are not desirable).
Comments: Ŷ, Ṽ, V̂ have the disadvantage of not being available on the Spanish typewriter.
Vn is particularly useable if tone is written with numbers dividing the syllables.
ŷ may be confusing to readers if a bar through the letter or a macron is also used.
B. Length may be written: VV, VhV, V̅, Ŷ, Ŷ, (V· and V: are not desirable).
Comment: V· and V: are not to be used because these dots can be confused with punctuation.
C. Tone may be written with:
1. Superscript numbers after the syllable
2. Diacritics over the vowel: á, â, à, â, á, ā (underline for mid or low tone)
Comments: The use of superscript or subscript numbers before the syllable, or subscript numbers after the syllable is not encouraged.
Numbers should be written at the end of the syllable and not before the final consonant of a syllable.
The numbers 1 to 5 (or 6) should extend from high to low. It is advisable to avoid the use of the number 1 and begin with number 2 as high. This has been done successfully in other places and avoids confusion of the 1 with punctuation marks.
In some instances, a single number such as 5 or 6 can be used to indicate a complex tone glide.
If there are two contrastive tones, mark only one. If there are three or more tones which must be marked, then symbolize all of them. It is not a good principle to “symbolize” something which is phonemically contrastive by the absence of a symbol.

D. Ballistic and Controlled syllables may be written: V and VV, or ∪ and V, or Vj and V.
E. Complex vowel nuclei need to be given separate consideration in each language.
Phonemic V?V or V? or ∪ may be written: V’ or Vh or V’V or VhV (VV probably should be used only for length) Phonemic VhV or Vh may be written: VjV or Vj

4.3. Acceptable symbols for consonants

They may be modified by the use of diacritics above or below the line, or by the use of digraphs. Dieresis, acute accent, or underline are acceptable. The best choice depends on the shape of the letter to be modified. Vowels lend themselves more readily to the use of diacritics, while consonants more often can be written with digraphs.

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Suggested symbols</th>
<th>Symbols to be avoided</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>v, b</td>
<td>w</td>
</tr>
<tr>
<td>w</td>
<td>hu, u, v, ü, may use -uh syllable final</td>
<td>w</td>
</tr>
<tr>
<td>θ</td>
<td>d, z</td>
<td>th</td>
</tr>
<tr>
<td>z</td>
<td>z, s</td>
<td></td>
</tr>
<tr>
<td>š</td>
<td>x, si, š, sy, sh</td>
<td>š</td>
</tr>
<tr>
<td>š</td>
<td>x, x, ź, s, š, xh</td>
<td>xr, ź, x</td>
</tr>
<tr>
<td>ž</td>
<td>x, zh, ll, y, x</td>
<td>ź, zx</td>
</tr>
<tr>
<td>ž</td>
<td>x, x, zh, ŋ, ų</td>
<td>ų, zr</td>
</tr>
<tr>
<td>ts</td>
<td>ts, tz</td>
<td>ds, dz</td>
</tr>
<tr>
<td>dz</td>
<td>dz, ds</td>
<td></td>
</tr>
<tr>
<td>tš</td>
<td>ch, tsi</td>
<td>ty, tx</td>
</tr>
<tr>
<td>tš</td>
<td>ch, čh, čh</td>
<td>chr, čh</td>
</tr>
<tr>
<td>dž</td>
<td>dx, dz, ds, ll, dy</td>
<td>dch</td>
</tr>
</tbody>
</table>
4.4. Modification of consonants

A. Palatalization: Ci, Cy, yC, (ñ, ll for specific phonemes)
B. Labialization: Cu, Cü, (avoid if possible Cw)
C. Aspiration: j, h, ŋ (if extra symbol is needed for two degrees of aspiration)
D. Glottalization and glottal stop: C’, CC, h, ‘ (the ‘ is not acceptable)

Comments: For a glottalized consonant, one may want to use apostrophe (C’), and use h for glottal stop between vowels.
For ?C, C?, or C’ one may prefer to use double consonants.
E. Length: CC, Ċ
F. Fortis and lenis
   Underline the fortis member:
      Ť and t
      ņ and n
      ř and l
   Double the fortis member:
      nn and n
      tt and t

4.5. Miscellaneous orthographic problems

A. Punctuation

It is recommended that, in writing the vernacular, all simple punctuation marks be used as in Spanish, that is, period, comma, interrogation marks (¿ … ?), exclamation marks, and, if needed, colon and semicolon. A language which has a question indicator either initially or finally should,
nevertheless, use the Spanish punctuation marks for questions. If there is no other reason, it does make the vernacular look more like Spanish. It should also help the person who is literate in his own language to more readily transfer these punctuation habits to Spanish.

For information on how various punctuation marks are used, see Wares, *Punctuation, proofreading, and printing*. Another available source of information on Spanish punctuation usage is the school textbook series, particularly the workbooks and readers in Spanish. Before teaching literacy classes, it will be valuable to know how the school children are taught to write and punctuate.

B. Proper nouns which can be preceded by a prefix should retain the capital letter, even in the middle of the word.

```plaintext
maPedro  ‘it’s Pedro’
ixTachihuín Dios  ‘the Word of God’
toTazin  ‘our Father’
```

C. Stress should be written, as in Spanish, with an acute accent over the vowel. In case it is necessary to use diacritics over the vowel, the accent may be written with an apostrophe preceding the stressed syllable.

D. Enclitics should probably be written together with the head word. In the case that this voids a stress rule, the enclitic can be written with a hyphen.

E. Loan words

Probably in each situation, the spelling will have to be worked out separately according to the opinions of literate speakers of the language and other cultural pressures. A general principle which has been followed in many languages in Mexico is that loan words, which have become obviously integrated into the vernacular, are spelled with the vernacular orthography. Words with only one or two at the most phonological changes might be better written as in Spanish, since they would look almost like Spanish anyway.

Among the Amuzgos, a poll concerning loan words was taken of some of the youngsters who were in primary and secondary school and who had also attended reading classes in Amuzgo. Their unexpected answer was: “If a Spanish word has been taken into our language then it’s our word and we want it spelled like our language.” So all loan words have been spelled as they are pronounced in Amuzgo.

It appears that there is a trend in some of the government agencies, as well as in some language groups themselves, toward preserving the uniqueness of the vernacular language.

F. Morphophonemic writing

Again, each problem must be decided in its own context. The following is a general rule of thumb taken from a lecture by Ray Gordon: “If allomorphs are grammatically conditioned they
probably should be written phonemically; if they are phonologically conditioned they should be written morphophonemically.” In the latter case the morpheme maintains the same shape in all environments.

Several of the literacy specialists in SIL have suggested that content words be written phonemically but function words might be better written morphophonemically. The functors are the words which are often learned as sight words, therefore, there is great advantage to having the form constant.

5. Stages in orthography approval

There are three stages of orthography development and approval. These will extend over at least three or four years of initial work in the language. This allows time for a number of speakers of the language to become literate in their own language. It allows time for testing of certain symbols and time for feedback of opinion from local people.

After several months in the new allocation, when the major part of the phonology has been described and approved by a linguistic consultant, the translator should make an official formulation of the orthography he plans to use. Either the linguistic consultant or an orthography consultant will be willing to help him make this formulation. It should be recorded on the “Statement of orthography” form and given to the orthography consultant for suggestions and approval. It should be accompanied by a copy of the phonology statement for the files.

5.1. Tentative orthography

At this time it will receive the status of tentative orthography. During this stage it is very flexible. If there are changes of symbolization due to a reinterpretation of the phonology, they are to be submitted for approval along with the updated phonological statement. The orthography will remain tentative, probably for at least a year. Thus, there is sufficient time allowed to establish that it is both readable and “writeable,” and to test the problem areas. It is expected that the results of testing will be tabulated and reports made to the orthography consultant on the conclusions reached. The orthography consultant will be responsible to approve any subsequent changes or symbols and the request for change of status of the orthography.

Extensive publishing is not encouraged during this stage. But it is expected that trial literacy materials, testing materials, and reading materials will be printed or mimeographed for use in teaching.

5.2. Working orthography

When all or most of the troublesome problems of phonology have been resolved and the analysis finalized, the orthography may be classified as a working orthography. This stage is not for testing but for proving that the orthography is adequate. There are no restrictions on the amount of publications during this stage.
5.3. Established orthography

When there have been no changes in the working orthography over a period of three years or more, it may be classified as an established orthography. This implies that not only the linguist, but also the indigenous community is satisfied with the alphabet. It implies that they are using it, are learning to read and write, and that there will probably be no more changes made. The orthography must have established status before a dictionary or the New Testament manuscripts are turned in for publication. It is recommended that the request for established status be made well in advance of that date.

6. Files and goals

The orthography file, which is kept in the literacy office, contains the orthography reports from each language, and all other related correspondence. Members are welcome to use the file to look up the orthographies of languages related to theirs. Literacy consultants find the files useful, in order to have the record of the symbols being used which must be taught in the literacy programs.

Therefore, it is important to all of us to have orthography reports complete, legibly written, and kept updated in our files. We also plan to collect as many phonological statements as possible of the present working and established orthographies and all phonological statements for the present tentative orthographies.

The orthography consultants are responsible to give advice and direction based on linguistic experience, awareness of similar problems in related dialects or languages, and awareness of problems likely to arise in the process of teaching reading and writing. It is important to us to know how the translator is progressing in testing and forming his orthography. The consultants hope to have some input as to how this is done, so that the symbols chosen will be the most accurate, consistent, and useful within the one alphabet being formed, as well as in the larger area of related languages and dialects.

Our ultimate purpose in formulating an orthography is to enable a community of preliterate to become literate in their own language. This goal includes the existence of a vital and ongoing program with enthusiastic readers and writers of their own materials.

Suggested reading on orthography problems:

See also Gudschinsky 1970b Bibliography (Literacy).

See also Hooley 1974 Bibliography (Literacy).


Day 38: The Big Picture

In today’s class we are going to present several specific arguments for why the feature geometry model of writing rules is better than the formalisms of classical generative phonology.

1. The formalism of ________ more directly describes assimilation as a sharing of the same articulatory gesture than does the changing of the value of a feature in a matrix; thus it is physiologically more intuitive.

2. Similarly, Feature Geometry makes natural processes ________ to formalize and unnatural processes more difficult to formalize, even more so than the classical formalisms did.

3. In other words, in the feature geometry tree we have a single feature or node simultaneously shared by two adjacent timing units or segments. So the formalism itself directly shows what is going on in our vocal tracts in a more explicit way.

4. Now there is no direct and simple way to spread a feature which is not already present in the environment.

5. In terms of the evaluation metric of classical generative phonology, the following rule would be considered equally as simple and natural as the voicing rule given above, since it requires the same number of features to be specified:

\[
[+\text{aspirated}] \rightarrow [\sim\text{consonantal}] / ___ [+\text{distributed}]
\]

• which is totally bizarre, and not attested in any language.

6. In terms of the classic model, which just counted up the number of things that have to be stipulated, all of the following rules would be considered equally complex:

\[
\begin{align*}
C & \rightarrow [+\text{voice}] / ___ [+\text{voice}] \\
C & \rightarrow [-\text{voice}] / ___ [+\text{voice}] \\
C & \rightarrow [+\text{round}] / ___ [+\text{voice}] \\
C & \rightarrow [-\text{round}] / ___ [+\text{voice}] \\
C & \rightarrow [+\text{back}] / ___ [+\text{voice}]
\end{align*}
\]

• (3) The arrangement of features allows us to manipulate more than one feature at a time as part of a bundle or node.
• In a language such as Indonesian, which has four primary points of articulation for consonants, a nasal assimilation rule would have to account for the following sequences:

mp  ηf  nt  ns  tiği  gk etc.

mb  nd  ịfị  ịg

• This would require two variable features to capture the relevant places of articulation:

\[
\begin{align*}
C & \rightarrow \[\alpha \text{ coronal}] / ___ C \\
[+\text{nasal}] & \rightarrow \[\alpha \text{ coronal}] / ___ C
\end{align*}
\]

• However, for languages such as Spanish or English, which exploit more point of articulation differences, we would need to include a third feature such as [distributed]:

\[
\begin{align*}
mp  ηf  nt  ns  ịfị  ịg & \rightarrow \[\alpha \text{ coronal}] / ___ C \\
[+\text{nasal}] & \rightarrow \[\alpha \text{ coronal}] / ___ C
\end{align*}
\]

• These innovations began an important movement which has been called “non-linear” phonology, cf. syllable structure, metrical structure, feature geometry, etc. The new views of non-linear autosegmental phonology eventually led to the Feature Geometry concept in the mid 1980’s.

Homework for Day 39:
• read Marlett, chapter 14, and review chapter 13
• review the exercises in Marlett, pages 76-77 (for practice only)
Nevertheless, we really don’t want to say that the nasal assimilation rule for Spanish is any less natural than the one for Indonesian, even though it uses one more feature.

With feature geometry, we can express these two rules in exactly the same way, capturing the fact that what is involved is really the exact same process in both languages:

\[
\begin{array}{c}
C \quad C \\
\mid \quad \mid \\
R \quad R \\
\mid \quad \mid \\
SL \quad SL \\
\mid \quad [+\text{nasal}] \\
\end{array}
\]

This is because when we spread a node such as Place, all of the other nodes and terminal features which are dominated by it come along for a free ride, regardless of whether they are specified explicitly, or just left implicit.

Similarly, when we \( \text{disassociate} \) a node in the tree, all of the features subsumed under it are also automatically deleted. For example, consider a language like Korean, which has contrastive voiced stops, aspirates, and ejectives. These surface unchanged in syllable onsets, but in coda position they are neutralized such that all of them are realized as plain, voiceless plosives. In the classical model, this rule would require a change in three separate features:

\[
[-\text{sonorant}] \rightarrow \left[ \begin{array}{c}
-\text{voice} \\
-\text{spread glottis} \\
-\text{constricted glottis}
\end{array} \right]_{\sigma}
\]

In the feature geometry model, on the other hand, all we have to do is delink the Laryngeal Node in the relevant environment:

\[
\begin{array}{c}
C \quad \sigma \\
\mid \\
R \\
\mid \\
SL \quad L \\
\mid \quad [-\text{son}]
\end{array}
\]

In other words, \( \text{spreading} \) = \( \text{assimilation} \).

Therefore, the model of feature geometry which we use must be set up in such a way that any group of two or more features which can be jointly assimilated by a single rule must be exhaustively and uniquely dominated by the same node somewhere in the feature tree.

These pressures have led to many revisions and modifications of the feature tree, and phonologists are still not in total agreement about which features and nodes (as well as their hierarchical arrangement) will ultimately lead to the ideal tree configuration.

Another way of looking at all this is the following: in the classical model, that fact that features such as [+anterior], [+coronal], and sometimes [+distributed] occurred together over and over again in many similar rules was to certain degree accidental. There was no independent, a priori explanation for why these combinations occurred rather than, say, [+anterior] and [–voice], or [+coronal] plus [–aspirated], etc.

In the feature geometry tree, on the other hand, we have a very principled and systematic explanation for this phenomenon: features like [+anterior] and [Coronal] frequently co-occur in phonological rules or processes since they are all attached to a common node.

(4) The feature geometry mechanism is more ______________________ than the classical notation was.

By more constrained, we mean more restricted and less powerful.

This is important to avoid unnatural abuses of the formalism.

For example, previously, it was possible to group together labials and dorsals but exclude coronals with the feature [–coronal].

However, this is unnatural since there is no language known which would form a natural class including [p] and [k] but excluding [t].

This is predicted by the current theory since now there is no way to formally refer to the absence of a node.

Similarly, features such as [distributed] (and [strident] in some models) have become more constrained by limiting them to coronal points of articulation.

Previously, [–distributed] included bilabials along with alveopalatals, but no known phonological process makes reference to this whole group (to the exclusion of other points of articulation).

(5) Feature geometry more directly explains and predicts why certain segments do and do not _______________ in the languages of the world.

In other words, feature geometry theory more adequately accounts for why the sounds that do occur, can occur, and why the sounds that do not occur, cannot occur.

For example, using a linear-type matrix, imagine a hypothetical language with a new sound that is characterized by the following combination of feature specifications:

\[
\begin{array}{c}
[+\text{round}] \\
[-\text{labial}] \\
\end{array}
\]
• In the classical model there is no way to directly avoid this contradiction, except by fixing it later with a redundancy rule:

\[ [+\text{round}] \rightarrow [+\text{labial}] \]

In the geometric model this is done for us automatically, since now the only possible way to access the feature [round] is by means of invoking the presence of the Labial Node. In other words, it is built into the theory by Universal Grammar to begin with.

• (6) The classical model of assimilation regularly the Obligatory Contour Principle, which says that at the melodic level, adjacent identical elements (or features or specifications) are prohibited.

• This principle is very important since it governs representations on the tonal tier. For example, the Obligatory Contour Principle requires that identical adjacent tones be multiply linked rather than associated one-to-one.

• In other words, there are many important reasons why we do things the way we do.

• A description of a language does not really account for anything, and it does not really explain anything.

• Linguistic formalisms make a claim about what is really happening in processing a language in our minds and in our vocal tracts, and they should help us explain why the world is the way it is.
Over the years in which I’ve taught phonology, I’ve collected the following list of “bloopers” from students’ answers on homeworks, quizzes, and tests. Read them and weep.

Assimilation occurs when a contoid is absorbed by a neighboring contoid.
I note that [h] and [?] occur together in one phoneme, and separately elsewhere.
Only consonants occur in the onset position.
Rule B must have applied first otherwise you could not get rule B to apply since it would not follow an obstruent.
Glottal Stop Deletion: the word-initial glottal stop is deleted when preceded by a nasal in the word-final position of a prefix.
Order of rules is not crucial, except in the word /seg/ where the suffix is applied immediately for [sege] and after rule 2 in [seks] and rule 6 must precede rule 4.
A stop deletes word-finally in the coda between a nasal and an [s], or after a sonorant consonant.
Syllables with voiceless consonants attract stress more than syllables with vowels.
Stress may be placed on the first syllable of a given language.
There are no syllabic vowels (so far in the data).
This contrast in analogous environment is due to assimilation.
Onset is obligatory except syllable-initially.
A glottal is inserted word-initially when the onset is a vowel.
[dl] is occurring vowel-medially.
[t] and [č] contrast in assimilating environment.
Beam Me Up Scotty Rule: when the last two root-final phones are [+voice] and proceed a [–voice] consonant, then a [+nasal] root-final consonant (if present) and the suffix are deleted.

Stress goes wherever it wants. Stress isn’t real.

Q: When a linguist is developing an orthography (practical writing system) for a previously unwritten language, why is it important that he or she determine which segments contrast?

A: There is less margin of error to do individual speech patterns accounted for phonologically scripted. Clearer patterns of non-redundancy occur.

A: So as to determine the stylistic content of a language based on the phonemic content, hence to see which orthographies actually are separate phonemes as compared to allophones.

Q: Give three different criteria for determining which of the allophones of a phoneme should be chosen as the underlying or basic form for representing the phoneme as a whole.

A: The grapheme for whichever allophone that the people whose language this appreciate and like should be taken seriously.

******************************************************************************

THINGS TO PONDER............

-Is there another word for synonym?
-Isn’t it a bit unnerving that doctors call what they do “practice”?
-When sign-makers go on strike, is anything written on their signs?
-Where do forest rangers go to “get away from it all?”
-Why isn’t there mouse-flavored cat food?
-What do you do when you see an endangered animal eating an endangered plant?
-If a parsley farmer is sued, can they garnish his wages?
-Would a fly without wings be called a walk?
-Why do they lock gas station bathrooms? Are they afraid someone will clean them?
-If a stealth bomber crashes in a forest, does it make a sound?
-If a turtle is found without a shell, is it homeless or naked?
-Why don’t sheep shrink when it rains?
-Can vegetarians eat animal crackers?
-If the police arrest a mime, should they tell him he has the right to remain silent?
-Why do they put Braille instructions on drive-through bank machines?
-How do they get the deer to cross at those yellow road signs?
-Why do they sterilize the needle when they are going to give a convict a lethal injection?
-Why did kamikaze pilots wear helmets?
-Is it true that cannibals don’t eat clowns because they taste funny?
Prose rules to formalize

This optional exercise is designed to give you additional practice with formal rules. It will not be collected nor graded. At our study session for the final exam we can work on some or all of these. At the moment there is no answer sheet available.

Assuming the following inventory of phonemic (underlying) segments in a hypothetical language, formalize the phonological processes stated below. Our emphasis at this point should be on feature geometry rules, but we can look at the corresponding classical notation for comparison, if there is time and interest.

1. Voiceless oral stops become voiced between vowels.
2. \(/n/\) becomes \([\eta]\) before velar stops.
3. Phrase-final vowels become voiceless.
4. A word-final obstruent is deleted.
5. \(/p/, /t/, /k/, and /q/\) become aspirated at the beginning of a word.
6. \(/t/\) becomes \([\check{c}]\) before \(/i/\) and \(/e/\); also, \(/d/\) becomes \([j]\) in the same environment.
7. Oral stops become voiced after a nasal consonant which is the final sound in the preceding morpheme.
8. Stress the final vowel of a word. (Some words end in one consonant; all others end in a vowel.) [classical notation only]
9. Insert the vowel [e] between two consonants.

10. Vowels are lengthened before voiced obstruents.

11. Between vowels and/or oral glides, /b/, /d/, and /g/ become fricatives.

12. High vowels become mid immediately preceding a morpheme-initial /q/ (a uvular stop).

13. Stress the second vowel in every word. [classical notation only]

14. /i/ and /e/ become rounded (but remain front) when followed by a rounded vowel, even if there is a consonant between the two vowels.

15. High and mid vowels become their corresponding lax vowels in closed syllables (syllables ending with a consonant).

16. A nasal consonant takes on the same point of articulation as (becomes homorganic to) a following oral consonant.

17. Obstruents devoice word-finally.

18. A high tone becomes a mid tone following a low tone.
Day 40: The Big Picture

In today’s class we are going talk a little bit about acoustic phonetics, and how it can help us in doing phonological analysis. We will illustrate by demonstrating some nice features (characteristics) of the Praat program.

Some nice features of

1. availability-can be freely downloaded from the following website: //www.fon.hum.uva.nl/praat/-both PC and Mac versions are available.
2. technical support-a Praat users’ group exists at the following site (enrollment is free): //uk.groups.yahoo.com/group/praat-users/-all 3183+ posts and responses are archived there, dating back to its creation in the year 2000
3. popularity-Praat appears to be the most widely-used and quoted acoustic analysis program in the world
   -in the two issues of the Journal of the International Phonetic Association published in 2006, among those articles which identified the software used in running experiments and/or analyzing phonetic data, out of 5 references, one was to the program PitchWorks, one was MacQuirer, and three were Praat.
   -in a similar survey of the three issues of the Journal of Phonetics for the year 2006, out of 11 references, 3 were to ESPS Waves, 2 were Kay CSL Elemetrics, and 6 were Praat.
4. input-can read and use sound files up to 2 gigabytes in size (approximately 3 hours of recorded speech).
5. output-can write selections to the following types of file formats: AIFF, AIFC, WAV, Next/Sun, and NIST.
6. help files-in addition to a general user’s manual, Praat includes learning modules about the following topics:
   (a) statistics
   (b) neural networks
   (c) Optimality Theory
   -it also has a general tutorial as well as tutorials specifically for writing formulas and scripts
7. user-speech synthesis-listening experiments (identification and discrimination tests)
   -its speech analysis functions include:
   (a) jitter, shimmer, and voice breaks
   (b) cochleagrams
   (c) excitation patterns
   (d) it can show voicing pulses
8. speed-very rapidly displays pitch, intensity, and formant tracks, individually or simultaneously
9. graphic displays-can create formant scatterplots as well as various kinds of diagrams, charts, and tables
10. customization-can use custom scripts to automatically perform frequently repeated tasks
11. extracting data-can use log input files to automatically extract data from prepared files

-many of these were personally answered by Paul Boersma, one of the co-authors of Praat

Day 40: Praat and acoustic phonetics
Dear Mom and Dad:

I know you worry about me a lot since I'm out here in the jungle, but as I've assured you before, I'm really in a quite pleasant and safe environment. Just to put your imaginative minds at ease, I've compiled some typical scenes from my everyday life to show you what the conditions here are like.

The Men's Dorm, where many of the bachelors on our center live, is often referred to as the Home for Unwed Boys (HUB). There we prepare exquisite native foods and keep the house in the most tidy of conditions. As you can see in this first picture, it was my turn one week to prepare breakfast for all of us guys.

The founder and president of the HUB is my close friend and associate Owen Blickensderfer. In this next shot you can see him playing with one of the many house pets we like to keep around.
We often find that after a hard day's work, the surrounding vegetation has grown so much that we have to discover a new route to get home. The big, strong, handsome chap you see helping me is the famous pioneer missionary trailblazer, Jeff Scott.

The food here is quite nourishing. On days when we can't think of anything else to fix, we usually fall back on our hearty recipe for stew.

To get our daily exercise, we often go down to the lake for a refreshing afternoon swim. Here you see Owen enjoying himself at the old swimming hole.