Class 5: Principles of rule application

To do
- Read Anderson chs. 9, 10 (SQs due Tuesday).
- Finish Hakha Lai assignment (due Tuesday).

Overview: Multiple application
The basic problem to be dealt with today is what to do with a form that, for some rule $A \rightarrow B / X_1 Y$, contains multiple instances of $XAY$, either because $XAY$ straightforwardly occurs twice in the form, or because there are multiple ways of interpreting $XAY$ (e.g., it contains parentheses).

1. Multiple, non-overlapping matches
SPE p. 344: “To apply a rule, the entire string is first scanned for segments that satisfy the environmental constraints of the rule. After all such segments have been identified in the string, the changes required by the rule are applied simultaneously.”

Example: Remember Palauan vowel reduction from last time? Add to our previous rules one that stresses the penult:

stress: $V \rightarrow [+stress] / \_ \_ (C_0 V) C_0 #$

reduction: $[V] \rightarrow \_\_\_-

shortening: $[V] \rightarrow [-\text{long}]$

o How would those rules apply to an underlying representation like /ʔabínjal/?

2. Multiple matches: one instance’s target is another’s environment
Example: optional schwa deletion French (data originally from Dell 1970\textsuperscript{1})

/suvõnir/ $\rightarrow$ [suvõnir] or [suvnir] ‘to remember’
/pasõra/ $\rightarrow$ [pasõra] or [pasra] ‘will pass’
/parvõnir/ $\rightarrow$ [parvõnir] *[parvnir] ‘to reach’
/suflõra/ $\rightarrow$ [suflõra] *[suflra] ‘will blow’
/āri#dõve#partir/ $\rightarrow$ [āri#dõve#partir] or [āri#dve#partir] ‘Henri had to go’
/3ak#dõve#partir/ $\rightarrow$ [3ak#dõve#partir] *[3ak#dve#partir] ‘Jacques had to go’

o Write a rule for schwa deletion (assuming that these data are correct!).

What does the quote from SPE above predict for this form: /ty#dɔvne/ ‘you were becoming’

Actual result is (supposedly—I’ve gotten different reactions from French speakers) [ty#dɔvne] or [ty#dvne] or [ty#dɔvne], but not *[ty#dvne]—discuss.

Example from Colin Wilson (seen in your study questions): Woleian

(Austronesian language from the Federated States of Micronesia with 1,631 speakers)

/mata/ mate ‘eye’ /mata+i/ metai ‘my eyes’
/mata+mami/ matemami ‘our (excl.) eyes’
/yafar/ yefar ‘shoulder’ /yafar+ai/ yaferai ‘our (incl.) shoulders’
/parasa/ perase ‘switch’ /parasa+rasa/ peraserase ‘splash-intrans.’
/marama/ merame ‘moon’ /marama+li/ maremali ‘moon of’

final-V raising: \[ V \rightarrow \{low\} / \_ \_ \# \]
dissimilation: \[ V \rightarrow \{low\} / \_ \_ \ C_{0}^{\{cons\} +\text{low}\} \]

What does the quote from SPE above say should happen to /marama+li/?

3. Possible solution I: directional application

Left-to-right: Scan the string for the leftmost eligible segment and apply the rule to it. Then scan the resulting form for the leftmost eligible segment, etc.

Does this work for Woleian? French?

Right-to-left: Same thing but start with the rightmost eligible segment.

Does this work for Woleian? French?

4. Possible solution II from Anderson (1974)\(^3\)

- Find all segments eligible for the rule and circle them.
- For each circled segment, underline the smallest environment that lets the segment meet the rule’s structural description.
- If the rule is optional, you may uncircle some of the eligible segments and de-underline their environments.

---


Ling 200A, Phonological Theory I. Fall 2005, Zuraw/Heinz
• If any circled segment is contained in some other circled segment’s underlined environment, uncircle (de-underline the environments of) as few segments as possible to get rid of these overlaps.
• Now apply the rule simultaneously to the remaining circled segments. (Of course, circling and underlining themselves have no theoretical status—this is just a convenient way to say that we are identifying two different types of thing, targets and environments)

○ What does Anderson’s proposal predict for the French string /ty#vudre#k#s#v#l#b#ado#/

‘you would like that what the beadle…’?

○ Does Anderson’s proposal help with Woleianian?

5. Tonkawa revisited (you read about it in K&K)
(Coahuiltecan language once spoken in Texas. Today’s Tonkawa people are based in Oklahoma.)

\[
\begin{array}{llll}
\text{/picena/} & \text{picn+o?} & \text{‘he cuts it’} & \text{picn+n+o?} & \text{‘he is cutting it’} \\
& \text{we+pcen+o?} & \text{‘he cuts them’} & \text{we+pcena+n+o?} & \text{‘he is cutting them’} \\
& \text{ke+pcen+o?} & \text{‘he cuts me’} & \text{ke+pcena+n+o?} & \text{‘he is cutting me’} \\
& \text{picen} & \text{‘castrated one; steer’} & & \\
\text{/notoxo/} & \text{notx+o?} & \text{‘he hoes it’} & \text{notx+n+o?} & \text{‘he is hoeing it’} \\
& \text{we+ntox+o?} & \text{‘he hoes them’} & \text{we+ntoxo+n+o?} & \text{‘he is hoeing them’} \\
& \text{ke+ntox+o?} & \text{‘he hoes me’} & \text{ke+ntoxo+n+o?} & \text{‘he is hoeing me’} \\
& \text{notox} & \text{‘hoe’} & & \\
\text{/netale/} & \text{netl+o?} & \text{‘he licks it’} & \text{netle+n+o?} & \text{‘he is licking it’} \\
& \text{we+ntal+o?} & \text{‘he licks them’} & \text{we+ntale+n+o?} & \text{‘he is licking them’} \\
& \text{ke+ntal+o?} & \text{‘he licks me’} & \text{ke+ntale+n+o?} & \text{‘he is licking me’} \\
\text{/naxace/} & \text{naxc+o?} & \text{‘he makes it a fire’} & \text{nxase+n+o?} & \text{‘he is making it a fire’} \\
& \text{we+nxac+o?} & \text{‘he makes a fire’} & \text{we+nxace+n+o?} & \text{‘he is making a fire’} \\
& \text{ke+nxac+o?} & \text{‘he makes me a fire’} & \text{ke+nxace+n+o?} & \text{‘he is making me a fire’} \\
\end{array}
\]

○ Recall K&K’s syncope rule Ch. 3: \( V \rightarrow \emptyset / CVC_\_CV \)

○ If we simplify the rule to \( V \rightarrow \emptyset / VC_\_CV \), what problems do we run into?

6. Minimal vs. maximal application
Something to think about in these and other cases of potential multiple application: is the rule applying as often as possible or as seldom as possible? Is this something we might want a theory to make reference to?
7. More than one target because of an abbreviatory convention

English stress in verbs and adjectives (there are exceptions, and there’s also a lot more to it…)

<table>
<thead>
<tr>
<th>eváde</th>
<th>ovéid</th>
<th>tormént</th>
<th>tomcént</th>
<th>rélish</th>
<th>jéliʃ</th>
</tr>
</thead>
<tbody>
<tr>
<td>supréme</td>
<td>sapjím</td>
<td>éléct</td>
<td>alékt</td>
<td>cóvet</td>
<td>kʰəvət</td>
</tr>
<tr>
<td>cajóle</td>
<td>kʰədʒóul</td>
<td>exist</td>
<td>øgzíst</td>
<td>devélop</td>
<td>døvélop</td>
</tr>
<tr>
<td>defý</td>
<td>dəfái</td>
<td>adápt</td>
<td>adəpt</td>
<td>stólid</td>
<td>stólid</td>
</tr>
<tr>
<td>caróuse</td>
<td>kʰəmáuz</td>
<td>collápse</td>
<td>kʰəlæps</td>
<td>cómmon</td>
<td>kʰúmən</td>
</tr>
<tr>
<td>confide</td>
<td>kʰənfáið</td>
<td>exháust</td>
<td>øgzíst</td>
<td>clandéstine</td>
<td>klaendéstin</td>
</tr>
</tbody>
</table>

- Formulate generalizations above the three columns. It may help to think of the traditional English distinction between long vowels ([eɪ, i, aɪ, oʊ, aʊ, u]—treat as [+long]) and short vowels ([æ, e, i, a, ə, o]—treat as [–long]).

- Let’s translate the generalizations into a rule schema.

- Expand the schema into rules.

- Try applying the rules to a word from the rightmost column—what problem could arise if the rules weren’t disjunctively ordered?

➤ Reminder: When a schema with parentheses is expanded, the resulting rules are disjunctively ordered. The $n^{th}$ subrule applies only if the $(n-1)^{th}$ was not applicable.

Because schemas with parentheses expand like this:

$$A \rightarrow B / X (Y) _-_Z$$

$$A \rightarrow B / XY _-_Z$$

$$A \rightarrow B / X _-_Z$$

and not in the reverse order, we could say that these schemas are greedy: they look for the longer match first.
8. **Tricky case from Latvian, if we have extra time**
(Indo-European language from Latvia with 1,500,000 speakers):

- **glide formation:**
  \[
  \begin{align*}
  \text{[-cons]} \rightarrow \text{[-syll]} / \_ & \_ \_ \_ \text{[+syll]} \\
  \text{[+high]} \rightarrow & \_ \_ \_ \_ \_ \_ \_ \_ \_ \\
  \end{align*}
  \]

- **truncation:**
  \[
  V \rightarrow \emptyset / \_ \_ \_ \_ \# \]

- First, remember the special convention about the + boundary: /__Y\ is really /__+(Y). That means that every rule is really a schema (can you see how?)!

- Apply the rules to these cases and discuss:
  - /iäi+a/ ‘rides’
  - /kur+ra/ ‘basket (gen. sg.)’
  - /au+ia/ ‘puts on (footgear)’

9. **Leftover item: some ways to test theories about what the learner values (besides typology)**

Each has its methodological strengths and weaknesses:

- Study errors in L1 and L2 acquisition: perhaps common errors reflect grammars that are preferred over the target grammar, and rare or nonexistent errors are rare because they reflect dispreferred grammars
- Give subjects artificial-language-learning tasks (Guest, Dell & Cole 2000; Pater & Tessier 2003; Pycha & al. 2003; Wilson 2003): try teaching subjects various “languages” and see which are easiest for them to learn
- Probe subjects’ existing grammar with novel cases, in an attempt to see which grammar, out of multiple possibilities, they really have (novel language games: Treiman 1983, Derwing & al. 1988, Pierrehumbert & Nair 1995; novel types of “wug” word: the “plural of Bach test” proposed by Lise Menn [Halle 1978], Zuraw 2005)
- Study results of naturally occurring extensions of the grammar to novel cases, as in literary invention (alliteration, puns, imperfect rhymes: Minkova 2001, Fleischhacker 2002b, Steriade 2003, Kawahara 2005) and loanword adaptation
- Compare the processing of two rules within a single grammar (Zhang & Lai in progress)—given that both have been learned, which is easier to use?
- Use a standard wug test (i.e., realistic wug words) to compare the applicability of two or more generalizations (Pertsova 2004)—did speakers learn them all equally well?

---

4 Actual outcomes: [jaj], [kurwja], [auj].
10. **Vignette based on leftover item: another case where computing brevity is hard**

Gildea & Jurafsky\(^5\) used algorithms to try to learn rules in the form of *finite-state transducers*, like this one, which approximates English tapping:\(t \rightarrow r \left[ \begin{array}{c} V +\text{stress} \end{array} \right] [r]_0^- V :\)

There are many interesting points in the theory, but one especially striking one is that one version of the algorithm learns this transducer instead:

The two transducers are equally brief, even though the rules they represent aren’t:

\[ t \rightarrow r \left[ \begin{array}{c} V +\text{stress} \end{array} \right] (r)^* \quad \text{vs.} \quad t \rightarrow r / \left[ \begin{array}{c} V +\text{stress} \end{array} \right] (r)^* \]

[there’s a good chance I didn’t get the second rule exactly right!]

(Which one is right for English? Consider what the second transducer would do on *sky-writing* and *gyrating*, which weren’t in its training set.)

---