## Class 12: Markedness and allomorph choice

## To do for next time

- Start revised Catalan assignment


## 1. Phonologically based allomorph selection

French again-alternative view
(based on discussion in Joan Mascaró (1996) External allomorphy and contractions in Romance, Probus 8, 181-205)

| 'friend' | 'husband' |  |
| :---: | :---: | :---: |
| [bel] ami | [bo] mari | pretty, nice |
| [nuvel] ami | [nuvo] mari | new |
| [vjej] ami | [vjø]mari | old |
| [sct] ami | [sø] mari | stupid |
| [kel] ami | [ kcl$]$ mari | what |
| [3oli] ami | [3oli] mari | pretty |

Despite the regularities in the correspondences between the two allomorphs, we wouldn't want to try to derive them both from a single UR. Rather, both could be listed, with the choice depending on phonological factors (whether the following word begins with a consonant or a vowel).

- What do you think about the hybrid cases of French? Compatible with allomorph-listing?


## Korean

|  | 'baby' | 'water' |
| :--- | :--- | :--- |
| subject | ko.gi-.ga | mu.R-i |
| object | ko.gi-.rtl | mu.R-tl |
| topic | ko.gi-.ntn | mu.R-mn |
| 'it's ...' | ko.gi-.je.jo | mu.R-i.e.jo |

## Other examples

- English a/an, [ $\left.\mathrm{\partial}_{\wedge}\right] /\left[\mathrm{\partial i}_{\mathrm{i}}\right]$
- Italian il/l'
- English more $X / X e r$
- English deverbal-noun -al/-ment


## 2. Allomorph selection as TETU

The key proposal of OT is that there can be output-oriented constraints that are violable, but nonetheless active.
Given a theory (or at least a partial inventory) of markedness conditions, we expect them to pop up here and there, even in languages that seem to violate them rather freely.

Reduplicative TETU: CORR-IO >> PHONO >> CORR-BR
Allomorphic TETU: CORR-IO >> PHONO, but the lexicon sometimes supplies an allomorph that satisfies PHONO:
(again, from Mascaró)

| $\{/ \mathrm{b} \varepsilon 1 /, / \mathrm{bo} /\}$ ami | MAX-C | DEP-C | NOCODA | ONSET |
| ---: | :---: | :---: | :---: | :---: |
| $\rightarrow$ bel ami |  |  |  |  |
| bo ami |  |  |  | $*!$ |


| /30li/ ami | MAX-C | DEP-C | NOCODA | ONSET |
| ---: | :---: | :---: | :---: | :---: |
| 30lit ami |  | $*!$ |  |  |
| $\rightarrow$ 30li ami |  |  |  | $*$ |


| \{/bel/, /bo/\} такі | MAX-C | DEP-C | NoCodA | Onset |
| :---: | :---: | :---: | :---: | :---: |
| bel такі |  |  | *! |  |
| $\rightarrow$ bo такі |  |  |  |  |


| $/ \mathrm{kel} /$ mabi | MAX-C | DEP-C | NoCODA | ONSET |
| ---: | :---: | :---: | :---: | :---: |
| $\rightarrow \mathrm{k} \mathrm{\varepsilon l} \mathrm{mabi}$ |  |  | $*$ |  |
| ko mabi | $*!$ |  |  |  |

Moral: even though French has no repair mechanism for hiatus or codas, it nevertheless avoids them when it can do so at no cost to faithfulness.

## 3. Case study: Tagalog nasal substitution for the millionth time

Nasal substitution

| a. | $h$ | hukbó | 'army' | pay-hukbó |
| :--- | :--- | :--- | :--- | :--- |
|  | $m$ | marká | 'mark' | 'military' |
|  | (no examples of $n$ ) |  | pay-marká | 'marker' |


| kaliskîs | 'scales' | pay-kaliskîs | 'tool for removg scales' |
| :---: | :---: | :---: | :---: |
| $?$ ?isda? | 'fish' | mai-ni-yisdá? | 'fisher' |
| ?ulól | 'silly' | may-Rulól | 'to fool someone' |
| $b$ mag-bigáj | 'to give' | ma-migáj | 'to distribute' |
| bigkás | 'pronouncing' | mam-bi-bigkás | 'reciter' |
| $d$ dalá:yin | 'prayer' | 2i-pa-naláy-in | 'to pray' |
| dinîg | 'audible' | pan-dinîg | 'sense of hearing' |
| $g$ gindáj | 'unsteadiness on feet' | pa-ŋi-ŋindáj | 'unsteadiness on feet' |
| gá:waj | 'witchcraft' | may-ga-gáwaj | 'witch' |

## 4. Phonos

Tagalog freely tolerates $m p, n t, n s, \eta k$
*NÇ >> \{IDENT-IO[VOICE], MAX-C, DEp-V, Uniformity, Ident-IO[NASAL], etc. $\}$
Tagalog also tolerates onset $\eta$ (cf. Japanese, English).
*[ $\eta \gg$ \{Ident-IO[PLACE], MAX-C, IdENT-IO[NASAL], etc. $\}$

## 5. Distribution of nasal substitution



- Nasal substitution is frequent when it eliminates a violation of *NC
- Nasal substitution is infrequent when it creates a violation of *[ $\eta$ (actually, I propose a scale $*[\eta \gg *[n \gg *[m$; the phonetic property that presumably makes [ y$]$ a bad onset is that the oral "side tube" is very short, so there are no low-enough antiformants to interfere with the vowel-like-ness of the sound; [n] has a slightly longer oral tube, but the lowest antiformant is still higher than [m]'s)


## 6. Loans from Spanish

This distribution is fairly productive-if you can say that of a distribution:


## 7. Cross-linguistic facts

(adapted from Newman 1984, p. 10)

(Caution: Tagalog is often described as being a Sama-Badjao- or Cebuano-type language, though the facts are more complicated; the same may be true for some other languages. You will see another interesting pattern in Timugon Murut in a future problem set.)

## 8. Case study: Finnish genitives

Anttila, Arto. 1997. Deriving variation from grammar. In Frans Hinskens, Roeland van Hout, and Leo Wetzels (eds.), Variation, Change and Phonological Theory, 35-68.
Amsterdam/Philadelphia, John Benjamins Publishing Company.

## 9. Distribution of Finnish genitives

Some stems always take the 'strong' genitive -iden, some always take the 'weak' genitive $-e n /-$ $j e n$, and some vary, but often with a preference one way or the other.

- Monosyllables always take the strong variant (InitialStress, *StressedLight)
/maa/ má-i.den *má.-jen
- Disyllabic stems ending in a light syllable always take the weak variant (InitialStress, *StressClash, *UnstressedHeavy)
/kala/ *ká.lo-i.den, *ká.lò-i.den ká.lo.-jen
- Disyllabic and longer stems ending in a heavy syllable always take the strong variant (IdentWeight?? These cases aren't really discussed.) /palttoo/ pált.to-i.den *pált.to.-jen
- Trisyllabic and longer stems ending in a light syllable vary.
- Those ending in a high vowel prefer the weak variant
/lemmikki/ ~lém.mik.kè-i.den lém.mik.ki.-en
- Those ending in a low vowel prefer the strong variant /sairaala/ sái.raa.lò-i.den $\sim$ sái.raa.lo.-jen
- Those ending in a mid vowel vary more freely (secondary stress is optional: *LAPSE must be freely ranked w.r.t. some anti-stress constraint)
/fyysikko/ fýy.si.kò-i.den fýy.sik.ko.-jen

Weird quirk: these generalizations refer to underlying vowel height

```
*HeavyHigh > > *HeavyMid >> *HEavyLow
*LIGHTLOW >> *LIGHTMID >> *LIGHTHIGH
*StRESSEDHIGH >> *STRESSEDMid >> *STRESSEDYLOW
*UNSTRESSEDLOW >> *UNSTRESSEDMID >> *UNSTRESSEDHIGH
(do we really need all four scales?)
```

- In trisyllabic and longer stems, there's also a tendency for a heavy antepenult to take the weak genitive and for a light antepenult to take the strong genitive. ("weight-clash/lapse" constraints: *H.H, *L.L)

Weight-clash considerations conflict with vowel-height considerations (corpus data reported by Anttila, for 3-, 4- and 5-syllable words combined):


Some categorical gaps in longer words:

| /ministeri/ | mí.nis.te.rè-i.den | mí.nis.te.ri.-en |
| :--- | :--- | :--- |
| /margariini/ | *már.ga.rì.ne-i.den | már.ga.rìi.ni.-en |
|  | *már.ga.rii.nè-i.den |  |
| /aleksanteri/ | á.lek.sàn.te.rì-i.den | á.lek.den |

## 10. Proposed grammar

Nearly stratal, but not quite

(plus transitivity)

There is variation in ranking within each row, and among all the constraints of the last three rows, except that *LIGHTO>>*LIGHTI and *UnSTRESSEDA>>*UnSTRESSEDO>>*UNSTRESSEDI.

- How can this work in terms of statistics?

