Today's topics:

- **Richness of the Base; inputs, URs**
- **Allophonic alternations in OT**

Background preparation:

- McCarthy (2008), sec 2.8

0. Today's objectives

After today's class, you should be able to:

- Explain, at a conceptual level, how to model predictable and unpredictable information in OT
- Define the term "Richness of the Base" (ROTB), and explain its implications for
  - modeling predictable information
  - inputs in our phonological model
- Propose and rank constraints for allophonic alternations in OT, considering ROTB

1. Predictable vs. unpredictable information

- Review: In generative grammar, there is a key conceptual distinction between **predictable** and **unpredictable** information
  - Unpredictable information ***must*** be
  - Predictable information (if productive) is

1. Predictable vs. unpredictable information

- Review: In generative grammar, there is a key conceptual distinction between **predictable** and **unpredictable** information
  - Unpredictable information ***must*** be learned, memorized, **stored in the lexicon**
  - Predictable information (if productive) is **enforced by the mental grammar**

- How does this look from the perspective of a constraint-based phonological model (OT)?

1. Predictable vs. unpredictable information

From a constraint-based perspective

- **Predictable** info: enforced by the grammar
  - Now this means that predictable information is enforced by the **constraints** as they are **ranked** in a particular language

- **Unpredictable** info: stored in mental lexicon
  - But now this also means that unpredictable information found in the UR/input form must **survive** in the **winning output** form

1. Predictable vs. unpredictable information

- What does it mean to ask if it is **predictable whether or not** syllables have **codas** in some language?
  - Some examples to explore this question

2. Modeling a language without codas

- Language #1: Syllables never have codas, codas avoided by deletion | What is the ranking?

/tip/	NoCoda	Dep	Max
→ a. ti			*
b. tip	*		
c. ti.pV		*	

(We know /tip/ because of an **alternation**: /tip+o/ → [tipo])

2. Modeling a language without codas

- Language #1: Syllables never have codas, codas avoided by deletion | What is the ranking?

/tip/	NoCoda	Dep	Max
→ a. ti			*
b. tip	*W		L
c. ti.pV		*W	L

(We know /tip/ because of an **alternation**: /tip+o/ → [tipo])

2. Modeling a language without codas

- Language #1: Syllables never have codas, codas avoided by deletion | **{ NoCoda, Dep } » Max**

/tip/	NoCoda	Dep	Max
→ a. ti			*
b. tip	*W		L
c. ti.pV		*W	L

(We know /tip/ because of an **alternation**: /tip+o/ → [tipo])

2. Modeling a language without codas

- Language #1: Syllables never have codas, codas avoided by deletion | What is the ranking?

/ba/	NoCoda	Dep	Max
→ a. ba			
b. bat	*	*	

2. Modeling a language without codas

- Language #1: Syllables never have codas, codas avoided by deletion | What is the ranking?

/ba/	NoCoda	Dep	Max
→ a. ba			
b. bat	*w	*w	

2. Modeling a language without codas

- Language #1: Syllables never have codas, codas avoided by deletion | (no rk arg here)

/ba/	NoCoda	Dep	Max
→ a. ba			
b. bat	*W	*W	

  - Ranking from above is *consistent* with this outcome: **{ NoCoda, Dep } » Max**

2. Modeling a language without codas

- Language #1: Syllables never have codas / deletion
  - Is it **predictable** or **unpredictable** whether a word chosen at random will have a coda or not?
  - What is the relationship between *markedness* and *faithfulness* constraints here?

2. Modeling a language without codas

- Language #1: Syllables never have codas / deletion
  - Is it **predictable** or **unpredictable** whether a word chosen at random will have a coda or not?
    - **Predictable**! Whether UR ends in a consonant or not, output will have **no coda**

      /tip/ → [ti] *and* /ba/ → [ba]
  - Relationship between *markedness*, *faithfulness*?
    - The **markedness** constraint enforcing the predictable pattern (NoCoda) dominates at least one faithfulness constraint that would preserve the contrast

3. Modeling a language that allows codas

- Language #2: Codas allowed | What is the ranking?

/kip/	DEP	MAX	NOCODA
→ a. kip			*
b. ki		*W	L
c. ki.pV	*W		L

/ma/	DEP	MAX	NOCODA
→ a. ma			
b. mat	*W		*W

3. Modeling a language that allows codas

- Language #2: Codas allowed | **{ DEP, MAX } » NOCODA**

/kip/	DEP	MAX	NOCODA
→ a. kip			*
b. ki		*W	L
c. ki.pV	*W		L

/ma/	DEP	MAX	NOCODA
→ a. ma			
b. mat	*W		*W

3. Modeling a language that allows codas

- Language #2: Codas are allowed
  - Is it **predictable** or **unpredictable** whether a word chosen at random will have a coda or not?
  - What is the relationship between *markedness* and *faithfulness* constraints here?

3. Modeling a language that allows codas

- Language #2: Codas are allowed
  - Is it **predictable** or **unpredictable** whether a word chosen at random will have a coda or not?
    - **Unpredictable**! If UR ends in a consonant, the output will have a coda; otherwise not

      /tip/ → [tip] *and* /ba/ → [ba]
  - What is the relationship between *markedness* and *faithfulness* constraints here?
    - All **faithfulness** constraints that preserve the contrast dominate the markedness constraint enforcing the pattern

4. Modeling predictable/unpredictable info

- Summary: Whether some phonological property is **predictable or unpredictable** depends on the **markedness vs. faithfulness** rankings
  - For **unpredictable** information to survive in the output form, *all* the relevant **faithfulness** constraints must dominate the markedness constraint that would remove that unpredictable information

4. Modeling predictable/unpredictable info

- Summary: Whether some phonological property is **predictable or unpredictable** depends on the **markedness vs. faithfulness** rankings
  - If the **markedness** constraint dominates even *one* faithfulness constraint, the winner will be unfaithful and the markedness constraint will always be satisfied
    - Every surface form will *avoid* the same phonological pattern (so the behavior is **predictable**)

4. Modeling predictable/unpredictable info

- Note that for input /ba/, candidate (a) [ba] has *no violations* of { MAX, DEP, NOCODA }
  - This means that /ba/ → [ba] is predicted to be the winner in *all languages*
    - How does this match what we know about cross-linguistic syllable typology?
  - /ba/ → [ba], *[bat] is an example of **harmonic bounding**
    - /ba/ → [ba] has a *proper subset* of the violations of /ba/ → [bat]

5. The OT principle of “Richness of the Base”

- Now consider Language #3 — see also A CVCV language
  - Similar to Language #1 in that no surface forms have codas (predictable)
  - But unlike Language #1, *no* alternations
    - Every morpheme always surfaces with no evidence for a final C: /pa/ always surfaces as [pa], /mifu/ always surfaces as [mi.fu], etc.
    - This means that every morpheme’s UR has the same segmental structure as its SR
  - What is the constraint ranking for this language?

5. The OT principle of “Richness of the Base”

- Language #3: No codas, no alternations
  - What ranking can we explicitly motivate by analyzing attested surface forms?

/mifu/	NoCoda	Dep	Max
→ a. mi.fu			
b. mif	*		*
c. mi.fut	*	*	

5. The OT principle of “Richness of the Base”

- Language #3: No codas, no alternations
  - What ranking can we explicitly motivate by analyzing attested surface forms?

/mifu/	NoCoda	Dep	Max
→ a. mi.fu			
b. mif	*W		*W
c. mi.fut	*W	*W	

5. The OT principle of “Richness of the Base”

- Language #3: No codas, no alternations
  - What ranking can we explicitly motivate by analyzing attested surface forms?
  - **Nothing!**

    We can’t motivate any rankings here, because there are **no conflicts** among these constraints

5. The OT principle of “Richness of the Base”

- However!

  If we are serious about the idea that **predictable** patterns are driven by **markedness** constraints, we must conclude that NoCoda » *Faithfulness*

  - NoCoda must dominate either Max or Dep, although **we don’t know which one**
  - Why don’t we know? Because real words in this language never show either epenthesis or deletion

5. The OT principle of “Richness of the Base”

- If we have NoCoda » *Faithfulness*, we have a grammar with the **power** to get rid of codas
  - Even if we give the grammar an input with a final consonant, the output will still have no coda

5. The OT principle of “Richness of the Base”

- But…how can we give the grammar an input with a final consonant, if there is no evidence that any morpheme ends in a consonant?
  - Here is where ***input and UR are not the same***
  - We can give the grammar a **hypothetical** input (not a real word) and consider *what it would do*

5. The OT principle of "Richness of the Base"

- We can give the grammar a **hypothetical** input (not a real word) and consider *what it would do*

/CVC/	NoCoda	Dep	Max
?→ a. CV_			*
?→ b. CV.CV		*	
c. CVC	*W	L(?)	L(?)

  - If Dep » Max (Max is lowest), candidate (a) will win
    If Max » Dep (Dep is lowest), candidate (b) will win

5. The OT principle of “Richness of the Base”

- We can give the grammar a **hypothetical** input (not a real word) and consider *what it would do*
  - If DEP » MAX (MAX is lowest), candidate (a) will win
    If MAX » DEP (DEP is lowest), candidate (b) will win
  - We don’t know which, but **one of them will**

    This is because **NoCoda must dominate** (at least) **one faithfulness constraint**, in order for the grammar to **enforce** the **predictable** information that codas are not allowed

5. The OT principle of "Richness of the Base"

- What this means: A grammar with NoCoda » *Faithfulness*$_1$ will **productively** get rid of codas, even if no existing morphemes show this alternation

- This example illustrates a key OT principle:
  - **Richness of the Base** (ROTB): There are no language-particular restrictions on input forms (Prince & Smolensky 1993)

5. The OT principle of “Richness of the Base”

- **Richness of the Base** (ROTB): There are no language-particular restrictions on input forms
  - Translation: If something is a possible input in one language (such as /CVC/), it is a **possible input in all languages**
  - This means we do not need devices such as “morpheme structure constraints” that allow the phonological grammar to specify what is or is not a *possible UR* in each language
- But see also: Gouskova, Maria. 2024. MSCs in positional neutralization: The problem of gapped inventories. *Phonology* 40: 229-265.

5. The OT principle of “Richness of the Base”

- In **rule-based phonology**, how would we model Language #3 (no codas; no alternations)?
  - We would **not** model this with a deletion (or insertion) **rule**, because there is no deletion (insertion) process in this language
    - Morphemes never have consonants in a position where they would *become* codas, so there is no need to “remove” codas
  - Instead, we would propose a **morpheme structure constraint**:
    “URs can never have CC or C#”

5. The OT principle of “Richness of the Base”

- In OT, Language #3 is modeled the **same way** as Language #1: NoCoda » *Faithfulness*

Language #1	*Language #3*
/tip/ [ti_], /tip+o/ [ti.**p**o]	/pa/ [pa], /mifu/ [mi.fu]
Consonant deletion can be observed	There are no visible C~Ø or Ø~V alternations
Analysis: NoCoda » Max (and Dep » Max)	Analysis: NoCoda » (Max or Dep)

- Predictable = **enforced by the constraint ranking**, whether we see an active “rule” (alternation) or not

6. ROTB and *productive* predictable patterns

- In general:
  If some structure is absent in a language, this tells us markedness (M) » *at least one* faithfulness (F)
  - codas in Hawai'ian
  - front rounded vowels in English

6. ROTB and *productive* predictable patterns

- Having an M » F ranking for some phonological structure makes a prediction:
  - It should be part of the native speaker's **knowledge** that this structure is illegal
  - Evidence for this?
    Consider loanword phonology or invented words; what does a native speaker do?
    - Hawai'ian: English *wine* [wai<u>n</u>] → [wai.n<u>a</u>]
    - English: French *menu* [...n<u>y</u>] → [...n<u>ju</u>]

6. ROTB and *productive* predictable patterns

- However, languages also have **accidental gaps**
  - Some structure just happens to be missing in the morphemes of the lexicon, but the grammar doesn't actually prohibit it
    - The absence of this structure is **not productive**
  - Example: [bw] onset clusters are extremely rare in English, but in experiments, native speakers do not treat them as ungrammatical

6. ROTB and *productive* predictable patterns

- We can model this difference in OT as follows:
  - True **productive gaps** have **M » F**
    - Given a “new” word or a loanword, native speakers actively avoid the structure
  - **Accidental gaps** do **not** have ~~M » F~~
    - Given a “new” word or a loanword, native speakers produce the structure faithfully
    - *Warning:* Sometimes speakers/languages use special faithfulness for loanwords — this could mean that the gap *is* productive, outside loanwords

7. ROTB and implications for “inputs”

- What is an **input** in an OT grammar?
  - Sometimes, an input is an **actual UR** of an attested morpheme or word
  - But sometimes, an input is a **hypothetical** input that we use to make the grammar is doing its job: the grammar must actively enforce productive predictable patterns

7. ROTB and implications for “inputs”

- McCarthy (2008), sec 2.8 | What inputs do we need to consider for an analysis?
  - Obvious: **URs that surface unfaithfully** — show that the grammar is enforcing something
  - Perhaps less obvious: **URs that surface faithfully** are informative when they violate markedness constraints

7. ROTB and implications for “inputs”

- McCarthy (2008), sec 2.8 | What inputs do we need to consider for an analysis?
  - Perhaps even less obvious: **Hypothetical inputs**, that don’t correspond to actual URs of the language, are sometimes needed in order to test the grammar’s ability to rule out ungrammatical forms

8. Application: Complementary distribution

- What kind of **descriptive generalization** can we make about the Spanish data set in McCarthy (2008), sec 2.8, exercise (30)?

8. Application: Complementary distribution

- **Descriptive generalization** (scratch work):
  - Voiced fricatives appear …
  - Voiced stops appear …

- Can we express either environment as a **featurally natural class**?
  - Is there a clear “elsewhere” case?
  - Will thinking about syllable structure help here?

8. Application: Complementary distribution

- **Descriptive generalization** (scratch work):
  - Voiced fricatives appear after vowels, liquids [l, r]
  - Voiced stops appear in initial position and after oral stops, nasals, and fricatives

- Can we express either environment as a **featurally statable class**?
  - Is there a clear "elsewhere" case?
  - Will thinking about syllable structure help here?

8. Application: Complementary distribution

- **Descriptive generalization**
  - Voiced fricatives appear after [+son, –nas]
  - Voiced stops appear elsewhere

- Can we propose constraints analogous to those for the Yoruba and Madurese nasalization patterns that McCarthy discusses (p 91)?
  - What ranking do we need for Spanish?
  - Later: What are the cross-linguistic (typological) consequences of proposing these constraints?

8. Application: Complementary distribution

- **F** constraint related to the **difference** between allophones

- **M** constraint forcing the **context-specific** allophone

- **M** constraint forcing the **elsewhere** allophone

8. Application: Complementary distribution

- **F** constraint related to the **difference** between allophones

  IDENT[±cont]

- M constraint forcing the **context-specific** allophone

  *VD ‘assign one * for every **sequence** of segments [+son, –nas] [–son, –cont, +voi]’

- M constraint forcing the **elsewhere** allophone

  *VOIFRIC ‘assign one * for every **segment** that is [–son, +cont, +voi]’

8. Application: Complementary distribution

- How are these constraints ranked for Spanish? (What rankings among them can we **prove**?)

	IDENT[±CONT]	*VOIFRIC	*VD

8. Application: Complementary distribution

- How are these constraints ranked for Spanish? (What rankings among them can we **prove**?)

/aba/	*VD	*VoiFric	Ident[±cont]
(a) [aba]	* W	L	L
→ (b) **[aβa]**		*	*

- What other kind of **input** should we try?

8. Application: Complementary distribution

- How are these constraints ranked for Spanish?
  M-context-specific » M-general » F

/aba/	*VD	*VoiFric	Ident[±cont]
(a) [aba]	* W	L	L
→ (b) **[aβa]**		*	*

/βa/	*VD	*VoiFric	Ident[±cont]
(a) [βa]		* W	L
→ (b) **[ba]**			*

- Note the key role of **ROTB** here

8. Application: Complementary distribution

- So… what is the UR for [b]~[β] in Spanish?
  - Could it be [b]?
  - Could it be [β]?
  - Could it be abstract/underspecified?
  - What information **must** this UR contain?

- Next time: The typological implications of our analysis of Spanish complementary distribution