

Predictable information in OT and 'Richness of the Base'

1. Predictable information and constraint ranking

Suppose a language only has morphemes with the shape /CV/, /CVCV/, /CVCVCV/, etc.

- Can we prove a ranking for **ONSET** and **NoCODA** with respect to **NoEPENTHESIS** and **NoDELETION** using any combination of morphemes from this language?
 - No! We cannot prove any ranking among these constraints with any morphemes in the lexicon. (Be sure you see why this is.)
- What do you think will happen if this language borrows a word with the shape /CVC/? How do you think the loanword will surface?
 - Most languages that never have codas also avoid having codas when they borrow words (at least at first — prolonged borrowing can change this pattern).
 - If a language adds vowels to avoid codas when borrowing words, what does this suggest about the relative ranking among **NoCODA**, **NoEPENTHESIS**, and **NoDELETION**?

2. 'Richness of the Base'

- A key idea of OT is: any **predictable** pattern that is part of the **productive** knowledge of a native speaker is something that the **grammar** (constraint ranking) **must enforce**.
 - If a language has no codas ever, that is not a mere coincidence based on what kinds of morphemes happen to be in the lexicon.
 - Instead, this reflects a ranking where **NoCODA** » **NoEPENTHESIS** *or* **NoCODA** » **NoDELETION** (although we may not know *which* of these rankings the language actually has!) — even though we can't directly prove this ranking using **existing morphemes** in the language.
- Consequence: If some structure (such as a coda) is completely illegal in a language, then even if no actual UR ever contains this structure, we still have to make sure the **grammar** (constraint ranking) is **robust** enough to get rid of the illegal structure.
 - This means we may need to argue for a constraint ranking based partly on a **hypothetical input** whose job is to represent a completely illegal structure.
 - For the example above: we need to consider a /CVC/ input and determine what ranking or rankings must be part of the grammar to prevent a [CVC] output from winning.
 - This principle has been called '**Richness of the Base**' (where 'rich' means 'unrestricted', and 'base' is an inconvenient older technical term) — this principle says that there can be **no language-particular restrictions on input forms**.

- Important: Richness of the Base *doesn't* mean that speakers have ridiculous URs in their actual lexicon (like /bbbbbbapppppp/ for [bap]). It *does* mean that we may have to consider (slightly) 'ridiculous' inputs (illegal structures) in our tableaux in order to rank as many constraints as possible.
 - This is why *input* is not always exactly the same thing as **UR**; some **hypothetical inputs** are **not actual URs** of the language, but are considered because they represent illegal patterns that the grammar must have a way to eliminate.