Optimal Syntax: Overview

01 Introduction, motivations, administrative details
02 Some Generative foundations: Procrastinate/Last Resort
03 Constraint hierarchies and grammaticality; 1st paper due
04 Markedness and faithfulness
05 Constraint interaction; 2nd paper due
06 Towards an Optimal typology – Crosslinguistic constraints and English
07 Prosody and Optimality; 3rd paper due
08 Language change – the view from Optimality
09 Correspondence, theory and application; 4th paper due
10 Syntactic effects of movement
11 Syntactic effects of control; 5th paper due
12 Radical ungrammaticality
13 Evolutionary linguistics – Optimal cues; 6th paper due
14 Wrapup

Grammar as an Input/Output system

Generative theory:
A grammar shall encompass all expressions of a language shall separate the expressions that belong to the language from those that do not

→ 1) generative grammar oversgenerates
2) filter eliminates

“ideal” separator: a simple generative mechanism called GEN

GEN allows grammatical and ungrammatical expressions
EVAL filters out ungrammatical ones

Universal Grammar (UG)
a) if a linguistic fact of an individual speaker is a fact in all known languages and
b) if this fact is not acquired as imitation of input data:
→ evidence that this fact comes from a specific feature of UG
UG allows for language variation, it limits the range of variation
→ UG defines parameters of possible variation

L1 Acquisition:
a) children draw from data with positive evidence rather than negative evidence (Guasti, 2004)
b) children are not informed about ill-formed/disallowed sentences
c) all mature speakers can judge whether a sentence is grammatical
d) some expected errors from generalization of input never occur

Cliticization constraint

a. Who do you want to see? input
b. Who do you wanna see? input

Cliticization constraint

c. Who do you want to come? infant output
d. *Who do you wanna come? impossible output

rules:
- describe mapping phonological representation to phonetic form
- to replace rules with a hierarchy of ranked, violable constraints

goal of rules/constraints: description of linguistic competence

English plural formation

basic shape /-z/ → cats /-s/ dogs /-z/ horses /-z/
constraint: sequences of obstruents must have the same value for voicing
remedy: devoicing, epenthesis

<table>
<thead>
<tr>
<th>phonological representation</th>
<th>[kæts]</th>
<th>[dɔgzs]</th>
<th>[hɔrzs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>devoicing</td>
<td>[kætv]</td>
<td></td>
<td>[hɔtv]</td>
</tr>
<tr>
<td>epenthesis</td>
<td></td>
<td>[tʃ]</td>
<td></td>
</tr>
</tbody>
</table>

Syllable structure features across languages

Hawaiian: only 1 consonant in sequence: wāhine (“woman”) borrowing: adapted ničer → ničwāhine

English: multiple consonants construct, gnumt, sixths

Berber: allows consonant-only words tīlt (“lock”)
non-occurring languages:
e.g. all consonants clustered at left/right every word begins with a consonant, etc.

→ these patterns do never occur for some reason

Patterns in natural language underlie constraints on all levels of linguistic study
**English vs. Yawelmani**

<table>
<thead>
<tr>
<th>Yawelmani /xathn/</th>
<th>English (“to eat”)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) words begin with exactly 1 C: apple is disallowed</td>
<td></td>
</tr>
<tr>
<td>2) at edges, only 1 C is allowed: strength is disallowed</td>
<td></td>
</tr>
<tr>
<td>3) no more than 1 V in sequence: alien is disallowed</td>
<td></td>
</tr>
<tr>
<td>4) word-internal CC possible but not necessary: instruct is disallowed</td>
<td></td>
</tr>
</tbody>
</table>

**Syllable tendencies (often violated)**
- begin with a C: ONSET
- have one V: PEAK
- end in a V: NoCODA
- have at most one C at the edges: *COMPLEX
- are composed of C and V: ONSET & PEAK

---

**Optimality theory basics**

- **a)** UG contains set of violable constraints
- **b)** Constraints characterize universals
- **c)** Constraint violations characterize markedness and variation

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>*COMPLEX</td>
<td>1</td>
</tr>
<tr>
<td>ONSET</td>
<td>2</td>
</tr>
<tr>
<td>PEAK</td>
<td>3</td>
</tr>
<tr>
<td>NoCODA</td>
<td>4</td>
</tr>
</tbody>
</table>

Optimal candidate: fewest lowest violations; only violation: of the lowest ranking constraint → each language has its own ranking of these constraints differences in ranking: systematic variation between languages

---

**Some Generative foundations**

Crosslinguistically, valid regularities of PF to be found in outputs

**Phonological representation:** abstract, facilitates word recognition

Cf. word/non-word pairs like *murse/nurse*

Basic assumption in Generative theory

**Phonemic principle:** differences in phonological form correspond with differences in the message result a) rules abstract away from surface differences result b) rules attribute as much as possible the same phonological shape (word, morpheme) in all environments therefore: rules of grammar

- state regularities of sound structure
- relate abstract phonological forms to their phonetic realizations

**Generative Phonology**

rules of the form $A \rightarrow B/C\_D$

- input sequences of the form CAD & assigns property B to A

Cf. impossible, immobile, imbalanced, involuntary, infallible, insecure

input: in- output: in-/im- + Adj

Proposed OT solutions:

- flaws in model that rules exhaustively characterize competence
cf. irregulars: learn - left, hear - hot, mean - meant

- not -ed suffix is devoiced but stem is devoiced

- insufficiency of rules

---


Generative rule system implies over specification

- theories needed
  - a) to define class of possible predicates CAD (structural descriptions)
  - b) to define class of possible operations (structural changes)

Cf. extension of epenthesis principle

pick: picked, rain - rained, wall - waited

→ linguistic competence represented by set of constraints

Motivation for difference in underlying and surface form

- defined from the result: to avoid ill-formed clusters

Cf. Spanish: 0 → [e]/-/C

English: aspiration, /p/ → rules serve to satisfy constraints on well-formedness of surface

→ locus of explanation is in a theory of constraints

---

**Constraints**

are effects of universal aspects of sound structure

- members of a set of constraints: typically in conflict
  - satisfying one constraint may violate another constraint

2 fundamentally conflicting demands in sound structure:

**Markedness:** tendency for phonetic forms to be pronounced in a simple, natural way (determined by speech organs, cognitive endowment etc.)

**Faithfulness:** tendency for distinguishing lexical properties to be preserved in the phonetic realizations
GEN and EVAL 1
How to compute a surface form corresponding to a given input? = comparison of all possible surface forms (a set of candidates) that could correspond to the input

OT proposes input, output and a relationship between the two → relationship mediated by GEN (generator) and EVAL (evaluator)

a) selection of the candidate that conforms best to constraints GEN produces set of candidates
b) EVAL identifies candidate with highest degree of harmony (i.e. that violates high-ranked constraints to smallest possible degree) → constraints can be violated in order to satisfy a higher ranked constraints

Functions of GEN and CON
Functions of GEN
Primary task: Generation processes of addition, deletion, rearrangement without restrictions
→ candidate set to be created: potentially infinite
Secondary task: to indicate input - output

Functions of CON
Universal set of constraints, part of innate knowledge
→ every language uses same set of constraints
Constraints: formal means that encode universals
Potential for a constraint to be violated: result of the position of this constraint in the particular language’s hierarchy
→ constraints are a measure for Markedness
→ each constraint is a Markedness statement

Ranking hierarchies
FaithC violation in Spanish
- features in Spanish: breaking up sC clusters
  0 → [e]%sC → esfera ("sphere"), hemisferio ("hemisphere"), not *herniesferio
- Spanish inf. marker: -er/-ir verb roots ending with cluster of two C, + inf. marker
  (1) absorb-er esculp-ir disting-uir

Adjectivization/Nominalization:
Adj suffix -to; N suffix -tor:
only single C:
(2) abso-to escul-tor distin-to

Independence phonological shape from morph. function

Functions of EVAL
- selects optimal candidate from set created by Gen
→ finds candidate that best satisfies ranked constraints

How to achieve best satisfaction:
- violations of lower ranked constraints are violated given that they help avoid violation of higher ranked constraints

Process account of Spanish FaithC violation
a) infinitive suffix is vowel-initial
→ C is syllabified as onset of that syllable
b) Adj/N suffixes are consonant-initial
→ final C of verb root cannot syllabify via CC onset or CC coda, thus violating *Complex
(3) ab.sor.to ab.sor.to ab.sor.to *ab.sor.to *ab.sor.to es.cul. to es.cul.to *es.cul.to *es.cul.to

→ Spanish does not allow *Complex violations
(cf. English, does not allow a C to syllabify in sicken)
→ Spanish does not allow epenthesis here
(cf. epenthesis in Japanese or Hawaiian)
Spanish FAITHC violation (continued)

Epenthetic vowel: could “rescue” a single C
Spanish option: to delete it
c) "lose a consonant" option
→ FAITHC is subordinate → in a CCC sequence: 1 left out thus
PEAK, FAITHV, *COMPLEX satisfied, only FAITHC violated

<table>
<thead>
<tr>
<th>/absorb/</th>
<th>FAITHV</th>
<th>PEAK</th>
<th>*COMPLEX</th>
<th>FAITHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>abs. orb. to</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
</tr>
<tr>
<td>abs. orb. to</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
</tr>
<tr>
<td>abs. orb. be to</td>
<td></td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
</tr>
</tbody>
</table>

English *COMPLEX violation

Nominalization of Adj.: -ness
added to Adj. to create N (happiness, sadness…)

Process account
a) –ness suffix is consonant-initial
→ critical environment: Adj. that are CC-final,
cf. damp, soft → damness, dampness
b) English creates complex Coda with damness
thus
PEAK, FAITHV, FAITHC satisfied, only *COMPLEX violated

<table>
<thead>
<tr>
<th>/damp ness/</th>
<th>FAITHV</th>
<th>PEAK</th>
<th>FAITHC</th>
<th>*COMPLEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>damp ness</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
</tr>
<tr>
<td>damp ness</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
</tr>
<tr>
<td>damp ness</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
<td>⌊</td>
</tr>
</tbody>
</table>

Constituents: Formal description

class membership defined by morphosyntactic properties

morphosyntactic properties:
a) inflection and derivation
b) position/distribution

Constituents:
all operations in grammar are constituent-based
all constituents belong to a limited set of grammatical categories

categories store grammatical properties
properties: usually morphosyntactic (i.e. N take plural, the,…)

Inflectional properties: Content words

English: inflectional properties only distinctive for N and V

Noun: number marking -s
problems: invariable plurals -sheep, irregulars (mice)
count/mass N (*golds), compounds coat hangers

Verb: 4 inflectional markings:
(1) -n (perfective/ past participle) (2) -s (3rd person)
(3) -ing (imperfective/ progressive participle) (4) -d (past tense)

problems: irregular past or perfectives
Content words:
- typically have antonyms
N, V, P, A, ADV

Inflectional marking: Categories

Inflection marking in English per categories

problems:
- inflection marking can be regular and irregular
- morphological evidence can be misleading
(A –comparison, un- l –ness etc.)
Inflectional marking: Types

<table>
<thead>
<tr>
<th>marking</th>
<th>singular</th>
<th>plural</th>
<th>case</th>
<th>tense</th>
<th>aspect</th>
<th>modality</th>
<th>person</th>
<th>degree</th>
<th>gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>num</td>
<td>nom</td>
<td>gen</td>
<td>dat</td>
<td>acc</td>
<td>voc</td>
<td>masc</td>
<td>neut</td>
<td>masc</td>
</tr>
</tbody>
</table>

Inflectional marking in English per types

→ specific inflection attaches to specific categories

INFL & COMP
- I or INFL: finite auxiliaries and infinitival to
- Complementizer, COMP or C

(1) John wonders [whether it will rain tonight]
[] = complement clause

whether: subordinating conjunction, Complementizer
C can be finite or non-finite:
whether, if, that… finite
for: infinite

(2) [For you to make the right decision] is unthinkable

Formal categories of language, constituents

Classify constituents using square brackets:

(3) All conclusive results show clearly that we should
manage to finish this daunting task in a long lunch break

Descriptions beyond constituent level 1

Syntactic evidence:
different categories have different distributions

(4) A: Where is the bookshop?
B: Down the street

properties of b:
*down the street was built yesterday
*down the street meal for Mary tasted bad

From here, the only way goes down/down the street

→ has P-like properties,
down determines the street
[PP [P down [D the [N street]]]]

Descriptions beyond constituent level 2

(5) A: What are you trying to do?
B: Help you

down is head of phrase down the street
down the street is a projection of down

Projection: expansion of a head word
Head: preposition, projects into PP

Constituent: a structural unit or component out of which
a phrase or clause is built up

Descriptions beyond constituent level 3

Phrases: expressions larger than a word which are
maximal projections
IP, CP, DP, NP, VP, AP, ADVP, PP

Clauses: expressions which contain a subject and a
predicate.
most cases: predicate has lexical (= nonauxiliary) verb
number of lexical verbs = number of clauses

Minimal projection: not a projection of another constituent
Heads = minimal projections

Maximal projection: constituent not contained in a larger
constituent with the same head