T-orders across categorical and probabilistic constraint-based phonology

Arto Anttila (Stanford University) and Giorgio Magri (CNRS)

AMP 2017, New York University, September 15-17
Introduction
□ Implicational universals:

\[ P \rightarrow \hat{P} \]

if a language has property \( P \), it also has property \( \hat{P} \)
□ **Implicational universals:**  

- $P \rightarrow \hat{P}$  
  if a language has property $P$, it also has property $\hat{P}$

□ **Examples:**

- $CCCV \rightarrow CCV \rightarrow CV$  
  if a language allows a complex margin cluster, it allows a simpler one

- $(/\text{cost.us}/, [\text{cos.us}]) \rightarrow (/\text{cost.me}/, [\text{cos.me}])$  
  if a dialect of English deletes $t/d$ before $V$, it also deletes it before $C$
Consider a typology \( \mathcal{T} \) of phonological grammars, construed as mappings from underlying representations to surface representations.
Consider a typology $\mathcal{T}$ of phonological grammars, construed as mappings from underlying representations to surface representations.

Consider the simplest antecedent and consequent properties $P$ and $\hat{P}$:

- $P$: the property of mapping an ur $x$ to a sr $y$: $(x, y)$
- $\hat{P}$: the property of mapping an ur $\hat{x}$ to a sr $\hat{y}$: $(\hat{x}, \hat{y})$
Consider a typology $\mathcal{T}$ of phonological grammars, construed as mappings from underlying representations to surface representations.

Consider the simplest antecedent and consequent properties $P$ and $\widehat{P}$:

$P$: the property of mapping an ur $x$ to a sr $y$: $(x, y)$

$\widehat{P}$: the property of mapping an ur $\widehat{x}$ to a sr $\widehat{y}$: $(\widehat{x}, \widehat{y})$

The implicational universal

$$(x, y) \xrightarrow{\overline{x}} (\widehat{x}, \widehat{y})$$

holds provided each grammar in $\mathcal{T}$ which maps $x$ to $y$ also maps $\widehat{x}$ to $\widehat{y}$.
Consider a typology $\mathcal{T}$ of phonological grammars, construed as mappings from underlying representations to surface representations.

Consider the simplest antecedent and consequent properties $P$ and $\hat{P}$:

- $P$: the property of mapping an ur $x$ to a sr $y$: $(x, y)$
- $\hat{P}$: the property of mapping an ur $\hat{x}$ to a sr $\hat{y}$: $(\hat{x}, \hat{y})$

The implicational universal

$$(x, y) \xrightarrow{\mathcal{T}} (\hat{x}, \hat{y})$$

holds provided each grammar in $\mathcal{T}$ which maps $x$ to $y$ also maps $\hat{x}$ to $\hat{y}$.

$\xrightarrow{\mathcal{T}}$ is a partial order called the **T-order** induced by $\mathcal{T}$ [Anttila and Andrus 2006]
Implicational universals can be statistical: [Guy 1991; Kiparsky 1993; Coetzee 2004]

- variable $t/d$ deletion is more frequent before $C$ than $V$
Implicational universals can be statistical: [Guy 1991; Kiparsky 1993; Coetzee 2004]

- variable $t/d$ deletion is more frequent before $C$ than $V$

Consider a typology $\mathcal{T}$ of probabilistic phonological grammars, construed as functions from underlying forms to probability distributions over surface forms
Implicational universals can be statistical: [Guy 1991; Kiparsky 1993; Coetzee 2004]

- variable \( t/d \) deletion is more frequent before C than V

Consider a typology \( \mathcal{I} \) of **probabilistic** phonological grammars, construed as functions from underlying forms to probability distributions over surface forms

The implicational universal

\[
(x, y) \xrightarrow{\mathcal{I}} (\hat{x}, \hat{y})
\]

holds provided each grammar in \( \mathcal{I} \) assigns a probability to \( (\hat{x}, \hat{y}) \) which is at least as large as the probability it assigns to \( (x, y) \)
Implicational universals can be statistical: [Guy 1991; Kiparsky 1993; Coetzee 2004]

- variable $t/d$ deletion is more frequent before C than V

Consider a typology $\mathcal{T}$ of **probabilistic** phonological grammars, construed as functions from underlying forms to probability distributions over surface forms.

The implicational universal

$$(x, y) \xrightarrow{\mathcal{T}} (\hat{x}, \hat{y})$$

holds provided each grammar in $\mathcal{T}$ assigns a probability to $(\hat{x}, \hat{y})$ which is at least as large as the probability it assigns to $(x, y)$.

The preceding categorical definition of T-orders is a special case of the probabilistic one.
T-orders are interesting/important because:

- impose strict limits on categorical and statistical phonological patterns
- “measure” the amount of typological structure
- follow from the phonological theory
- need not be learned, cannot be subverted by learning
- model categorical and gradient phonotactic judgments

[Becker et al. 2011]

[Anttila 2008]
- T-orders are interesting/important because:
  - impose strict limits on categorical and statistical phonological patterns
  - “measure” the amount of typological structure
  - follow from the phonological theory
  - need not be learned, cannot be subverted by learning
  - model categorical and gradient phonotactic judgments

- We develop formal theory of T-orders in constraint-based phonology:
  - two categorical frameworks: categorical OT (\(\overset{OT}{\rightarrow}\))
  - categorical HG (\(\overset{HG}{ightarrow}\))
  - four probabilistic frameworks: stochastic OT (\(\overset{sOT}{ightarrow}\))
  - stochastic (or noisy) HG (\(\overset{sHG}{ightarrow}\))
  - partial order OT (\(\overset{poOT}{ightarrow}\))
  - Max Ent (\(\overset{ME}{ightarrow}\))

and we explore its phonological implications
Formal results 1: a complete characterization of the relationships among T-orders in these six frameworks

- T-orders indeed allow for cross-framework comparisons of typological structure, even bridging across categorical and probabilistic frameworks
Formal results 1: a complete characterization of the relationships among T-orders in these six frameworks

- T-orders indeed allow for cross-framework comparisons of typological structure, even bridging across categorical and probabilistic frameworks

Formal results 2: an almost complete characterization of T-orders in the six frameworks in terms of just constraint violation profiles of the antecedent mapping, the consequent mapping, and their losers

- help us understand what it really means that a T-order holds
- allow us to compute T-orders without computing the entire typology
Formal results 1: a complete characterization of the relationships among T-orders in these six frameworks

- T-orders indeed allow for cross-framework comparisons of typological structure, even bridging across categorical and probabilistic frameworks.

Formal results 2: an almost complete characterization of T-orders in the six frameworks in terms of just constraint violation profiles of the antecedent mapping, the consequent mapping, and their losers

- Help us understand what it really means that a T-order holds
- Allow us to compute T-orders without computing the entire typology.

Phonological implications 1: a surprising equivalence result for OT and HG T-orders under the 2-candidate condition.
Formal results 1: a complete characterization of the relationships among T-orders in these six frameworks
- T-orders indeed allow for cross-framework comparisons of typological structure, even bridging across categorical and probabilistic frameworks

Formal results 2: an almost complete characterization of T-orders in the six frameworks in terms of just constraint violation profiles of the antecedent mapping, the consequent mapping, and their losers
- help us understand what it really means that a T-order holds
- allow us to compute T-orders without computing the entire typology

Phonological implications 1: a surprising equivalence result for OT and HG T-orders under the 2-candidate condition

Phonological implications 2: first ever analytical results on MaxEnt T-orders, showing that they display some counterintuitive properties
Formal results 1
The HG typology is usually at least as large as the OT typology hence, a HG T-order entails the OT T-order and the vice versa fails in the general case.
The HG typology is usually at least as large as the OT typology hence, a HG T-order entails the OT T-order and the vice versa fails in the general case.

But see below for more on this!
\( (x, y) \xrightarrow{\text{ME}} (\hat{x}, \hat{y}) \)

\( (x, y) \xrightarrow{\text{HG}} (\hat{x}, \hat{y}) \)

\( (x, y) \xrightarrow{\text{sHG}} (\hat{x}, \hat{y}) \)

\( (x, y) \xrightarrow{\text{poOT}} (\hat{x}, \hat{y}) \)

\( (x, y) \xrightarrow{\text{OT}} (\hat{x}, \hat{y}) \)

\( (x, y) \xrightarrow{\text{sOT}} (\hat{x}, \hat{y}) \)

\( □ \) T-orders in OT, stochastic OT, and partial order OT are identical
□ T-orders in OT, stochastic OT, and partial order OT are identical

□ despite the fact that the OT typology is finite
  the partial order OT typology is much larger
  the stochastic OT typology is infinite
T-orders in OT, stochastic OT, and partial order OT are identical

despite the fact that the OT typology is finite

the partial order OT typology is much larger

the stochastic OT typology is infinite

Analogous equivalence holds for HG and stochastic HG
□ T-orders in OT, stochastic OT, and partial order OT are identical

□ despite the fact that the OT typology is finite

□ the partial order OT typology is much larger

□ the stochastic OT typology is infinite

□ Analogous equivalence holds for HG and stochastic HG

□ Thus, stochastic OT/HG are probabilistic variants of categorical OT/HG which don’t tamper with the categorical typological structure
A ME T-order entails an HG T-order
□ A ME T-order entails an HG T-order

□ The vice versa fails even with 1 constraint and 1 loser!
□ A ME T-order entails an HG T-order
□ The vice versa fails even with 1 constraint and 1 loser!
□ Thus, although ME and stochastic HG both look like “innocuous” probabilistic variants of HG, ME actually loses much of the typological structure imposed by HG
Formal results 2
Because of the equivalences above, we only need constraint conditions for a T-order in HG, OT, and MaxEnt.
Because of the equivalences above, we only need constraint conditions for a T-order in HG, OT, and MaxEnt.

Focus on the antecedent mapping \((x, y)\) of the T-order \((x, y) \rightarrow (x', y')\).
Because of the equivalences above, we only need constraint conditions for a T-order in HG, OT, and MaxEnt.

Focus on the antecedent mapping \((x, y)\) of the T-order \((x, y) \rightarrow (\hat{x}, \hat{y})\)

consider a corresponding loser mapping \((x, z)\)

compute the antecedent difference vector:

\[
\begin{bmatrix}
C_1(x, z) - C_1(x, y) \\
C_2(x, z) - C_2(x, y) \\
\vdots \\
C_n(x, z) - C_n(x, y)
\end{bmatrix}
\]

The consequent difference vectors are defined analogously, as pitting the consequent mapping \((\hat{x}, \hat{y})\) against one of its losers \((\hat{x}, \hat{z})\)
Because of the equivalences above, we only need constraint conditions for a T-order in HG, OT, and MaxEnt.

Focus on the antecedent mapping \((x, y)\) of the T-order \((x, y) \rightarrow (\hat{x}, \hat{y})\)
consider a corresponding loser mapping \((x, z)\)
compute the **antecedent difference vector**:

\[
\text{violations of the loser } (x, z) - \text{violations of the antecedent } (x, y) = \begin{bmatrix}
C_1(x, z) - C_1(x, y) \\
C_2(x, z) - C_2(x, y) \\
\vdots \\
C_n(x, z) - C_n(x, y)
\end{bmatrix}
\]
Because of the equivalences above, we only need constraint conditions for a T-order in HG, OT, and MaxEnt.

Focus on the antecedent mapping \((x, y)\) of the T-order \((x, y) \rightarrow (\hat{x}, \hat{y})\) and consider a corresponding loser mapping \((x, z)\).

Compute the antecedent difference vector:

\[
\begin{bmatrix}
C_1(x, z) - C_1(x, y) \\
C_2(x, z) - C_2(x, y) \\
\vdots \\
C_n(x, z) - C_n(x, y)
\end{bmatrix}
\]

The consequent difference vectors are defined analogously, as pitting the consequent mapping \((\hat{x}, \hat{y})\) against one of its losers \((\hat{x}, \hat{z})\).
Suppose there are only $n = 2$ constraints:
- the antecedent difference vectors can be plotted as points in the plane.
Suppose there are only $n = 2$ constraints:

- the antecedent difference vectors can be plotted as points in the plane
- The red region is their corresponding cone
Suppose there are only $n = 2$ constraints:

- the antecedent difference vectors can be plotted as points in the plane
- The red region is their corresponding cone
- The points which are larger than some point in this cone yield the gray region
Suppose there are only \( n = 2 \) constraints:
- The antecedent difference vectors can be plotted as points in the plane
- The red region is their corresponding cone
- The points which are larger than some point in this cone yield the gray region

The HG T-order \( (x, y) \xrightarrow{\text{HG}} (\hat{x}, \hat{y}) \) holds (for any \( n \)) iff each consequent difference vector lives in this gray region: it is larger than some vector in the cone generated by the antecedent difference vectors.
Suppose there are only $n = 2$ constraints:
- the antecedent difference vectors can be plotted as points in the plane
- The red region is their corresponding cone
- The points which are larger than some point in this cone yield the gray region

The HG T-order $(x, y) \xrightarrow{HG} (\hat{x}, \hat{y})$ holds (for any $n$) iff each consequent difference vector lives in this gray region: it is larger than some vector in the cone generated by the antecedent difference vectors

Follows from the *Hyperplane Separation theorem* of convex geometry through straightforward algebra

*[Boyd and Vandenbergh 2004]*
Suppose there are only $n = 2$ constraints:
- the antecedent difference vectors can be plotted as points in the plane
- The red region is their corresponding cone
- The points which are larger than some point in this cone yield the gray region

The HG T-order $(x, y) \xrightarrow{\text{HG}} (\hat{x}, \hat{y})$ holds (for any $n$) iff each consequent difference vector lives in this gray region: it is larger than some vector in the cone generated by the antecedent difference vectors.

Follows from the *Hyperplane Separation theorem* of convex geometry through straightforward algebra.

While $(x, y) \xrightarrow{\text{HG}} (\hat{x}, \hat{y})$ is expensive to check directly (universal quantification over weights), the geometric characterization above is easy to check (polyhedral feasibility problem).
Suppose there are only $n = 2$ constraints:
- the antecedent difference vectors can be plotted as points in the plane
- The red region is their corresponding cone
- The points which are larger than some point in this cone yield the gray region

The HG T-order $(x, y) \xrightarrow{\text{HG}} (\hat{x}, \hat{y})$ holds (for any $n$) iff each consequent difference vector lives in this gray region: it is larger than some vector in the cone generated by the antecedent difference vectors.

Follows from the Hyperplane Separation theorem of convex geometry through straightforward algebra.

While $(x, y) \xrightarrow{\text{HG}} (\hat{x}, \hat{y})$ is expensive to check directly (universal quantification over weights), the geometric characterization above is easy to check (polyhedral feasibility problem).

We thus advertise new Python code to establish HG T-orders.
Suppose again there are only $n = 2$ constraints:
- the antecedent difference vectors can be plotted as points in the plane
- The red region is their **polyhedron**
- The points which are larger than some point in this polyhedron yield the gray region
Suppose again there are only $n = 2$ constraints:

- the antecedent difference vectors can be plotted as points in the plane
- The red region is their **polyhedron**
- The points which are larger than some point in this polyhedron yield the gray region

If the ME T-order $(x, y) \xrightarrow{\text{ME}} (\hat{x}, \hat{y})$ holds (for any $n$), then:

**C1:** each consequent difference vector lives in this gray region: it is larger than some vector in the polyhedron generated by the antecedent difference vectors

**C2:** the number of candidates of the antecedent ur $x$ is at least as large as the number of candidates of the consequent ur $\hat{x}$
Suppose again there are only $n = 2$ constraints:

- The antecedent difference vectors can be plotted as points in the plane.
- The red region is their **polyhedron**.
- The points which are larger than some point in this polyhedron yield the gray region.

If the ME T-order $(x, y) \xrightarrow{\text{ME}} (\hat{x}, \hat{y})$ holds (for any $n$), then:

**C1:** each consequent difference vector lives in this gray region: it is larger than some vector in the polyhedron generated by the antecedent difference vectors.

**C2:** the number of candidates of the antecedent ur $x$ is at least as large as the number of candidates of the consequent ur $\hat{x}$.

C1 is also sufficient when both $x$ and $\hat{x}$ have at most three candidates.
Suppose again there are only $n = 2$ constraints:

- the antecedent difference vectors can be plotted as points in the plane
- The red region is their **polyhedron**
- The points which are larger than some point in this polyhedron yield the gray region

If the ME T-order $(x, y) \xrightarrow{\text{ME}} (\hat{x}, \hat{y})$ holds (for any $n$), then:

- **C1**: each consequent difference vector lives in this gray region: it is larger than some vector in the polyhedron generated by the antecedent difference vectors
- **C2**: the number of candidates of the antecedent ur $x$ is at least as large as the number of candidates of the consequent ur $\hat{x}$

C1 is also sufficient when both $x$ and $\hat{x}$ have at most three candidates

We derive a more involved sufficient condition which is nonetheless stronger than C1+C2
Phonological applications 1
Geminate devoicing in Japanese

[Kawahara 2006; Pater 2016]
Lyman’s Law: One voiced obstruent per word

Recent loanwords violate Lyman’s Law:
(a) \([\text{bobu}]\) ‘Bob’ \([\text{giga}]\) ‘giga’

Voiced geminates are allowed in recent loanwords:
(b) \([\text{web:u}]\) ‘web’ \([\text{wip:u}]\) ‘whipped (cream)’

Singleton + geminate voiced obstruent yield optional devoicing:
(c) \([\text{gud:o}] \sim [\text{gut:o}]\) ‘good’
\([\text{dog:u}] \sim [\text{dok:u}]\) ‘dog’
Constraints for Japanese (Pater 2016)

**IDENT-VOICE**  Assign a violation for an output segment that differs from its input correspondent in [voice].

**OCP-VOICE**  Assign a violation for a two voiced obstruents within the same word.

**VCE-GEM**  Assign a violation for a voiced obstruent geminate.
### Geminate devoicing in HG (Pater 2016)

<table>
<thead>
<tr>
<th></th>
<th>1.5</th>
<th>1</th>
<th>1</th>
<th>Harmony</th>
</tr>
</thead>
<tbody>
<tr>
<td>/bobu/</td>
<td>IDENT-VOICE</td>
<td>OCP-VOICE</td>
<td>*VCE-GEM</td>
<td></td>
</tr>
<tr>
<td>(1a) bobu</td>
<td>-1</td>
<td></td>
<td></td>
<td>-1</td>
</tr>
<tr>
<td>(1b) bopu/pobu</td>
<td>-1</td>
<td></td>
<td></td>
<td>-1.5</td>
</tr>
<tr>
<td>/web:u/</td>
<td>IDENT-VOICE</td>
<td>OCP-VOICE</td>
<td>*VCE-GEM</td>
<td></td>
</tr>
<tr>
<td>(2a) web:u</td>
<td></td>
<td>-1</td>
<td></td>
<td>-1</td>
</tr>
<tr>
<td>(2b) wep:u</td>
<td></td>
<td>-1</td>
<td></td>
<td>-1.5</td>
</tr>
<tr>
<td>/dog:u/</td>
<td>IDENT-VOICE</td>
<td>OCP-VOICE</td>
<td>*VCE-GEM</td>
<td></td>
</tr>
<tr>
<td>(3a) dog:u</td>
<td></td>
<td>-1</td>
<td></td>
<td>-2</td>
</tr>
<tr>
<td>(3b) dok:u</td>
<td></td>
<td>-1</td>
<td></td>
<td>-1.5</td>
</tr>
</tbody>
</table>
**Predicted typology. OT languages are framed.**

<table>
<thead>
<tr>
<th>OT/HG</th>
<th>OT/HG</th>
<th>OT/HG</th>
<th>OT/HG</th>
<th>HG</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>L2</td>
<td>L3</td>
<td>L4</td>
<td>L5</td>
</tr>
<tr>
<td>bobu</td>
<td>bopu/pobu</td>
<td>bobu</td>
<td>bopu/pobu</td>
<td>bobu</td>
</tr>
<tr>
<td>web:u</td>
<td>web:u</td>
<td>wep:u</td>
<td>wep:u</td>
<td>web:u</td>
</tr>
<tr>
<td>dog:u</td>
<td>dok:u</td>
<td>dok:u</td>
<td>dok:u</td>
<td>dok:u</td>
</tr>
</tbody>
</table>
**T-order: OT = HG**

Graph 1

Graph 2
Theoretical result

OT and HG T-orders can be identical even when the typologies are different.

• In particular, OT and HG T-orders are identical if there are only two candidates (Japanese).

• But this does not exhaust all such cases (e.g., Swedish obstruent voicing, Lombardi 1999, work in progress)
Let’s add two harmonically bounded candidates

<table>
<thead>
<tr>
<th></th>
<th>1.5</th>
<th></th>
<th></th>
<th></th>
<th>Harmony</th>
</tr>
</thead>
<tbody>
<tr>
<td>/bobu/</td>
<td>IDENT-VOICE</td>
<td>OCP-VOICE</td>
<td>*VCE-GEM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1a) bobu</td>
<td></td>
<td>-1</td>
<td></td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>(1b) bopu/pobu</td>
<td>-1</td>
<td></td>
<td></td>
<td>-1.5</td>
<td></td>
</tr>
<tr>
<td>/web:u/</td>
<td>IDENT-VOICE</td>
<td>OCP-VOICE</td>
<td>*VCE-GEM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2a) web:u</td>
<td></td>
<td>-1</td>
<td></td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>(2b) wep:u</td>
<td>-1</td>
<td></td>
<td></td>
<td>-1.5</td>
<td></td>
</tr>
<tr>
<td>/dog:u/</td>
<td>IDENT-VOICE</td>
<td>OCP-VOICE</td>
<td>*VCE-GEM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3a) dog:u</td>
<td></td>
<td>-1</td>
<td></td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>(3b) dok:u</td>
<td>-1</td>
<td></td>
<td></td>
<td>-1.5</td>
<td></td>
</tr>
<tr>
<td>(3c) tog:u</td>
<td>-1</td>
<td></td>
<td></td>
<td>-2.5</td>
<td></td>
</tr>
<tr>
<td>(3d) tok:u</td>
<td>-2</td>
<td></td>
<td></td>
<td>-3</td>
<td></td>
</tr>
</tbody>
</table>
Maxent T-order

Graph 1

Graph 2

{\textlangle bobu, bopu/pobu\textrangle} \rightarrow \{\textlangle dogu, doku\textrangle\} 

{\textlangle webu, wepu\textrangle} \rightarrow \{\textlangle bobu, bobu\textrangle\}

{\textlangle dogu, dogu\textrangle} \rightarrow \{\textlangle webu, webu\textrangle\}
Theoretical result

- In Maxent T-orders, an input with fewer candidates cannot entail one with more candidates.

Even harmonically bounded candidates count: adding or subtracting them affects the T-order.
Phonological applications 2
CV syllabification

[Prince and Smolensky 2004]
# CV syllabification

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
<th>ONSET</th>
<th>*CODA</th>
<th>MAX</th>
<th>DEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>CV</td>
<td>CV</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>CVC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VC</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>CVC</td>
<td>CV</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CVC</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VC</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>VC</td>
<td>CV</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>CVC</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>VC</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>4.</td>
<td>V</td>
<td>CV</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CVC</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>VC</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
OT/HG T-order
Maxent T-order
What does this mean?

**Example**: The fact that \((\text{CVC}, \text{CVC}) \rightarrow (\text{CV}, \text{CV})\) does not hold in Maxent means that there are Maxent grammars where the faithful mapping \((\text{CV}, \text{CV})\) has a smaller probability than the faithful mapping \((\text{CVC}, \text{CVC})\)!

The same holds of the following entailments, which are true in OT/HG, but not in Maxent:

\[
\begin{align*}
(\text{VC}, \text{V}) & \rightarrow (\text{CV}, \text{CV}) \\
(\text{VC}, \text{CVC}) & \rightarrow (\text{CV}, \text{CV}) \\
(\text{V}, \text{V}) & \rightarrow (\text{CV}, \text{CV})
\end{align*}
\]
We now know the following

- OT, HG, and Maxent all have T-orders.
- OT and HG T-orders are identical in certain special cases.
- T-orders can be efficiently computed for all three theories, for HG and Maxent for the first time.
- Maxent T-orders behave in a counterintuitive way: they are sensitive to the number of candidates, including harmonically bounded ones.
- Given the standard CV syllabification constraints, Maxent predicts a non-Jakobsonian syllable typology.
Thank you!
## Summary: OT vs. HG

<table>
<thead>
<tr>
<th>Typology</th>
<th>T-orders</th>
<th>Cands</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>same</td>
<td>same</td>
<td>2</td>
<td>Finnish vowel harmony</td>
</tr>
<tr>
<td>same</td>
<td>same</td>
<td>&gt;2</td>
<td>CV syllable typology</td>
</tr>
<tr>
<td>same</td>
<td>different</td>
<td>2</td>
<td>CANNOT EXIST</td>
</tr>
<tr>
<td>same</td>
<td>different</td>
<td>&gt;2</td>
<td>CANNOT EXIST</td>
</tr>
<tr>
<td>different</td>
<td>same</td>
<td>2</td>
<td>Japanese geminate devoicing</td>
</tr>
<tr>
<td>different</td>
<td>same</td>
<td>&gt;2</td>
<td>Swedish obstruent devoicing</td>
</tr>
<tr>
<td>different</td>
<td>different</td>
<td>2</td>
<td>CANNOT EXIST</td>
</tr>
<tr>
<td>different</td>
<td>different</td>
<td>&gt;2</td>
<td>Spanish obstruent lenition/fortition</td>
</tr>
</tbody>
</table>

Anttila, Arto, and Curtis Andrus. 2006. T-orders. manuscript and software.


