Harmonic Grammar growing into Optimality Theory

Maturation as the strict domination limit (or vice-versa)

including joint work with Klaas Seinhorst (University of Amsterdam)

Tamás Biró

ELTE Eötvös Loránd University

BLINC @ ELTE, June 2, 2017
Learning and maturation

A standard idea in contemporary linguistics:

- Children learning a language:
  setting parameters or re-ranking constraints.
- Hence / because: children speak a typologically different language that resides in the typology of adult languages.

Is it really so?

Goal of this talk: to show how maturation (as opposed to learning) can be included into OT.
Learning and maturation

A standard idea in contemporary linguistics:

- Children learning a language:
  setting parameters or re-ranking constraints.
- Hence / because: children speak a typologically different language that resides in the typology of adult languages.

Is it really so?

Goal of this talk: to show how maturation (as opposed to learning) can be included into OT.
Overview

1. Optimality Theory and Harmonic Grammar
2. From HG to OT: the strict domination limit
3. Consonant cluster simplification in Dutch
4. Summary: learning and maturation
Overview

1. Optimality Theory and Harmonic Grammar
2. From HG to OT: the strict domination limit
3. Consonant cluster simplification in Dutch
4. Summary: learning and maturation
Optimality Theory in a broad sense

- Underlying representation $\rightarrow$ candidate set.
- Surface representation $=$ optimal element of candidate set.
- Optimality: most harmonic. What is “harmony”?
Optimality Theory in a narrow sense

\[ \text{Gen}(\sigma\sigma\sigma\sigma) = \{[\text{suuu}], [\text{usuu}], [\text{uusu}], [\text{uuus}]\}. \]

<table>
<thead>
<tr>
<th>/σσσσ/</th>
<th>EARLY</th>
<th>LATE</th>
<th>NONFINAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>[s u u u]</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>[u s u u]</td>
<td>1!</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>[u u s u]</td>
<td>2!</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>[u u u s]</td>
<td>3!</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

\[ \text{SF}(\sigma\sigma\sigma\sigma) = [\text{suuu}]. \]
Optimality Theory in a narrow sense

\[ \text{Gen} (\sigma\sigma\sigma) = \{[suuu], [usuu], [uusu], [uuus]\}. \]

<table>
<thead>
<tr>
<th>/\sigma\sigma\sigma/</th>
<th>NONFINAL</th>
<th>LATE</th>
<th>EARLY</th>
</tr>
</thead>
<tbody>
<tr>
<td>[s u u u]</td>
<td>0</td>
<td>3!</td>
<td>0</td>
</tr>
<tr>
<td>[u s u u]</td>
<td>0</td>
<td>2!</td>
<td>1</td>
</tr>
<tr>
<td>[u u s u]</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>[u u u s]</td>
<td>1!</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

\[ \text{SF}(\sigma\sigma\sigma) = [uuus]. \]
Optimality Theory in a narrow sense

\[ \text{Gen}(\sigma\sigma\sigma\sigma) = \{[suuu], [usuu], [uusu], [uuus]\}. \]

<table>
<thead>
<tr>
<th>/σσσσ/</th>
<th>NONFINAL</th>
<th>LATE</th>
<th>EARLY</th>
</tr>
</thead>
<tbody>
<tr>
<td>[u u s u]</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>[u s u u]</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>[s u u u]</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>[u u u s]</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Harmony in terms of **lexicographic order**:

[uusu] ≻ [usuu] ≻ [suuu] ≻ [uuus]. Hence, SF(σσσσ) = [uusu].
Harmonic Grammar, symbolic approach

\[ \text{Gen}(\sigma\sigma\sigma\sigma) = \{[suuu], [usu], [usu], [uuus]\} \].

Constraint \( C_k \) is assigned weight \( w_k \). Harmony \( H(A) \) of cand. \( A \):

\[
\begin{array}{|c|c|c|c|c|}
\hline
/\sigma\sigma\sigma\sigma/ & \text{NONFINAL} & \text{LATE} & \text{EARLY} & -H(A) = \\
\hline
w_k = & 9 & 3 & 1 & \sum_k w_k \cdot C_k[A] \\
\hline
[u uu uu] & 0 & 1 & 2 & 9 \cdot 0 + 3 \cdot 1 + 1 \cdot 2 = 5 \\
[u s u uu] & 0 & 2 & 1 & 9 \cdot 0 + 3 \cdot 2 + 1 \cdot 1 = 7 \\
[s uu uu] & 0 & 3 & 0 & 9 \cdot 0 + 3 \cdot 3 + 1 \cdot 0 = 9 \\
[u uu us] & 1 & 0 & 3 & 9 \cdot 1 + 3 \cdot 0 + 1 \cdot 3 = 12 \\
\hline
\end{array}
\]

By comparing the integer/real numbers in \( H \):

\( [usu] \succ [usu] \succ [suuu] \succ [uuus] \). Hence, \( \text{SF}(\sigma\sigma\sigma\sigma) = [usu] \).
Harmonic Grammar, connectionist approach

Boltzmann machine:

- The “energy” (negative harmony) of the network:
  \[ -H(A) = \sum_{i,j=1}^{N} s_i \cdot W_{i,j} \cdot s_j \]

- Input nodes clamped (fixed).
- Output nodes read after optimization of \( H(A) \).

- \( s_i \): activation of node \( i \).
- \( W_{i,j} \): connection strength between nodes \( i \) and \( j \).
Constraint $C_k$ is a set of $W_{i,j}$ partial connection strengths.

The “energy” (negative harmony) of the network:

$$-H(A) = \sum_{i,j=1}^{N} s_i \cdot W_{i,j} \cdot s_j = \sum_{k=1}^{N} w_k \cdot C_k[A] = \sum_{k=1}^{n} w_k \cdot W^k_{i,j} = \sum_{k=1}^{n} w_k \cdot W^k_{i,j} = \sum_{k=1}^{n} w_k \cdot C_k[A].$$
Harmonic Grammar, connectionist approach

Constraint $C_k$ is a set of $W_{i,j}$ partial connection strengths.

$$W_{i,j} = \sum_{k=1}^{n} w_k \cdot W_{i,j}^k$$

The “energy” (negative harmony) of the network:

$$-H(A) = \sum_{i,j=1}^{N} s_i \cdot W_{i,j} \cdot s_j = \sum_{k=1}^{N} w_k \cdot \sum_{i,j=1}^{n} s_i \cdot W_{i,j}^k \cdot s_j = \sum_{k=1}^{n} w_k \cdot C_k[A].$$
Summary thus far:

We have three approaches:

- **Optimality Theory**: symbolic, optimization w.r.t. lexicographic order.

- **Symbolic Harmonic Grammar**: optimization w.r.t. $\geq$ relation among real numbers.

- **Connectionist Harmonic Grammar**: optimization w.r.t. $\geq$ relation among real numbers.

*Typological predictions? Learnability? Cognitive plausibility?*

**Question**: how to get OT in a “connectionist brain”?

(A major issue for Smolensky’s *Integrated Connectionist/Symbolic Architecture.*)
Summary thus far:

We have three approaches:

- **Optimality Theory**: symbolic, optimization w.r.t. lexicographic order.

- **Symbolic Harmonic Grammar**: optimization w.r.t. $\geq$ relation among real numbers.

- **Connectionist Harmonic Grammar**: optimization w.r.t. $\geq$ relation among real numbers.

Typological predictions? Learnability? **Cognitive plausibility?**

**Question**: how to get OT in a “connectionist brain”?

(A major issue for Smolensky’s *Integrated Connectionist/Symbolic Architecture.*)
Summary thus far:

We have three approaches:

- Optimality Theory: symbolic, optimization w.r.t. lexicographic order.

- Symbolic Harmonic Grammar: optimization w.r.t. $\geq$ relation among real numbers.

- Connectionist Harmonic Grammar: optimization w.r.t. $\geq$ relation among real numbers.

Typological predictions? Learnability? Cognitive plausibility?

Question: how to get OT in a “connectionist brain”? (A major issue for Smolensky’s Integrated Connectionist/Symbolic Architecture.)
Overview

1. Optimality Theory and Harmonic Grammar
2. From HG to OT: the strict domination limit
3. Consonant cluster simplification in Dutch
4. Summary: learning and maturation
Exponential Harmonic Grammar, or $q$-HG

- **Optimality Theory** minimizes a vector of violations:

\[
H(A) = \begin{bmatrix}
C_n & C_{n-1} & \ldots & C_i & \ldots & C_1 \\
r_n (= n) & r_{n-1} & \ldots & r_i & \ldots & r_1 (= 1) \\
C_1[A] & C_2[A] & \ldots & C_i[A] & \ldots & C_n[A]
\end{bmatrix}
\]

- **Harmonic Grammar** minimizes a weighted sum of violations:

\[
H(A) = \sum_{i=1}^{n} w_i \cdot C_i[A].
\]

- “Standard” HG: weights $w_i = \text{ranks } r_i$.
- **Exponential HG**: weights are ranks exponentiated, fixed base ($e = 2.7182\ldots$) $w_i = e^{r_i}$.
- **$q$-HG**: weights are ranks exponentiated, with (variable) base $q$ $w_i = q^{r_i}$. 
Exponential Harmonic Grammar, or \( q \)-HG

- **Optimality Theory** minimizes a vector of violations:

\[
H(A) = \begin{array}{ccccccc}
C_n & C_{n-1} & \ldots & C_i & \ldots & C_1 \\
\frac{r_n}{r_{n-1}} & \frac{r_{n-1}}{r_{i-1}} & \ldots & \frac{r_i}{r_{i-1}} & \ldots & \frac{r_1}{r_1}
\end{array}
\]

- **Harmonic Grammar** minimizes a weighted sum of violations:

\[
H(A) = \sum_{i=1}^{n} w_i \cdot C_i[A].
\]

- "Standard" HG: weights \( w_i = \text{ranks } r_i \).
- Exponential HG: weights are ranks exponentiated, fixed base \( e = 2.7182\ldots \)
  \( w_i = e^{r_i} \).
- \( q \)-HG: weights are ranks exponentiated, with (variable) base \( q \)
  \( w_i = q^{r_i} \).
Exponential Harmonic Grammar, or $q$-HG

- **Optimality Theory** minimizes a vector of violations:

\[
H(A) = \begin{pmatrix}
C_n & C_{n-1} & \ldots & C_i & \ldots & C_1 \\
\mathsf{r}_n(=n) & \mathsf{r}_{n-1} & \ldots & \mathsf{r}_i & \ldots & \mathsf{r}_1(=1)
\end{pmatrix}
\]

- **Harmonic Grammar** minimizes a weighted sum of violations:

\[
H(A) = \sum_{i=1}^{n} w_i \cdot C_i[A].
\]

- “Standard” HG: weights $w_i = \text{ranks } \mathsf{r}_i$.
- **Exponential HG**: weights are ranks exponentiated, fixed base ($e \approx 2.7182$ . . .)
  \[w_i = e^{\mathsf{r}_i}.\]
- **$q$-HG**: weights are ranks exponentiated, with (variable) base $q$
  \[w_i = q^{\mathsf{r}_i}.\]
Exponential Harmonic Grammar, or $q$-HG

**Optimality Theory** minimizes a vector of violations:

$$H(A) = \begin{bmatrix}
C_n & C_{n-1} & \ldots & C_i & \ldots & C_1 \\
r_n (= n) & r_{n-1} & \ldots & r_i & \ldots & r_1 (= 1) \\
C_1[A] & C_2[A] & \ldots & C_i[A] & \ldots & C_n[A]
\end{bmatrix}$$

**Harmonic Grammar** minimizes a weighted sum of violations:

$$H(A) = \sum_{i=1}^{n} w_i \cdot C_i[A].$$

- "Standard" HG: weights $w_i = \text{ranks } r_i$.
- **Exponential HG**: weights are ranks exponentiated, fixed base ($e = 2.7182 \ldots$) $w_i = e^{r_i}$.
- **$q$-HG**: weights are ranks exponentiated, with (variable) base $q$ $w_i = q^{r_i}$. 
Strict domination in OT is $q$-HG in the $q \to +\infty$ limit

- In $q$-HG:
  \[
  H(A) = q^r_n \cdot C_n[A] + \ldots + q^r_i \cdot C_i[A] + \ldots + q^r_1 \cdot C_1[A]
  \]

- Or simply (if $r_i = i - 1$):
  \[
  H(A) = q^{n-1} \cdot C_n[A] + \ldots + q^2 \cdot C_i[A] + q^1 \cdot C_i[A] + q^0 \cdot C_1[A]
  \]
  \[
  H(A) = 2^{n-1} \cdot C_n[A] + \ldots + 4 \cdot C_i[A] + 2 \cdot C_i[A] + 1 \cdot C_1[A]
  \]
  \[
  H(A) = 3^{n-1} \cdot C_n[A] + \ldots + 9 \cdot C_i[A] + 3 \cdot C_i[A] + 1 \cdot C_1[A]
  \]
  \[
  H(A) = 10^{n-1} \cdot C_n[A] + \ldots + 100 \cdot C_i[A] + 10 \cdot C_i[A] + 1 \cdot C_1[A]
  \]

- Main difference between OT and HG is strict domination.

- If $q$ grows large, $q$-HG turns into OT, because…
Strict domination in OT is $q$-HG in the $q \to +\infty$ limit

- In $q$-HG:

  $$H(A) = q^n \cdot C_n[A] + \ldots + q^i \cdot C_i[A] + \ldots + q^1 \cdot C_1[A]$$

- Or simply (if $r_i = i - 1$):

  $$H(A) = q^{n-1} \cdot C_n[A] + \ldots + q^2 \cdot C_i[A] + q^1 \cdot C_i[A] + q^0 \cdot C_1[A]$$

  $$H(A) = 2^{n-1} \cdot C_n[A] + \ldots + 4 \cdot C_i[A] + 2 \cdot C_i[A] + 1 \cdot C_1[A]$$

  $$H(A) = 3^{n-1} \cdot C_n[A] + \ldots + 9 \cdot C_i[A] + 3 \cdot C_i[A] + 1 \cdot C_1[A]$$

  $$H(A) = 10^{n-1} \cdot C_n[A] + \ldots + 100 \cdot C_i[A] + 10 \cdot C_i[A] + 1 \cdot C_1[A]$$

- Main difference between OT and HG is strict domination.

- If $q$ grows large, $q$-HG turns into OT, because...
Strict domination in OT is $q$-HG in the $q \to +\infty$ limit

- 1.5-HG has *ganging-up cumulativity*:

<table>
<thead>
<tr>
<th>$w = \left[ \begin{array}{c} C_3 \ C_2 \ C_1 \ H \end{array} \right]$</th>
<th>2.25</th>
<th>1.5</th>
<th>1</th>
<th>2.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A1$</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A2$</td>
<td>1</td>
<td>1</td>
<td></td>
<td>2.5</td>
</tr>
</tbody>
</table>

- 1.5-HG also has *counting cumulativity*:

<table>
<thead>
<tr>
<th>$w_i = \left[ \begin{array}{c} C_3 \ C_2 \ C_1 \ H \end{array} \right]$</th>
<th>2.25</th>
<th>1.5</th>
<th>1</th>
<th>2.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A1$</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A3$</td>
<td>1</td>
<td>2</td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

(Cf. Jäger and Rosenbach 2006)
Strict domination in OT is $q$-HG in the $q \rightarrow +\infty$ limit

- But OT does not have *ganging-up cumulativity*:

<table>
<thead>
<tr>
<th></th>
<th>$C_3$</th>
<th>$C_2$</th>
<th>$C_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>

- OT does not have *counting cumulativity* either:

<table>
<thead>
<tr>
<th></th>
<th>$C_3$</th>
<th>$C_2$</th>
<th>$C_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td></td>
<td></td>
<td>**</td>
</tr>
</tbody>
</table>

(Regarding Stochastic OT, cf. Jäger and Rosenbach 2006)
Strict domination in OT is $q$-HG in the $q \rightarrow +\infty$ limit

- 3-HG does not have *ganging-up cumulativity*:

  $\begin{array}{cccc}
  & C_3 & C_2 & C_1 & H \\
  A1 & 9 & 3 & 1 & \\
  \text{A2} & 1 & 1 & 4 & \\
  \end{array}$

- 3-HG does not have *counting cumulativity*, either:

  $\begin{array}{cccc}
  & C_3 & C_2 & C_1 & H \\
  A1 & 9 & 3 & 1 & \\
  \text{A3} & 1 & 2 & \\
  \end{array}$

(Cf. Jäger and Rosenbach 2006)
Strict domination in OT is $q$-HG in the $q \to +\infty$ limit

As we have known it since Prince and Smolensky 1993, strict domination in OT can be reproduced using $q$-HG with sufficiently large $q$:

$$q - 1 \geq C_k[A] \quad \text{for all } k \text{ and } A.$$
Overview

1. Optimality Theory and Harmonic Grammar
2. From HG to OT: the strict domination limit
3. Consonant cluster simplification in Dutch
4. Summary: learning and maturation
Word initial consonant cluster simplification in Dutch

Klaas Seinhorst collecting data from CHILDES (Laura):

Cf. Becker and Tessier (2011)
Word initial consonant cluster simplification in Dutch

Using logistic regression or probit regression:

<table>
<thead>
<tr>
<th>cluster</th>
<th>simplifies to</th>
<th>lower boundary (age in days)</th>
<th>upper boundary (age in days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>kl-</td>
<td>k-</td>
<td>894.32</td>
<td>1010.55</td>
</tr>
<tr>
<td>sl-</td>
<td>l-</td>
<td>943.82</td>
<td>1028.68</td>
</tr>
<tr>
<td>st-</td>
<td>t-</td>
<td>962.60</td>
<td>1076.23</td>
</tr>
<tr>
<td>zw-</td>
<td>z-</td>
<td>1344.24</td>
<td>1551.39</td>
</tr>
</tbody>
</table>

Table: 95% confidence intervals of the locations of the inflection points.

Differences among kl, sl and st: statistically not significant. But differences between zw and each of the three others: $p < 10^{-11}$!
Word initial consonant cluster simplification: OT

The traditional account in OT: learning $\equiv$ constraint re-ranking.

- **Before learning: Markedness $\gg$ Faithfulness**

<table>
<thead>
<tr>
<th>/klɛin/</th>
<th>NO\textsc{Complex} Onset</th>
<th>FAITHF</th>
<th>*[l]</th>
<th>*[k]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[klɛin]</td>
<td>!</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[kɛin]</td>
<td>!</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[lɛin]</td>
<td>!</td>
<td>*</td>
<td>*</td>
<td>!</td>
</tr>
</tbody>
</table>

- **After learning: Faithfulness $\gg$ Markedness**

<table>
<thead>
<tr>
<th>/klɛin/</th>
<th>FAITHF</th>
<th>NO\textsc{Complex} Onset</th>
<th>*[l]</th>
<th>*[k]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[klɛin]</td>
<td>!</td>
<td>!</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[kɛin]</td>
<td>!</td>
<td>!</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[lɛin]</td>
<td>!</td>
<td>!</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
Questions to the traditional account:

- Child is exposed to huge amount of evidence way before correct production. Why no learning?

- If only *NoComplexOnset* and *Faithf* are involved, why significant difference for *zw* onset?

- If cluster-specific constraints: factorial typology predicted.
Word initial consonant cluster simplification: $q$-HG

An alternative approach:

- Child has acquired $\text{FAITHF} \gg \text{NOCOMPLEXONSET}$ much earlier, probably already at pre-linguistic age.
- Relative ranks $*[w] \gg *[s] \gg *[l] \gg *[z] \gg *[k] \gg *[t]$ motivated by *natural phonology* (? feedback appreciated!).
- No more ranking needed. For instance,

<table>
<thead>
<tr>
<th>$C_i$</th>
<th>$\text{FAITHF}$</th>
<th>$\text{NOCOMPLEXONSET}$</th>
<th>$*[w]$</th>
<th>$*[s]$</th>
<th>$*[l]$</th>
<th>$*[z]$</th>
<th>$*[k]$</th>
<th>$*[t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_i$</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$(1.1)^{r_i}$</td>
<td>2.14</td>
<td>1.95</td>
<td>1.77</td>
<td>1.61</td>
<td>1.46</td>
<td>1.33</td>
<td>1.21</td>
<td>1.1</td>
</tr>
<tr>
<td>$2^{r_i}$</td>
<td>256</td>
<td>128</td>
<td>64</td>
<td>32</td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>
Word initial consonant cluster simplification: $q$-HG

Before maturation: small $q$, e.g., $q = 1.1$ (NB: Faithfulness $\gg$ Markedness!)

<table>
<thead>
<tr>
<th>/klɛin/</th>
<th>FAITHF</th>
<th>NoCOMPLCONS</th>
<th>*[l]</th>
<th>*[k]</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>w$_i$ =</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[klɛin]</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>4.62</td>
</tr>
<tr>
<td>[kɛin]</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
<td>3.35</td>
</tr>
<tr>
<td>[lɛin]</td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
<td>3.60</td>
</tr>
</tbody>
</table>

After maturation: large $q$, e.g., $q = 2$

<table>
<thead>
<tr>
<th>/klɛin/</th>
<th>FAITHF</th>
<th>NoCOMPLCONS</th>
<th>*[l]</th>
<th>*[k]</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>w$_i$ =</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[klɛin]</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>148</td>
</tr>
<tr>
<td>[kɛin]</td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
<td>260</td>
</tr>
<tr>
<td>[lɛin]</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td>272</td>
</tr>
</tbody>
</table>
Word initial consonant cluster simplification: $q$-HG

- $q$ is a function of age, e.g. age $\propto \log(q)$.

- $[xy]$ produced by $q$-HG, if $q$ is s.t. $q^c + q^x + q^y = q^f + q^y$ or larger:

<table>
<thead>
<tr>
<th>/xy/</th>
<th>FaithF</th>
<th>*ComplOns</th>
<th>*[x]</th>
<th>*[y]</th>
<th>$H$ for given $q$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r_i =$</td>
<td>$f$</td>
<td>$c$</td>
<td>$x$</td>
<td>$y$</td>
</tr>
<tr>
<td></td>
<td>$w_i =$</td>
<td>$q^f$</td>
<td>$q^c$</td>
<td>$q^x$</td>
<td>$q^y$</td>
</tr>
<tr>
<td>[xy]</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$q^c + q^x + q^y$</td>
</tr>
<tr>
<td>[y]</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>$q^f + q^y$</td>
</tr>
<tr>
<td>[x]</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$q^f + q^x$</td>
</tr>
</tbody>
</table>

- Critical age function of deleted segment [x], but not remaining [y].
- If $f > c$, $x > y$, then: step function predicted.
- To get S-shaped curve, use Stochastic OT.
Word initial consonant clusters: stochastic $q$-HG
Overview

1. Optimality Theory and Harmonic Grammar
2. From HG to OT: the strict domination limit
3. Consonant cluster simplification in Dutch
4. Summary: learning and maturation
Learning vs. Maturation?

Learning:
- Knowledge acquired from surrounding linguistic data
- Source of cross-linguistic variation
- Features in the child’s language shared by other adult languages

Maturation:
- Skills emerging due to general development
- Universal developmental paths
- Features in child’s language not appearing in any adult language
Learning vs. Maturation?

**Learning:**
- Knowledge acquired from surrounding linguistic data
- Source of cross-linguistic variation
- Features in the child’s language shared by other adult languages

**Maturation:**
- Skills emerging due to general development
- Universal developmental paths
- Features in child’s language not appearing in any adult language
Learning vs. Maturation?

Learning from surrounding linguistic data:
- Features in the child’s language shared by other adult languages
  - Child learning English produces “Italian-like” pro-drop
    → “Pro-drop” parameter not yet switched.
  - Child learning English deleting codas
    → *CODA markedness not yet demoted below FAITHFULNESS.

Maturation due to general development:
- Features in child’s language not appearing in any adult language
  - Long distance place agreement in consonant harmony?
  - Erroneous pronoun resolution?
Learning vs. Maturation?

Learning from surrounding linguistic data:
- Features in the child’s language shared by other adult languages
  - Child learning English produces “Italian-like” pro-drop
    → “Pro-drop” parameter not yet switched.
  - Child learning English deleting codas
    → *CODA markedness not yet demoted below FAITHFULNESS.

Maturation due to general development:
- Features in child’s language not appearing in any adult language
  - Long distance place agreement in consonant harmony?
  - Erroneous pronoun resolution?
Learning vs. Maturation!

Modelling learning and modelling maturation: shouldn’t they be different?

Learning from surrounding linguistic data:
- Setting parameters
- Re-ranking constraints

Maturation due to general development:
- Restrictions on working memory, speed of mental computation…
- Varying $q$ in $q$-HG?
Learning vs. Maturation!

Modelling learning and modelling maturation: shouldn’t they be different?

Learning from surrounding linguistic data:
- Setting parameters
- Re-ranking constraints

Maturation due to general development:
- Restrictions on working memory, speed of mental computation…
- Varying $q$ in $q$-HG?
Thank you for your attention!

Tamás Biró:
http://birot.web.elte.hu/

This research was supported by a Veni Grant (project number 275-89-004) offered by the Netherlands Organisation for Scientific Research (NWO), as well as by a Marie Curie FP7 Integration Grant (no. PCIG14-GA-2013-631599, “MeMoLI”), 7th EU Framework Programme.