Being stressed by stress
Competence, performance and learning
of metrical stress
in Optimality Theory

Tamás Biró

ACLC, University of Amsterdam (UvA)

BGU, March 13, 2012
How to observe stress pattern in (Dutch) fast speech?

Quiz item:

Q4  President Bush is een typische

A1  intellectueel
A2  amerikaan
A3  taalkundige

‘President Bush is a typical’
‘intellectual’
‘American’
‘linguist’

How to observe stress pattern in (Dutch) fast speech?

Quiz item:

Q4  President Bush is een typische 'President Bush is a typical'

A1  intellectueel 'intellectual'
A2  amerikaan 'American'
A3  taalkundige 'linguist'

How to observe stress pattern in (Dutch) fast speech?

Forms and frequencies observed by Schreuder and Gilbers:

<table>
<thead>
<tr>
<th>Word</th>
<th>Fast: 0.67</th>
<th>Fast: 0.38</th>
<th>Slow: 0.13</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘camera’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘publisher’</td>
<td>0.96</td>
<td>0.81</td>
<td>0.20</td>
</tr>
<tr>
<td>‘study grant’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘perfectionist’</td>
<td>0.33</td>
<td>0.62</td>
<td>0.87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Word</th>
<th>Fast: 0.33</th>
<th>Slow: 0.04</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘study grant’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘beat reduction’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Shift to right | Shift to right | Shift to left
Stressing questions

- Account for stress patterns within some linguistic framework.
- Account for changing patterns: normal vs. fast speech.
  - Account for fast speech forms. Account for their frequencies.
  - Both forms present. Different word types have different frequencies.
  - Gradual change between ‘normal’ and ‘fast’.
- Account for language acquisition:
  - Despite difficulties caused by framework (*viz.* feet).
  - (Despite difficulties caused by speech errors.)
- (Account for language change: iterated learning.)
Stressing questions

- Account for stress patterns within some linguistic framework.

- Account for changing patterns: normal vs. fast speech.
  - Account for fast speech forms. Account for their frequencies.
  - Both forms present. Different word types have different frequencies.
  - Gradual change between ‘normal’ and ‘fast’.

- Account for language acquisition:
  - Despite difficulties caused by framework (viz. feet).
  - (Despite difficulties caused by speech errors.)

- (Account for language change: iterated learning.)
Stressing questions

- Account for stress patterns within some linguistic framework.
- Account for changing patterns: normal vs. fast speech.
  - Account for fast speech forms. Account for their frequencies.
  - Both forms present. Different word types have different frequencies.
  - Gradual change between ‘normal’ and ‘fast’.
- Account for language acquisition:
  - Despite difficulties caused by framework (viz. feet).
  - (Despite difficulties caused by speech errors.)
- (Account for language change: iterated learning.)
Stressing questions

- Account for stress patterns within some linguistic framework.

- Account for changing patterns: normal vs. fast speech.
  - Account for fast speech forms. Account for their frequencies.
  - Both forms present. Different word types have different frequencies.
  - Gradual change between ‘normal’ and ‘fast’.

- Account for language acquisition:
  - Despite difficulties caused by framework (viz. feet).
  - (Despite difficulties caused by speech errors.)

- (Account for language change: iterated learning.)
Overview

1. Optimality Theory: competence
2. Simulated Annealing for Optimality Theory: performance
3. Learning in Optimality Theory: problems
4. Conclusions
Overview

1. Optimality Theory: competence
2. Simulated Annealing for Optimality Theory: performance
3. Learning in Optimality Theory: problems
4. Conclusions
Traditional Optimality Theory

- Underlying form → a set of candidates (potential surface forms).
- Constraints: elementary functions defined on the candidate set.
- Grammatical form predicted by this grammar: candidate that violates the highest ranked constraints the least.
- Serial evaluation approach:
Traditional Optimality Theory

- Underlying form $\mapsto$ a set of candidates (potential surface forms).
- Constraints: elementary functions defined on the candidate set.
- Grammatical form predicted by this grammar: candidate that violates the highest ranked constraints the least.

Grammar = constraint hierarchy

- Optimality Theory is a P&P-style model of competence.
- Constraints are universal “principles”.
- Constraint ranking is language-dependent “parameter”.

Tamás Biró Being stressed by stress 8/37
Traditional Optimality Theory

Prosodic phonology:

\[
\begin{bmatrix}
\text{IP} \\
Wd \ 
\sigma \  [\dot{\sigma}\sigma] \ [\dot{\sigma}] \ 
\sigma \  [\sigma\dot{\sigma}]
\end{bmatrix}
\]

- Intonational Phrase → Prosodic Word → Foot → Syllable → ...
- Stressed syllable: head (‘strong’) syllable of a word.
- Main stress: head syllable of head (‘strong’) foot.
Traditional Optimality Theory

Proposal of Schreuder and Gilbers:

Andante speech:

<table>
<thead>
<tr>
<th>/fóto+toestel/</th>
<th>OUTPUT-OUTPUT CORRESPONDENCE</th>
<th>FOOT REPULSION</th>
<th>PARSE SYLLABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>(fóto)(tòestel)</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>(fóto)toe(štel)</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

NB: Further output forms in the language: fóto and tóestel.
Traditional Optimality Theory

Proposal of Schreuder and Gilbers:

Allegro speech:

<table>
<thead>
<tr>
<th>/fóto+toestel/</th>
<th>Foot Repulsion</th>
<th>Output-Output Correspondence</th>
<th>Parse Syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>(fóto)(tòestel)</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(fóto)toe(štel)</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

NB: Further output forms in the language: *fóto* and *tóestel*. 

Tamás Biró

Being stressed by stress
Traditional Optimality Theory

Grammar = constraint hierarchy

- Different grammar for fast speech. Hence, different competence in fast speech?
- How to account for both forms produced?
  - frequencies depending on speech rate, and
  - frequencies depending on word.

Constraint re-ranking

- Ad hoc re-ranking, motivated by analogy in music.
- Principled framework: Boersma’s *Stochastic Optimality Theory.*
Traditional Optimality Theory

Grammar = constraint hierarchy

- Different grammar for fast speech. Hence, different competence in fast speech?
- How to account for both forms produced?
  - frequencies depending on speech rate, and
  - frequencies depending on word.

Constraint re-ranking

- Ad hoc re-ranking, motivated by analogy in music.
- Principled framework: Boersma’s *Stochastic Optimality Theory*. 
Traditional Optimality Theory

Grammar = constraint hierarchy

- Different grammar for fast speech. Hence, different competence in fast speech?
- How to account for both forms produced?
  - frequencies depending on speech rate, and
  - frequencies depending on word.

Constraint re-ranking

- Ad hoc re-ranking, motivated by analogy in music.
- Principled framework: Boersma’s Stochastic Optimality Theory.
**Traditional Optimality Theory**

Grammar = constraint hierarchy

- Different grammar for fast speech. Hence, different competence in fast speech?
- How to account for both forms produced?
  - frequencies depending on speech rate, and
  - frequencies depending on word.

**Constraint re-ranking**

- Ad hoc re-ranking, motivated by analogy in music.
- Principled framework: Boersma’s *Stochastic Optimality Theory.*
Overview

1. Optimality Theory: competence
2. Simulated Annealing for Optimality Theory: performance
3. Learning in Optimality Theory: problems
4. Conclusions
Errors of the mental computation

COMPETENCE

PERFORMANCE

static knowledge
Optimality Theory

processing in the brain
Simulated Annealing for OT
Errors of the mental computation

static knowledge  processing in the brain
Optimality Theory  Simulated Annealing for OT
Errors of the mental computation

static knowledge
Optimality Theory

processing in the brain
Simulated Annealing for OT
Errors of the mental computation

**Optimality Theory**

Grammar competence model
grammatical form = $E^*$ (globally) optimal candidate

SA-OT implementation performance model
produced forms = globally or locally optimal candidates

A grammar is a Harmony function on the candidate set, defined by the ranked constraints. Global optimum: more harmonic than all other candidates. Local optimum: more harmonic than its neighbours.
Basic idea of Simulated Annealing

Step 1 – introducing landscape:

- **Horizontal**: universal *neighbourhood structure* (a.k.a. *topology*) on the universal candidate set.
- **Vertical**: grammar-dependent harmony (violation profile of the constraints).
- Random walk in this landscape.
Basic idea of Simulated Annealing

Step 2 – walking in this landscape:

- Pick a random neighbour of your position.
- If neighbour is more optimal: move.
- If less optimal: move in the beginning, don’t move later.
  (Exponential expression applied to vector-valued target function.)
Basic idea of Simulated Annealing

Step 3 – performing a random walk on this landscape:

- Start random walk from some initial position.
- End position returned as output of the algorithm: produced form.
- Hopefully, global optimum (grammatical form) is found. Yet...
- Neighbourhood structure → local optima, where random walker can get stuck. Performance errors.
Step 4 – Precision of the algorithm

- **Precision** of the algorithm: chance of ending up in global optimum, and hence returning grammatical form.
  - Precision of the algorithm depends on its speed.
  - Trade precision for speed – just like human mind!
<table>
<thead>
<tr>
<th>Level</th>
<th>its product</th>
<th>its model</th>
<th>the product in the model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competence in narrow sense: static knowledge of the language</td>
<td>grammatical form</td>
<td>standard OT grammar</td>
<td>globally optimal candidate</td>
</tr>
<tr>
<td>Dynamic language production process</td>
<td>acceptable or attested forms</td>
<td>SA-OT algorithm</td>
<td>local optima</td>
</tr>
<tr>
<td>Performance in its outmost sense</td>
<td>acoustic signal, etc.</td>
<td>(phonetics, pragmatics)</td>
<td>??</td>
</tr>
</tbody>
</table>
How to predict stress pattern in (Dutch) fast speech?

Landscape:

Basic steps that connect neighbours:
- Move foot boundary: [s] ⇔ [su]; [s] ⇔ [us].
- Change head of foot: [su] ⇔ [us].
- Insert/delete monosyllabic foot: [s] ⇔ u
How to predict stress pattern in (Dutch) fast speech?

Hierarchy:

\[ \text{ALIGN-LEFT} \gg \text{OOC}_{z=2} \gg \text{FOOTREPULSION} \gg \text{PARSESYLL} \gg \text{TROCHAIC} \]

- Global optimum: [s]u[su].
- Local optima: [s]u[su] and [su]u[s].
- Local optimum [su]u[s] has less harmonic neighbours: [su]uu, [su][s][s], [us]u[s], [s]uu[s], [su][us].
How to predict stress pattern in (Dutch) fast speech?

Local optimum [su]u[s] has less harmonic neighbours: [su]uu, [su][s][s], [us]u[s], [s]uu[s], [su][us].

<table>
<thead>
<tr>
<th>Word</th>
<th>ALIGN-LEFT</th>
<th>OOC_{z=2}</th>
<th>FTRPULS</th>
<th>PARSE S</th>
<th>TROCHAIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>~</td>
<td>[su]u[s]</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>[su]uu</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>[su][s][s]</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>[us]u[s]</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>[s]uu[s]</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>[su][us]</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
How to predict stress pattern in (Dutch) fast speech?

http://www.birot.hu/sa-ot/

http://www.birot.hu/OTKit/
How to predict stress pattern in (Dutch) fast speech?

<table>
<thead>
<tr>
<th></th>
<th>fo.to.toe.stel</th>
<th>uit.ge.ve.rij</th>
<th>stu.die.toe.la.ge</th>
<th>per.fec.tio.nist</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>‘camera’</td>
<td>‘publisher’</td>
<td>‘study grant’</td>
<td>‘perfectionist’</td>
</tr>
<tr>
<td>susu</td>
<td>suuu</td>
<td>susuu</td>
<td>usus</td>
<td></td>
</tr>
<tr>
<td>fó.to.tòe.stel</td>
<td>fast: 0.82</td>
<td>fast: 0.65 / 0.67</td>
<td>fast: 0.55 / 0.38</td>
<td>fast: 0.49 / 0.13</td>
</tr>
<tr>
<td></td>
<td>slow: 1.00</td>
<td>slow: 0.97 / 0.96</td>
<td>slow: 0.96 / 0.81</td>
<td>slow: 0.91 / 0.20</td>
</tr>
<tr>
<td>fó.to.toe.stèl</td>
<td>fast: 0.18</td>
<td>fast: 0.35 / 0.33</td>
<td>fast: 0.45 / 0.62</td>
<td>fast: 0.39 / 0.87</td>
</tr>
<tr>
<td></td>
<td>slow: 0.00</td>
<td>slow: 0.03 / 0.04</td>
<td>slow: 0.04 / 0.19</td>
<td>slow: 0.07 / 0.80</td>
</tr>
</tbody>
</table>

Simulated / observed (Schreuder) frequencies.

In the simulations, $T_{\text{step}} = 3$ used for fast speech and $T_{\text{step}} = 0.1$ for slow speech.
Overview

1. Optimality Theory: competence

2. Simulated Annealing for Optimality Theory: performance

3. Learning in Optimality Theory: problems

4. Conclusions
The language acquisition problem

COMPETENCE

PERFORMANCE

COMPETENCE

PERFORMANCE

Tamás Biró

Being stressed by stress
Learning from competence?
Learning from performance!
Learning in Optimality Theory

General idea:

- Speaker-teacher wants to say *underlying form* $uf$.
- Speaker-teacher’s grammar produces *surface form* $sf$.
- Listener-learner hears *surface form* $sf = winner form$ $w$.
- Listener-learner’s grammar would produce $uf$ as *loser form* $l$.
- Listener-learner updates her grammar, in order to produce $w$, and not $l$:

  Winner-preferring constraints are promoted and loser-preferring constraints are demoted in hierarchy hypothesized by the learner.
Learning hidden structures (e.g., metrical feet)

But foot boundaries are not pronounced:

- Speaker-teacher wants to say: *ab.ra.ka.dab.ra* (underlying form).
- Speaker-teacher’s grammar produces: *[àb.ra].ka.[dáb.ra]* (surface form).
- Speaker-teacher utters: *àb.ra.ka.dáb.ra* (overt form).
- Listener-learner hears: *àb.ra.ka.dáb.ra* (overt form).
- Listener-learner hesitates:
  what is the grammatical surface form in the target languages?

* [àb].ra.ka.[dáb].ra,  [àb.ra].ka.[dáb].ra
* [àb].ra.[ka.dáb].ra,  [àb].ra.ka.[dáb.ra], etc.?
Learning hidden structures (e.g., metrical feet)

But foot boundaries are not pronounced:

- Speaker-teacher wants to say: \textit{ab.ra.ka.dab.ra} (underlying form).
- Speaker-teacher’s grammar produces: \textit{[àb.ra].ka.[dáb.ra]} (surface form).
- Speaker-teacher utters: \textit{àb.ra.ka.dáb.ra} (overt form).
- Listener-learner hears: \textit{àb.ra.ka.dáb.ra} (overt form).
- Listener-learner hesitates:
  - what is the grammatical surface form in the target languages?

\textit{[àb].ra.ka.[dáb].ra, [àb.ra].ka.[dáb].ra}
\textit{[àb].ra.[ka.dáb].ra, [àb].ra.ka.[dáb.ra], etc.?}
Learning hidden structures (e.g., metrical feet)

But foot boundaries are not pronounced:

- Speaker-teacher wants to say: \textit{ab.ra.ka.dab.ra} (underlying form).
- Speaker-teacher’s grammar produces: \textit{[àb.ra].ka.[dáb.ra]} (surface form).
- Speaker-teacher utters: \textit{àb.ra.ka.dáb.ra} (overt form).
- Listener-learner hears: \textit{àb.ra.ka.dáb.ra} (overt form).
- Listener-learner hesitates: what is the grammatical surface form in the target languages?

\textit{[àb].ra.ka.[dáb].ra, [àb.ra].ka.[dáb].ra}
\textit{[àb].ra.[ka.dáb].ra, [àb].ra.ka.[dáb.ra], etc.?}
Robust Interpretive Parsing (Tesar and Smolensky)

<table>
<thead>
<tr>
<th></th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>l 1.</td>
<td>ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>w 2.</td>
<td>[àb.ra].[ka.dáb].ra</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>l 3.</td>
<td>[àb.ra].ka.[dáb.ra]</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Teacher: FOOTREPULSION ≫ TROCHAIC ≫ NONFINAL, producing grammatical form 🕊: [àb.ra].ka.[dáb.ra].

Learner: NONFINAL ≫ TROCHAIC ≫ FOOTREPULSION, producing loser: ab.ra.[ka.dáb].ra.

Learner hears àb.ra.ka.dáb.ra. Two possible candidates. The winner must have been, the best one, [àb.ra].[ka.dáb].ra.

Compare w and l. Promote w-preferring constraints: none. Demote l-preferring constraints: FOOTREPULSION → deadlock!
Robust Interpretive Parsing (Tesar and Smolensky)

<table>
<thead>
<tr>
<th>/ab.ra.ka.dab.ra/</th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>l 1. ab.ra.[ka.dáb].ra</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>w 2. [àb.ra].[ka.dáb].ra</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3. [àb.ra].ka.[dáb.ra]</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

- **Teacher:** FOOTREPULSION \(\gg\) TROCHAIC \(\gg\) NONFINAL, producing grammatical form \(\Rightarrow\): [àb.ra].ka.[dáb.ra].

- **Learner:** NONFINAL \(\gg\) TROCHAIC \(\gg\) FOOTREPULSION, producing loser: ab.ra.[ka.dáb].ra.

- Learner hears àb.ra.ka.dáb.ra. Two possible candidates. The winner must have been, the best one, [àb.ra].[ka.dáb].ra.

- Compare w and l. Promote w-preferring constraints: none. Demote l-preferring constraints: FOOTREPULSION \(\rightarrow\) deadlock!
**Robust Interpretive Parsing** (Tesar and Smolensky)

<table>
<thead>
<tr>
<th>/ab.ra.ka.dab.ra/</th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>l 1. ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w 2. [àb.ra].[ka.dáb].ra</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>3. [àb.ra].ka.[dáb.ra]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Teacher:** FOOTREPULSION $\gg$ TROCHAIC $\gg$ NONFINAL, producing grammatical form $\Rightarrow$: [àb.ra].ka.[dáb.ra].
- **Learner:** NONFINAL $\gg$ TROCHAIC $\gg$ FOOTREPULSION, producing loser: ab.ra.[ka.dáb].ra.
- **Learner hears** àb.ra.ka.dáb.ra. Two possible candidates. 
  The winner must have been, the best one, [àb.ra].[ka.dáb].ra.
- Compare w and l. Promote w-preferring constraints: none. Demote l-preferring constraints: FOOTREPULSION $\rightarrow$ deadlock!
### Robust Interpretive Parsing

(Tesar and Smolensky)

<table>
<thead>
<tr>
<th></th>
<th>/ab.ra.ka.dab.ra/</th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>l 1.</td>
<td>ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w 2.</td>
<td>[àb.ra].[ka.dáb].ra</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>📷 3.</td>
<td>[àb.ra].ka.[dáb].ra</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Teacher**: FOOTREPULSION $\gg$ TROCHAIC $\gg$ NONFINAL, producing grammatical form $\Rightarrow$: [àb.ra].ka.[dáb].ra.

**Learner**: NONFINAL $\gg$ TROCHAIC $\gg$ FOOTREPULSION, producing loser: ab.ra.[ka.dáb].ra.

**Learner hears** àb.ra.ka.dáb.ra. Two possible candidates. The winner must have been, the best one, [àb.ra].[ka.dáb].ra.

- Compare w and l. Promote w-preferring constraints: none.
- Demote l-preferring constraints: FOOTREPULSION $\rightarrow$ deadlock!
Robust Interpretive Parsing  (Tesar and Smolensky)

<table>
<thead>
<tr>
<th>/ab.ra.ka.dab.ra/</th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>l 1. ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w 2. [àb.ra].[ka.dáb].ra</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>3. [àb.ra].ka.[dáb.ra]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Teacher: FOOTREPULSION ≫ TROCHAIC ≫ NONFINAL, producing grammatical form ‐→: [àb.ra].ka.[dáb.ra].
- Learner: NONFINAL ≫ TROCHAIC ≫ FOOTREPULSION, producing loser: ab.ra.[ka.dáb].ra.
- Learner hears àb.ra.ka.dáb.ra. Two possible candidates. The winner must have been, the best one, [àb.ra].[ka.dáb].ra.
- Compare w and l. Promote w-preferring constraints: none. Demote l-preferring constraints: FOOTREPULSION → deadlock!
Revised *Robust Interpretive Parsing* (Biró, under review)

<table>
<thead>
<tr>
<th>/ab.ra.ka.dab.ra/</th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>l 1. ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. [àb.ra].[ka.dáb].ra</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>3. [àb.ra].ka.[dáb.ra]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

- Teacher: **FOOTREPULSION** $\gg$ **TROCHAIC** $\gg$ **NONFINAL**, $\rightarrow$ [àb.ra].ka.[dáb.ra].
- Learner: **NONFINAL** $\gg$ **TROCHAIC** $\gg$ **FOOTREPULSION**, $\rightarrow$ l ab.ra.[ka.dáb].ra.
- Learner: w is either [àb.ra].[ka.dáb].ra or [àb.ra].ka.[dáb.ra].
- Calculate (weighted) average, as *winner violation profile*. Compare it to loser. Promote w-preferring constraints: **TROCHAIC**. Demote l-preferring constraints: **FOOTREPULSION** and **NONFINAL**.
- $\rightarrow$ solution: **TROCHAIC** $\gg$ **NONFINAL** $\gg$ **FOOTREPULSION**. Learner’s new grammar different from, but equivalent to teacher’s!
Revised *Robust Interpretive Parsing* (Biró, under review)

<table>
<thead>
<tr>
<th></th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPUSSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>1.</td>
<td>ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>[àb.ra].[ka.dáb].ra</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3.</td>
<td>[àb.ra].ka.[dáb.ra]</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

- Teacher: FOOTREPUSSION $\gg$ TROCHAIC $\gg$ NONFINAL, $\rightarrow$ [àb.ra].ka.[dáb.ra].
- Learner: NONFINAL $\gg$ TROCHAIC $\gg$ FOOTREPUSSION, $\rightarrow$ l ab.ra.[ka.dáb].ra.
- Learner: w is either [àb.ra].[ka.dáb].ra or [àb.ra].ka.[dáb.ra].
- Calculate (weighted) average, as *winner violation profile*. Compare it to loser. Promote w-preferring constraints: TROCHAIC. Demote l-preferring constraints: FOOTREPUSSION and NONFINAL.
- $\rightarrow$ solution: TROCHAIC $\gg$ NONFINAL $\gg$ FOOTREPUSSION. Learner’s new grammar different from, but equivalent to teacher’s!
Revised *Robust Interpretive Parsing* (Biró, under review)

<table>
<thead>
<tr>
<th></th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>[àb.ra].[ka.dáb].ra</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3</td>
<td>[àb.ra].ka.[dáb.ra]</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Teacher: FOOTREPULSION $\gg$ TROCHAIC $\gg$ NONFINAL, $\rightarrow$ l [àb.ra].ka.[dáb.ra].

Learner: NONFINAL $\gg$ TROCHAIC $\gg$ FOOTREPULSION, $\rightarrow$ l ab.ra.[ka.dáb].ra.

Learner: w is either [àb.ra].[ka.dáb].ra or [àb.ra].ka.[dáb.ra].

Calculate (weighted) average, as *winner violation profile*.

Compare it to loser. Promote w-preferring constraints: TROCHAIC. Demote l-preferring constraints: FOOTREPULSION and NONFINAL.

$\rightarrow$ solution: TROCHAIC $\gg$ NONFINAL $\gg$ FOOTREPULSION.

Learner’s new grammar different from, but equivalent to teacher’s!
Revised *Robust Interpretive Parsing* (Biró, under review)

<table>
<thead>
<tr>
<th>/ab.ra.ka.dab.ra/</th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>l 1. ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. [àb.ra].[ka.dáb].ra</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3. [àb.ra].ka.[dáb.ra]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

- Teacher: **FOOTREPULSION** ≫ **TROCHAIC** ≫ **NONFINAL**, → I [àb.ra].ka.[dáb.ra].
- Learner: **NONFINAL** ≫ **TROCHAIC** ≫ **FOOTREPULSION**, → I ab.ra.[ka.dáb].ra.
- Learner: w is either [àb.ra].[ka.dáb].ra or [àb.ra].ka.[dáb.ra].
- Calculate (weighted) average, as *winner violation profile*. Compare it to loser. Promote w-preferring constraints: **TROCHAIC**. Demote l-preferring constraints: **FOOTREPULSION** and **NONFINAL**.
- → solution: **TROCHAIC** ≫ **NONFINAL** ≫ **FOOTREPULSION**.
  Learner’s new grammar different from, but equivalent to teacher’s!
Revised *Robust Interpretive Parsing* (Biró, under review)

<table>
<thead>
<tr>
<th>/ab.ra.ka.dáb.ra/</th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>l 1. ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. [àb.ra].[ka.dáb].ra</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3. [àb.ra].ka.[dáb.ra]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

- **Teacher:** FOOTREPULSION ≫ TROCHAIC ≫ NONFINAL, → /ab.ra.ka.[dáb.ra]/.
- **Learner:** NONFINAL ≫ TROCHAIC ≫ FOOTREPULSION, → l ab.ra.[ka.dáb].ra.
- **Learner:** w is either [àb.ra].[ka.dáb].ra or [àb.ra].ka.[dáb.ra].
- **Learner:** w is either [àb.ra].[ka.dáb].ra or [àb.ra].ka.[dáb.ra].
- **Calculate (weighted) average, as winner violation profile.**
  - Compare it to loser. Promote w-preferring constraints: TROCHAIC.
  - Demote l-preferring constraints: FOOTREPULSION and NONFINAL.
- **→ solution:** TROCHAIC ≫ NONFINAL ≫ FOOTREPULSION.

*Learner’s new grammar different from, but equivalent to teacher’s!*

Tamás Biró

Being stressed by stress
Revised *Robust Interpretive Parsing* (Biró, under review)

<table>
<thead>
<tr>
<th></th>
<th>/ab.ra.ka.dab.ra/</th>
<th>NONFINAL</th>
<th>TROCHAIC</th>
<th>FOOTREPULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>ab.ra.[ka.dáb].ra</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>[àb.ra].[ka.dáb].ra</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3.</td>
<td>[àb.ra].ka.[dáb.ra]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ w_{0.5} \]

- **Teacher:** FOOTREPULSION $\gg$ TROCHAIC $\gg$ NONFINAL, $\rightarrow$ \[àb.ra].ka.[dáb.ra].
- **Learner:** NONFINAL $\gg$ TROCHAIC $\gg$ FOOTREPULSION, $\rightarrow$ l  ab.ra.[ka.dáb].ra.
- **Learner:** w is either [àb.ra].[ka.dáb].ra or [àb.ra].ka.[dáb.ra].
- **Calculate** (weighted) average, as *winner violation profile.* Compare it to loser. Promote w-preferring constraints: TROCHAIC. Demote l-preferring constraints: FOOTREPULSION and NONFINAL.
- $\rightarrow$ **solution:** TROCHAIC $\gg$ NONFINAL $\gg$ FOOTREPULSION. *Learner’s new grammar different from, but equivalent to teacher’s!*
Success rate of learning a random target grammar, as a function of parameter \(K_{\text{max}}\), for different update rules. Random initial grammar and random target grammar, with twelve constraints.
Overview

1. Optimality Theory: competence
2. Simulated Annealing for Optimality Theory: performance
3. Learning in Optimality Theory: problems
4. Conclusions
Conclusions

- **Optimality Theory**: model of linguistic competence. Grammatical form = (globally) optimal candidate.


- Learning an OT grammar poses challenges, e.g., due to hidden structures. But solutions exist!

- (Omitted from this talk: OT in other domains (e.g., syntax), learning in presence of performance errors, iterated learning.)
Conclusions

- **Optimality Theory**: model of linguistic competence. Grammatical form = (globally) optimal candidate.

- **Simulated Annealing for Optimality Theory**: model of performance. Produced form = locally optimal candidate. Predicting frequencies. More error, if SA-OT is run faster.

- Learning an OT grammar poses challenges, e.g., due to hidden structures. But solutions exists!

- (Omitted from this talk: OT in other domains (e.g., syntax), learning in presence of performance errors, iterated learning.)
Conclusions

- **Optimality Theory**: model of linguistic competence. Grammatical form = (globally) optimal candidate.

- **Simulated Annealing for Optimality Theory**: model of performance. Produced form = locally optimal candidate. Predicting frequencies. More error, if SA-OT is run faster.

- Learning an OT grammar poses challenges, e.g., due to hidden structures. But solutions exists!

- (Omitted from this talk: OT in other domains (e.g., syntax), learning in presence of performance errors, iterated learning.)
Conclusions

- **Optimality Theory**: model of linguistic competence. Grammatical form = (globally) optimal candidate.

- **Simulated Annealing for Optimality Theory**: model of performance. Produced form = locally optimal candidate. Predicting frequencies. More error, if SA-OT is run faster.

- Learning an OT grammar poses challenges, e.g., due to hidden structures. But solutions exists!

- (Omitted from this talk: OT in other domains (e.g., syntax), learning in presence of performance errors, iterated learning.)
Thank you for your attention!

Tamás Biró:

t.s.biro@uva.nl

Work supported by: