Optimality Theoretical morphology for hybrid grammars:

Implementing Burzio's Output-Output Faithfulness

Tamás Biró

ELTE Eötvös Loránd University

International Morphology Meeting 18, Budapest, May 13, 2018

"Genuine" morphology and Optimality Theory?

Why Optimality Theory at the *International Morphology Meeting*?

Question: *Is there place for analogy in Optimality Theory?*

I recommend **Luigi Burzio's Surface-to-Surface Morphology:**
"when your representations turn into constraints".

Question: *Is there place for co-morphologies in Optimality Theory?*

Sure! Let me present you an attempt!

"Genuine" morphology and Optimality Theory?

Why Optimality Theory at the *International Morphology Meeting*?

Question: *Is there place for analogy in Optimality Theory?*

I recommend **Luigi Burzio's Surface-to-Surface Morphology:**
"when your representations turn into constraints".

Question: *Is there place for co-morphologies in Optimality Theory?*

Sure! Let me present you an attempt!

"Genuine" morphology and Optimality Theory?

Why Optimality Theory at the *International Morphology Meeting*?

Question: *Is there place for analogy in Optimality Theory?*

I recommend **Luigi Burzio's Surface-to-Surface Morphology:** "when your representations turn into constraints".

Question: *Is there place for co-morphologies in Optimality Theory?*

Sure! Let me present you an attempt!

"Genuine" morphology and Optimality Theory?

Why Optimality Theory at the *International Morphology Meeting*?

Question: *Is there place for analogy in Optimality Theory?*

I recommend **Luigi Burzio's Surface-to-Surface Morphology:** "when your representations turn into constraints".

Question: *Is there place for co-morphologies in Optimality Theory?*

Sure! Let me present you an attempt!

"Genuine" morphology and Optimality Theory?

Why Optimality Theory at the *International Morphology Meeting*?

Question: *Is there place for analogy in Optimality Theory?*

I recommend **Luigi Burzio's Surface-to-Surface Morphology:** "when your representations turn into constraints".

Question: *Is there place for co-morphologies in Optimality Theory?*

Sure! Let me present you an attempt!

Overview

1. Aboh's *hybrid morphologies* and Yiddish plural formation
2. Burzio's Output-Output Faithfulness
3. Implementing OOF in hybrid morphologies
4. Summary and conclusions

Overview

Enoch O. Aboh's "hybrid grammars"

"Contrary to the tradition in linguistics that has singled out creoles as the archetype of language creation out of contact, I have argued in this book that ***each instance of acquisition involves language contact of some sort,*** *viz., contact of different idiolects which sometimes also involves different sociolects, dialects, or even languages (. . .) [C]hildren* ***learn to master multiple linguistic subsystems that are in contact and may 'cross-breed'*** *to produce new variants, which may subsequently serve as inputs for new learners. (. . .) [L]anguage learning is always imperfect: The learners' motivation is thus not to replicate the target language faithfully but to develop learning hypotheses that bring them close enough to the target to guarantee successful communication and* ***membership in the community****. Accordingly, learners do not derive identical grammars* ***from the pool****. (. . .)* ***[L]anguage change is a perpetual phenomenon contingent on learning.***" (Aboh 2015:313-4; bold are mine.)

Enoch O. Aboh's "hybrid grammars"

"Thus we must assume that change occurs at two levels: (i) the individual level, and (ii) the population level. ***Change at the individual level*** *is contingent on acquisition: each learner develops a grammar that is close enough to the target to allow communication. In synchrony, communities manage this type of variation intrinsic to acquisition by developing conventions and norms that speakers try to converge to.* ***Change at the population or community level****, however, is what diachronic studies are concerned with. It occurs when a significant number of speakers converge toward a new grammar that eventually spreads through the whole population (and may become the norm for subsequent learners)."* (Aboh 2015:314; bold are mine.)

Yiddish (Eastern) as a "hybrid grammar"

- Old High German base
- \+ Byzantine Greek, Old Romance; Slavic; modern German, English...
- \+ Semitic (Hebrew and Aramaic) component
  making up some 12-20% of the vocabulary (Kahn 2015:691).

- Semitic component constitutes a clear linguistic subsystem: eg.,
  - periphrastic verbs with a Hebrew participle:
    moykhl zayn 'to forgive' (lit. 'to be forgiving'), *mekadesh zayn* 'to sanctify'
  - orthography: Hb words spelled in Hb way (note the high level of literacy)
  - plural morphology (momentarily)
    etc.

Yiddish (Eastern) as a "hybrid grammar"

- Old High German base
- \+ Byzantine Greek, Old Romance; Slavic; modern German, English...
- \+ Semitic (Hebrew and Aramaic) component
  making up some 12-20% of the vocabulary (Kahn 2015:691).

- Semitic component constitutes a clear linguistic subsystem: eg.,
  - periphrastic verbs with a Hebrew participle:
    moykhl zayn 'to forgive' (lit. 'to be forgiving'), *mekadesh zayn* 'to sanctify'
  - orthography: Hb words spelled in Hb way (note the high level of literacy)
  - plural morphology (momentarily)
    etc.

Yiddish (Eastern) as a "hybrid grammar"

- Old High German base + Greek, Romance, Slavic...
- \+ Semitic (Hebrew and Aramaic) component, a linguistic subsystem.
- Possible origins / causes of the Semitic component:
  1. Religious and cultural register related to Judaism, Jewish culture.
  2. Identity marker (Judeo-languages as ethnolects, religiolects).
  3. High-prestige words percolating down from the most educated.
  4. Effect of (early) L2 acquisition.
- In traditional Jewish society: Hebrew acquired by most boys as early L2, starting at age 3 to 5 (mainly passive, written modality).

Yiddish (Eastern) plural formation

- Germanic plural markers: -∅ (*fish*, plur. *fish* 'fish'; with umlaut: *hant*, pl *hent* 'hand'), -ər (*kind*, pl. *kinder* 'child'; with umlaut: *land*, pl. *lender* 'land'), -(ə)n (*delegat*, pl. *delegatn* 'delegate') and -s (*lebn*, pl. *lebns* 'life').

- Some words of Slavic origin: -əs (*slup*, pl. *slupəs* 'pole, post').

- Semitic (Hebrew and Aramaic) component: -əs (*soyd*, pl. *soydəs* 'secret', *mokəm*, pl. *məkojməs* 'place'), -im (or -əm; *nign*, pl. *nigunəm* 'melody', *lamdn*, pl. *lamdonəm* 'learned man').

Yiddish (Eastern) plural formation

Linguistic features originating in etymologically different components of the hybrid grammar may recombine:

- Germanic and Hebraic co-morphologies coexist peacefully, and correlate strongly with the etymology of the vocabulary. But...
- sometimes pluralized differently from Hebrew:
  shabes ('Sabbat'), pl. *shabosəm*, not **shabosəs*, cf. Hebrew *šabbātōṯ*, *balebos* ('landlord, etc.'), pl. *balebatəm*, not **baleybos*.
- sometimes Hebraic plural attached to non-Semitic words:
  pojər, pl. *pojərəm* 'farmer' (cf. Modern German *Bauer*)
  and *doktər*, pl. *doktoyrəm* 'physician'.

Note, however, the limited scope of this kind of recombination!

Yiddish (Eastern) plural formation

- Words in the Hebraic component: *rabi*, pl. *rabonem* 'rabbi', *soyfər*, pl. *soyfrəm* 'scribe', *lamd(ə)n*, pl. *lamdonəm* 'learned man', etc.
- . . . and many more Hebrew words (*qal* active participle) follow pattern (in Yiddishized form): CoCəC, plural CoCCəm, as well as CoCCəC, plural CoCCoCəm.
- Note also that many of these words refer to professions.
- Similar phonological structure and similar semantics. Therefore, **by analogy**: *pojər*, pl. *pojərəm*, and *doktər*, pl. *doktoyrəm*.

Overview

Surface-to-Surface Morphology (Burzio 2002)

Turning representations into constraints:

Representations (lexical items) such as [A B C D].

- **Representational Entailments (RE):**

  "Mental representations of linguistic expressions constitute sets of entailments – a representation with the structure AB generating the entailments A⇒B, [and] B⇒A."

- Cf. to Hebbian learning: "cells that fire together wire together".

  "When an axon of a cell A is near enough to excite cell B or repeatedly or persistently takes part in firing it, some growth or metabolic change takes place in both cells such that A's efficiency, as one of the cells firing B, is increased"

Surface-to-Surface Morphology (Burzio 2002)

Turning representations into constraints:

- **Representational Entailments (RE):**

  "Mental representations of linguistic expressions constitute sets of entailments – a representation with the structure AB generating the entailments A⇒B, [and] B⇒A."

- Lexical item [A B C D] introduces **entailment A⇒B**, which is
  - satisfied by candidates [A B ∗∗] and [¬A ∗∗ ∗],
  - violated by candidates [A ¬B ∗∗] (include A, but not B).
- Burzio's version of OUTPUT-TO-OUTPUT FAITHFULNESS:
  # of entailments violated by the candidate, summed over all entailments by all lexical items in the entire lexicon.

Surface-to-Surface Morphology (Burzio 2002)

Example:

- Suppose lexical items currently in the lexicon:
  [A B C D], [A X Y D], [Z X U W].
- Entailments introduced: A⇒B, A⇒C, A⇒D, B⇒A, B⇒C, B⇒D, . . . , A⇒X, A⇒Y, A⇒D (again!), X ⇒Y, X⇒D . . . and many more.
- Let us evaluate candidate [A B U W]
  for constraint OUTPUT-OUTPUT FAITHFULNESS:

stars assigned to candidate = # of entailments violated,
such as A⇒C, A⇒D (twice!), B⇒C, B⇒D, . . .

Overview

Burzio's OO-Faithfulness for Yiddish

Thus far:

- Burzio's entailments (gradient attraction):
  a formalism to handle analogy output-output faithfulness.
- Yiiddish *pojər*, pl. *pojərəm* and *doktər*, pl. *doktoyrəm* as examples.

Now combining the two:

- **Word-level (micro-level) entailment:**
  - "If a word has singular form soyfər, then its plural ends in -əm."
  - "If a word means 'rabbi', then its plural ends in -əm."
- **Lexicon-level (macro-level) entailment:**
  - "If a word has singular form CoC(C)əC, then its plural ends in -əm."
  - "If a word refers to a profession, then its plural ends in -əm."
- The sum of these entailments acts as a **violable OT constraint**.

Burzio's OO-Faithfulness for Yiddish

Thus far:

- Burzio's entailments (gradient attraction):
  a formalism to handle analogy output-output faithfulness.
- Yiiddish *pojər*, pl. *pojərəm* and *doktər*, pl. *doktoyrəm* as examples.

Now combining the two:

- **Word-level (micro-level) entailment:**
  - "If a word has singular form soyfər, then its plural ends in -əm."
  - "If a word means 'rabbi', then its plural ends in -əm."
- **Lexicon-level (macro-level) entailment:**
  - "If a word has singular form CoC(C)əC, then its plural ends in -əm."
  - "If a word refers to a profession, then its plural ends in -əm."
- The sum of these entailments acts as a **violable OT constraint**.

Multi-agent simulation

- An agent equipped with:
  - + a grammar: Burzio's OOF;
  - + a lexicon:
    4 type A plural, 3 type B plural,
    1 type A plural analogous to type B words;
  - + a production algorithm: OT with exhaustive search;
  - + a learning algorithm: batch learning of lexical items;
  - + a level of education: $0 \leq \ell \leq 1$
    (either constant or following some distribution).

- $N = 250$ agents forming a generation

- Iterative learning with 10 consecutive generations

Lexical item

String grammar:

$$\text{form} \in \Phi^{\phi} = \{0,1,2,3,4\}^5, \quad \text{meaning} \in \Sigma^{\sigma} = \{a,b,c,d,\ldots\}^{\sigma},$$
$$\text{number} \in \{\textsc{sg}, \textsc{pl}\}, \quad \text{co-phonology} \in \{\textsc{a}, \textsc{b}\}.$$

E.g.: $\lambda_1 = \begin{bmatrix} \text{form} & [1230] \\ \text{meaning} & \text{'abcd'} \\ \text{number} & \textsc{sg} \\ \text{co-phonology} & \textsc{a} \end{bmatrix}$, $\lambda_2 = \begin{bmatrix} \text{form} & [2101] \\ \text{meaning} & \text{'abcd'} \\ \text{number} & \textsc{pl} \\ \text{co-phonology} & \textsc{a} \end{bmatrix}$

- Gen maps a singular form to all plural forms with same meaning and co-phonology, and any form $\in \Phi^{\phi}$.

Analogy constraint

A pair of lexical items:

$$\lambda_{\text{sg}} = \begin{bmatrix} \text{form} & [1230] \\ \text{meaning} & \text{'abcd'} \\ \text{number} & \text{SG} \\ \text{co-phonology} & \text{A} \end{bmatrix}, \quad \lambda_{\text{pl}} = \begin{bmatrix} \text{form} & [1031] \\ \text{meaning} & \text{'abcd'} \\ \text{number} & \text{PL} \\ \text{co-phonology} & \text{A} \end{bmatrix}$$

- Plural formation pattern: $\lambda_{\text{pl}}|\mathit{form} - \lambda_{\text{sg}}|\mathit{form} = [0, -2, 0, 1]$.

- Burzio's entailments:

if meaning[3]='c', then plural_formation_pattern [2]$= -2$,
if $\lambda_{\text{sg}}|$form[4]=0, then plural_formation_pattern [3]$= 0$.

Analogy constraint

Analogy constraint applied to candidate (input l_{sg}, output l_{pl}):

- Calculate plural formation pattern $l_{\text{pl}}|form - l_{\text{sg}}|form$.
- Collect all lexical items $(\lambda_{i,\text{pl}}, \lambda_{i,\text{sg}})$ with both singular and plural forms stored in the lexicon.
- Test each entailment of each $(\lambda_{i,\text{pl}}, \lambda_{i,\text{sg}})$ applied to $(l_{\text{sg}}, l_{\text{pl}})$.
- C_{analogy}(input l_{sg}, output l_{pl}) = number of entailment violations.

(for given l_{sg} and for all corresponding l_{pl})

Analogy constraint **with co-phonologies**

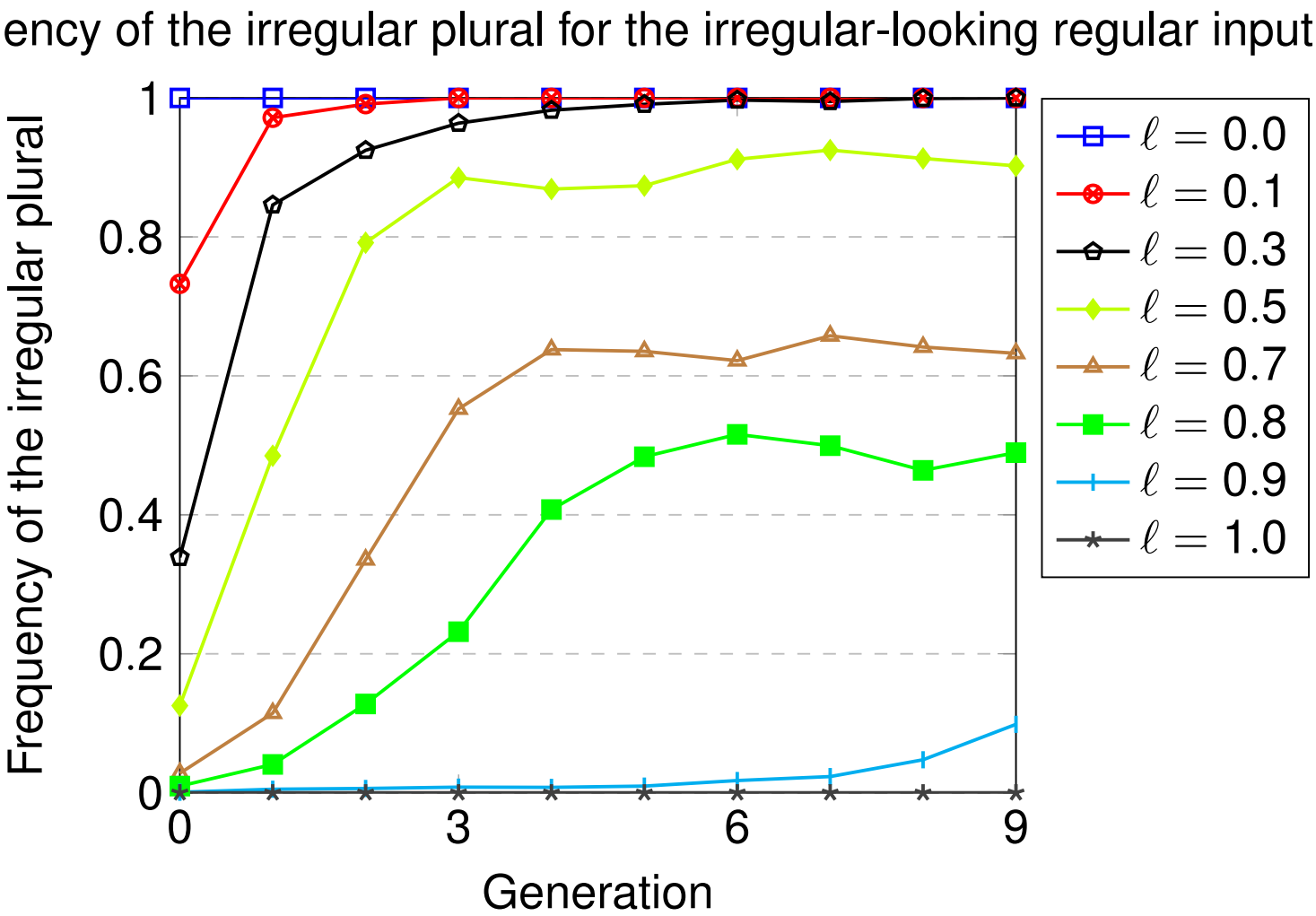
- $C_{\text{analogy}}\Big($ $\begin{bmatrix} \text{form} & [2120] \\ \text{meaning} & \text{'cdbb'} \\ \text{number} & \text{SG} \\ \text{co-phonology} & \text{A} \end{bmatrix}$, $\begin{bmatrix} \text{form} & [1031] \\ \text{meaning} & \text{'cdbb'} \\ \text{number} & \text{PL} \\ \text{co-phonology} & \text{A} \end{bmatrix}$ $\Big) = ?$

- Apply all entailments by all lexical item pairs, such as

$\begin{bmatrix} \text{form} & [1230] \\ \text{meaning} & \text{'abcd'} \\ \text{number} & \text{SG} \\ \text{co-phonology} & \text{A} \end{bmatrix}$, $\begin{bmatrix} \text{form} & [1231] \\ \text{meaning} & \text{'abcd'} \\ \text{number} & \text{PL} \\ \text{co-phonology} & \text{A} \end{bmatrix}$ pl. form. pat.: $[0, 0, 0, 1]$.

- if either they belong to the same co-phonology,
  or random number $r \in [0 \ldots 1]$ generated $>$ *level of education* ℓ.

Results


Frequency of the irregular plural for the irregular-looking regular input
Frequency of the irregular plural
Generation
0
0.2
0.4
0.6
0.8
1
0
3
6
9
$\ell = 0.0$
$\ell = 0.1$
$\ell = 0.3$
$\ell = 0.5$
$\ell = 0.7$
$\ell = 0.8$
$\ell = 0.9$
$\ell = 1.0$

Observations

If those less educated in the society had "hybridized" the grammar, the most educated members of the next generation would also acquire it.

- Plots probably reach state of equilibrium after a few generations. Oscillations around the equilibrium: due to small learning sample?
- Higher *level of education* ℓ:
  fewer re-analysis of the irregular-looking regular form.
- With a probability $< 0.1\%$, plural of the irregular words are also formed according to the regular pattern.
- Yet, no third plural formation class (e.g., a mixture of the regular and irregular patterns) emerge. Due to particular details of this toy grammar, or is it necessity of the framework?

Overview

What have we achieved?

Formalizing and computationally modeling analogical phenomena and paradigms in historical morphology is certainly feasible.

- Optimality Theory, analogy, and language change meet again:
- Burzio's formalism implemented in a multi-agent computer simulations with iterative learning.

Additional possibilities:

- More OT constraints, stochastic production, etc.
- Realistic lexicon. Non-trivial social network.
- Non-toy grammar.

But the complexity of the model would be intractable! Reproducing historical sociolinguistics *in silico* is by far not self-evident.

What have we achieved?

Formalizing and computationally modeling analogical phenomena and paradigms in historical morphology is certainly feasible.

- Optimality Theory, analogy, and language change meet again:
- Burzio's formalism implemented in a multi-agent computer simulations with iterative learning.

Additional possibilities:

- More OT constraints, stochastic production, etc.
- Realistic lexicon. Non-trivial social network.
- Non-toy grammar.

But the complexity of the model would be intractable! Reproducing historical sociolinguistics *in silico* is by far not self-evident.

Thank you for your attention!

Tamás Biró:
tamas[dot]biro[at]btk[dot]elte[dot]hu
http://birot.web.elte.hu/, http://birot.hu/OTKit/

This research was supported by a *Marie Curie FP7 Integration Grant* (no. PCIG14-GA-2013-631599, "MeMoLI"), *7th EU Framework Programme*.