Stem spaces and regularity in verbal inflection

Olivier Bonami
Université Paris-Sorbonne / LLF
olivier.bonami@paris-sorbonne.fr

Fabio Montermini
CLLE-ERSS, CNRS & Université de Toulouse
fabio.montermini@univ-tlse2.fr

Understanding the Architecture of the Mental Lexicon
Pisa, 24-26 November 2011
Introduction

• An overview of some recent research on inflection in Romance languages
• A reassessment of the developments morphology has known in the last 20 years
• Morphomes, stem spaces as means of structuring paradigms
• Integrating cognitively plausible models with formal linguistic treatments
Introduction

- Structuring paradigms
- Implications in paradigms
- Stem spaces and entropy
- Stem spaces and graphs
- Conclusions
Structuring paradigms

- Structuralist (and many generative) models of morphology considered formal unity as a principle: (almost) all formal variation is reduced by posing (morpho)phonological rules modifying a basic form.

- However, we the linguists know the very reason why this variation is observed, since it is the reflect of historical changes the language has gone through.

- A consequence of this was the projection in syncrony of historical changes. However, to our knowledge, no serious discussion of the cognitive premises and consequences of this has never been put forth.
• Such notions as morphomes (Aronoff 1994, Maiden 2005, 2009) are a way of structuring paradigms according to purely morphological principles; i.e. independent of external (phonological, semantic) motivations.
### Structuring paradigms

- Morphomic distribution of verbal forms in Nepali

**Long negative present of BIRSANU (‘forget’)***

<table>
<thead>
<tr>
<th></th>
<th>M.SG</th>
<th>F.SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>birsādinā</td>
<td>birsādinā</td>
<td>birsādajnaū</td>
</tr>
<tr>
<td>2.LOW</td>
<td>birsādajnas</td>
<td>birsādinas</td>
<td>birsādajnau</td>
</tr>
<tr>
<td>2.MID</td>
<td>birsādajnau</td>
<td>birsādinau</td>
<td>birsādajnau</td>
</tr>
<tr>
<td>3.LOW</td>
<td>birsādajna</td>
<td>birsādina</td>
<td>birsādajnan</td>
</tr>
<tr>
<td>3.MID</td>
<td>birsādajnan</td>
<td>birsādinan</td>
<td>birsādajnan</td>
</tr>
<tr>
<td>HIGH</td>
<td></td>
<td></td>
<td>birsanuhūdajna</td>
</tr>
</tbody>
</table>

3 grades of honorifics for persons 2-3; high grade honorific neutralizes person-gender-number (Bonami & Boyé 2010)
Structuring paradigms

• Morphomic distribution of verbal forms in Nepali

Long negative present of BIRSANU (‘forget’)

<table>
<thead>
<tr>
<th></th>
<th>M.SG</th>
<th>F.SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>birsãd-i-na-ã</td>
<td>birsãd-i-na-ã</td>
<td>birsãd-aj-na-aũ</td>
</tr>
<tr>
<td>2.LOW</td>
<td>birsãd-aj-na-s</td>
<td>birsãd-i-na-s</td>
<td>birsãd-aj-na-au</td>
</tr>
<tr>
<td>2.MID</td>
<td>birsãd-aj-na-au</td>
<td>birsãd-i-na-au</td>
<td>birsãd-aj-na-au</td>
</tr>
<tr>
<td>3.LOW</td>
<td>birsãd-aj-na</td>
<td>birsãd-i-na</td>
<td>birsãd-aj-na-n</td>
</tr>
<tr>
<td>3.MID</td>
<td>birsãd-aj-na-n</td>
<td>birsãd-i-na-n</td>
<td>birsãd-aj-na-n</td>
</tr>
</tbody>
</table>
Structuring paradigms

• All the inflected forms are considered as equal: what counts in a paradigm are global relations between all the forms, not local relations between a form and the hypothetical basic form of a lexeme it derives from.

• Such structures as morphomes have proven to be useful to account for:
  - diachronic change (Maiden 2005, 2009)
  - psycho-computational models (Pirrelli 2000; Bonami et al. 2008)
Structuring paradigms

• The most neutral way of addressing the question of paradigmatic relations is the “Paradigm Cell Filling Problem” (Ackerman et al. 2009; Malouf & Ackerman 2010):

  “Given exposure to an inflected wordform of a novel lexeme, what licenses reliable inferences about the other wordforms in its inflectional family?”
Knowledge of implicative patterns relating cells in a paradigm is relevant.

- This knowledge is best characterized in information-theoretic terms

- The reliability of implicative patterns relating paradigm cell A to paradigm cell B is measured by the conditional entropy of cell B knowing A.

(Ackerman et al. 2009; Malouf & Ackerman 2010)
• In theory, a solution to the PCFP should consist in determining how any word-form in a paradigm can be deduced from any other form, i.e. in describing the relations between all the forms of a lexeme (e.g. for a French verb $51 \times 50 = 2,550$ relations).

• Probably, such a distributed view is the most plausible model of morphological competence from a cognitive perspective.

• However, what linguists should be interested in is describing and formalizing languages and the differences between them.

• The model we propose aims at accounting for higher level properties of such a system.
A simple inventory of all form-to-form relations in a paradigm is a huge task.

Moreover, it is probable that it hides some generalizations.

In fact, morphologists have elaborated several theoretical objects aiming at giving structure to paradigms, intuitively seen as networks of relations between equal forms; morphomes are one of these objects.
In theory, all the relations between forms in a paradigm should be considered as equal. In practice, apart from cases of extremely simple inflectional systems, languages do not function like this.

Several phenomena (allomorphy, syncretism, inflectional class repartition...) influence the degree of predictability of each relation.
# Implications in paradigms

- Measuring the existence of structure in implicative relations.
- When paradigms are small, the structure is transparent

## French adjectives

<table>
<thead>
<tr>
<th></th>
<th>M Sg</th>
<th>M Pl</th>
<th>F Sg</th>
<th>F Pl</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>‘hard’</td>
<td>dyʁ</td>
<td>dyʁ</td>
<td>dyʁ</td>
<td>dyʁ</td>
<td>77%</td>
</tr>
<tr>
<td>‘legal’</td>
<td>legal</td>
<td>lego</td>
<td>legal</td>
<td>legal</td>
<td>3%</td>
</tr>
<tr>
<td>strong’</td>
<td>ʃɔʁ</td>
<td>ʃɔʁ</td>
<td>ʃɔʁt</td>
<td>ʃɔʁt</td>
<td>20%</td>
</tr>
</tbody>
</table>

Approximative figures from the lexique.org database
• Entropy measures of predictability between the forms of French adjectives

Dataset: 4,972 fully inflected adjectives from the lexique.org database. Phonetic transcriptions corrected by hand by the authors.
Implications in paradigms

- With large paradigms, measures become uninterpretable
Stem spaces and entropy

• Stem spaces for French (Bonami & Boyé 2003, 2007, Boyé 2009), Italian (Pirrelli & Battista 2000, Montermini & Boyé f.c.), Spanish (Boyé & Cabredo-Hofherr 2006), Catalan (Guerrero f.c.) were designed by hand as a way of capturing qualitatively implicative structure in paradigms.

• It turns out that there is a strong correlation between the hand-designed stem spaces and quantitative properties of implicational relations.

• In the case of French, slots in the stem space correspond almost exactly to the collections of paradigm cells related by fully predictable implicational relations.
### Stem spaces and entropy

<table>
<thead>
<tr>
<th>Stem Spaces</th>
<th>entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1</td>
<td></td>
</tr>
<tr>
<td>s2</td>
<td></td>
</tr>
<tr>
<td>s3</td>
<td></td>
</tr>
<tr>
<td>s4</td>
<td></td>
</tr>
<tr>
<td>s5</td>
<td></td>
</tr>
<tr>
<td>s6</td>
<td></td>
</tr>
<tr>
<td>s7</td>
<td></td>
</tr>
<tr>
<td>s8</td>
<td></td>
</tr>
<tr>
<td>s9</td>
<td></td>
</tr>
<tr>
<td>s10</td>
<td></td>
</tr>
<tr>
<td>s11</td>
<td></td>
</tr>
</tbody>
</table>

Legend:
- **inf**: Infinitive
- **pst.pcp**: Past Perfect Continuous
- **prs**: Present
- **prs.pcp**: Present Perfect Continuous
- **sbj**: Subject
- **sbjv**: Subjective
- **imp**: Imperfect
- **imp.2**: Imperfect 2
- **imp.3**: Imperfect 3
- **imp.4**: Imperfect 4
- **imp.5**: Imperfect 5
- **fut**: Future
- **fut.1**: Future 1
- **fut.2**: Future 2
- **fut.3**: Future 3
- **fut.4**: Future 4
- **fut.5**: Future 5
- **fut.6**: Future 6
- **cond**: Conditional
- **cond.1**: Conditional 1
- **cond.2**: Conditional 2
- **cond.3**: Conditional 3
- **cond.4**: Conditional 4
- **cond.5**: Conditional 5
- **cond.6**: Conditional 6
- **pst**: Past
- **pst.1**: Past 1
- **pst.2**: Past 2
- **pst.3**: Past 3
- **pst.4**: Past 4
- **pst.5**: Past 5
- **pst.6**: Past 6
- **pst.sbj**: Past Subject
- **pst.sbj.1**: Past Subject 1
- **pst.sbj.2**: Past Subject 2
- **pst.sbj.3**: Past Subject 3
- **pst.sbj.4**: Past Subject 4
- **pst.sbj.5**: Past Subject 5
- **pst.sbj.6**: Past Subject 6
- **prs.pcp**: Present Perfect Continuous
- **prs.pcp**: Present Perfect Continuous
Stem spaces and entropy

- Two types of exceptions to this observation
  - local uncertainties introduced by a single ambiguous pair, giving rise to very low entropy (<.01)

\[
\text{PRS 1 SG} \ [\varepsilon] \rightarrow \text{PRS 2 SG} \ [\varepsilon] \ ('hate')
\]

\[
\text{PRS 1 SG} \ [\varepsilon] \rightarrow \text{PRS 2 SG} \ [a] \ ('have')
\]

- phonological opacity

\[
/kadʁ+jɔ̃/ \quad \text{cadrions 'tally'} \rightarrow \text{IPF 1 PL} \ [kadʁijɔ̃]
\]

\[
/kadʁij+jɔ̃/ \quad \text{quadrillons 'cover'}
\]

\[
\text{IPF 1 PL} \ [kadʁijɔ̃] \rightarrow \text{IPF 1 SG} \ [kadʁijɛ]
\]

\[
\text{IPF 1 PL} \ [kadʁijɔ̃] \rightarrow \text{IPF 1 SG} \ [kadʁijɛ]
\]

\[
\text{H (IPF 1 PL} \rightarrow \text{IPF 1 SG)} \approx .44
\]

\[
\text{H (PRS 1 SG} \rightarrow \text{PRS 2 SG)} < .003
\]

19
The notion of a stem space is based on the observation that, in a paradigm, some relations are more predictive than others.

There are clusters of forms that are in systematic covariation, i.e. they systematically display the maximum degree of interpredictability.

Conversely, the relations between these clusters may or may not display a high degree of interpredictability.

Paradigms are considered as organised into stem spaces, a pattern defining the (morphomic) distribution of stems among word-forms.
### Stem spaces and graphs

<table>
<thead>
<tr>
<th></th>
<th>1SG</th>
<th>2SG</th>
<th>3SG</th>
<th>1PL</th>
<th>2PL</th>
<th>3PL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PRES IND</strong></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>IMPF IND</strong></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>PRES SBJ</strong></td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td><strong>FUT IND</strong></td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td><strong>PRES CND</strong></td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td><strong>PAST IND</strong></td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td><strong>IMPF SBJ</strong></td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td><strong>IMPER</strong></td>
<td>5</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### INFIN

<table>
<thead>
<tr>
<th>INFIN</th>
<th>PRES PART</th>
<th>PAST PART</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M SG</td>
<td>F SG</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>12</td>
</tr>
</tbody>
</table>
Stem spaces and graphs

- In theory, 56 relations are relevant for filling a 8-stem space:
Stem spaces and graphs

- French verb forms

<table>
<thead>
<tr>
<th></th>
<th>I PL Pres Ind (S1)</th>
<th>3 PL Pres Ind (S2)</th>
<th>I Sg Pres Ind (S3)</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>lavõ</td>
<td>lav</td>
<td>lav</td>
<td>‘wash’</td>
<td></td>
</tr>
<tr>
<td>muʁɔ̃</td>
<td>mœʁ</td>
<td>mœʁ</td>
<td>‘die’</td>
<td></td>
</tr>
<tr>
<td>savõ</td>
<td>sav</td>
<td>sɛ</td>
<td>‘know’</td>
<td></td>
</tr>
<tr>
<td>byvõ</td>
<td>bwav</td>
<td>bwa</td>
<td>‘drink’</td>
<td></td>
</tr>
</tbody>
</table>
## Stem spaces and graphs

- French verb forms

<table>
<thead>
<tr>
<th>I PL PRES IND (S1)</th>
<th>3 PL PRES IND (S2)</th>
<th>1 SG PRES IND (S3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>
# Stem spaces and graphs

- French verb forms

<table>
<thead>
<tr>
<th>I Pl Pres Ind (S1)</th>
<th>3 Pl Pres Ind (S2)</th>
<th>1 Sg Pres Ind (S3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>A</td>
</tr>
</tbody>
</table>

Not attested
Stem spaces and graphs

- Knowing S1 in addition to S2 does not improve the chances of predicting S3
**Stem spaces and graphs**

- In fact, most of the relations display a low degree of predictability

<table>
<thead>
<tr>
<th></th>
<th>s1</th>
<th>s2</th>
<th>s3</th>
<th>s4</th>
<th>s5</th>
<th>s6</th>
<th>s7</th>
<th>s8</th>
<th>s9</th>
<th>s10</th>
<th>s11</th>
<th>s12</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1</td>
<td>0.23</td>
<td>0.47</td>
<td>0.00</td>
<td>0.47</td>
<td>0.01</td>
<td>0.23</td>
<td>0.68</td>
<td>0.39</td>
<td>0.59</td>
<td>0.40</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>s2</td>
<td>0.09</td>
<td>0.29</td>
<td>0.09</td>
<td>0.29</td>
<td>0.10</td>
<td>0.01</td>
<td>0.45</td>
<td>0.43</td>
<td>0.43</td>
<td>0.43</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>s3</td>
<td>0.35</td>
<td>0.50</td>
<td>0.33</td>
<td>0.00</td>
<td>0.33</td>
<td>0.51</td>
<td>0.67</td>
<td>0.40</td>
<td>0.71</td>
<td>0.37</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>s4</td>
<td>0.00</td>
<td>0.23</td>
<td>0.48</td>
<td>0.48</td>
<td>0.01</td>
<td>0.23</td>
<td>0.28</td>
<td>0.39</td>
<td>0.59</td>
<td>0.40</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>s5</td>
<td>0.35</td>
<td>0.49</td>
<td>0.00</td>
<td>0.34</td>
<td>0.35</td>
<td>0.52</td>
<td>0.69</td>
<td>0.40</td>
<td>0.71</td>
<td>0.37</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>s6</td>
<td>0.00</td>
<td>0.20</td>
<td>0.47</td>
<td>0.00</td>
<td>0.47</td>
<td>0.23</td>
<td>0.68</td>
<td>0.39</td>
<td>0.59</td>
<td>0.39</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>s7</td>
<td>0.10</td>
<td>0.02</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
<td>0.11</td>
<td>0.29</td>
<td>0.42</td>
<td>0.56</td>
<td>0.42</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>s8</td>
<td>0.42</td>
<td>0.50</td>
<td>0.68</td>
<td>0.44</td>
<td>0.68</td>
<td>0.44</td>
<td>0.49</td>
<td>0.76</td>
<td>0.86</td>
<td>0.72</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>s9</td>
<td>0.13</td>
<td>0.30</td>
<td>0.23</td>
<td>0.13</td>
<td>0.23</td>
<td>0.13</td>
<td>0.30</td>
<td>0.76</td>
<td>0.29</td>
<td>0.08</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>s10</td>
<td>0.31</td>
<td>0.29</td>
<td>0.11</td>
<td>0.31</td>
<td>0.11</td>
<td>0.31</td>
<td>0.29</td>
<td>0.72</td>
<td>0.28</td>
<td>0.30</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>s11</td>
<td>0.18</td>
<td>0.36</td>
<td>0.29</td>
<td>0.18</td>
<td>0.29</td>
<td>0.18</td>
<td>0.36</td>
<td>0.83</td>
<td>0.13</td>
<td>0.40</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>s12</td>
<td>0.17</td>
<td>0.36</td>
<td>0.26</td>
<td>0.17</td>
<td>0.26</td>
<td>0.17</td>
<td>0.35</td>
<td>0.80</td>
<td>0.11</td>
<td>0.36</td>
<td>0.07</td>
<td></td>
</tr>
</tbody>
</table>
Stem spaces and graphs

- In fact, most of the relations display a low degree of predictability
In fact, most of the relations display a low degree of predictability.
Stem spaces and graphs

• Taking into account the relations allows drawing dependency graphs of paradigmatic relations

Graph of relations of LAVER (‘wash’)
Stem spaces and graphs

• Taking into account the relations allows drawing dependency graphs of paradigmatic relations

Graph of relations of BOIRE (‘drink’)
Stem spaces and graphs

- Taking into account the relations allows drawing dependency graphs of paradigmatic relations

Graph of relations of TAPIR (‘cover’)

diagram
Stem spaces and graphs

• Taking into account the relations allows drawing dependency graphs of paradigmatic relations
Stem spaces and graphs

- Full graph of relations for regular verbs in French

from Boyé (2010)
Conclusion

• The autonomy of morphology (of purely morphological relations) poses the problem of reducing the complexity of inflectional systems by relying only on morphological relations.

• Although fully connectionist (surface-to-surface, word-based) models may be cognitively plausible, we need more in order to account for the variety of languages and of phenomena (recurrent patterns of allomorphy / syncretism; inflectional classes and subclasses...).

• Stem spaces and graphs (among others) are means of reducing the complexity of computation needed in inflectional morphology while maintaining the claim that lexemes are complex objects.

• Formally, they are best represented as networks of connected forms rather than unique phonological objects that undergo (morpho)phonological operations.
References


Montermini F., Boyé G. f.c.: “Stem relations and inflection class assignment in Italian”. Word Structure.