Analytic and Synthetic Verb Forms in Irish — An Agreement-Based Implementation in LFG

Sebastian Sulger

KONVENS2010, Saarbrücken
Irish Verbs

... may occur in two different forms:

- analytic:
  - person/number features not expressed by verb
  - independent subject pronoun needed

```
shalaigh mé mo lámha.  *shalaigh mo lámha.
‘I messed up my hands.’
```
Irish Verbs

... may occur in two different forms:

- **analytic:**
  - person/number features not expressed by verb
  - independent subject pronoun needed

  \[\text{shalaigh m´ e mo l´ amha.} \quad \text{*shalaigh mo l´ amha.}\]
  \[\text{dirty. Pret I my hand. Pl} \quad \text{dirty. Pret my hand. Pl}\]
  ‘I messed up my hands.’

- **synthetic:**
  - person/number features expressed by suffixes on verb
  - independent subject pronoun not needed — even disallowed

  \[\text{shalaíomar ár l´ amha.} \quad \text{*shalaíomar muid ár l´ amha.}\]
  \[\text{dirty. Pret. 1P. Pl our hand. Pl} \quad \text{dirty. Pret. 1P. Pl we our hand. Pl}\]
  ‘We messed up our hands.’
Irish Verbs

To complicate matters:

- distribution of analytic vs. synthetic forms varies across paradigms:

Verb *tuig* ‘understand’

<table>
<thead>
<tr>
<th>Features</th>
<th>Present</th>
<th>Imperfect</th>
<th>Future</th>
</tr>
</thead>
<tbody>
<tr>
<td>1P.Sg</td>
<td>tuigim</td>
<td>thuiginn</td>
<td>tuigfidh</td>
</tr>
<tr>
<td>2P.Sg</td>
<td>tuigeann</td>
<td>thuigteá</td>
<td>tuigfidh</td>
</tr>
<tr>
<td>3P.Sg</td>
<td>tuigeann</td>
<td>thuigeadh</td>
<td>tuigfidh</td>
</tr>
<tr>
<td>1P.Pl</td>
<td>tuigimid</td>
<td>thuigimis</td>
<td>tuigfimid</td>
</tr>
<tr>
<td>2P.Pl</td>
<td>tuigeann</td>
<td>thuigeadh</td>
<td>tuigfidh</td>
</tr>
<tr>
<td>3P.Pl</td>
<td>tuigeann</td>
<td>thuigidis</td>
<td>tuigfidh</td>
</tr>
</tbody>
</table>
Irish Verbs

To complicate matters:

- ‘analytic form + pronoun’ is blocked if synthetic form exists which realizes the same features ("Morphological Blocking"):

  Tuigim an fhadhb.  *Tuigeann mé an fhadhb.
  understand.Pres.1P.Sg the problem understand.Pres l the problem

  ‘I understand the problem.’

→ Question: How can this interaction between verb forms be captured?
Implementation challenges

An implementation of Irish verbs is thus faced with two tasks:

1. block redundant use of pronoun together with synthetic form
2. block ‘analytic form + pronoun’ if synthetic form is available

To my knowledge, there is no efficient, working implementation for these two tasks.
An LFG implementation

I propose an approach within Lexical-Functional Grammar — key assumptions are:

- For both forms – synthetic and analytic – I assume that the pronoun is *pro*, realizing person/number features.
- *pro* agrees with the verb in person/number.
- Provided a detailed morphological analysis, simple agreement equations can be used to rule out ungrammatical strings.
An LFG implementation

My implementation is based on two components:

1. a detailed morphological analyzer written in XFST dealing with verbal morphology
   - lists both analytic and synthetic verb forms
   - analyses include morphosyntactic features such as person, number

2. a computational LFG grammar ruling out ungrammatical sentences using two short agreement templates
   - independent pronouns blocked when occurring with synthetic forms
   - analytic forms are blocked where inappropriate

The implementation is situated within the context of the ParGram project, devoted to the development of wide coverage LFG grammars.
Sneak Peek — Analysis output by the grammar:

*Tuigeann mé an fhadhb.
understand.Pres l the problem
‘I understand the problem.’

→ Analysis fails: no agreement in (^ PERS) of subject
An LFG implementation

Sneak Peek — Analysis output by the grammar:

Tuigeann sé an fhadhb.
understand.Pres he the problem
‘He understands the problem.’

→ Analysis valid: agreement in (\(^ PERS \)) and (\(^ NUM \))
For the details of my solution — come see my poster!