

Specifying Grammatical Points

Micheál Mac Lochlainn and Gearóid Ó Néill

Ollscoil na Gaillimhe, Gaoth Dobhair, Ollscoil Luimnigh, Luimneach
muscrae@yahoo.ie, gearoid.oneill@ul.ie

Abstract

The system provides a language for the grammarian to specify grammatical points of Irish for use in computer assisted language systems. There is a basic knowledge base, which can be used in the specification of other grammatical points. Included are predefined entities, functions and the provision for defining the rules. The rules, together with the other information, are used to form the basis of a quasi-natural language explanation. Information on pronunciation as well as a facility to record phrases is provided. The dictionary “An Foclóir Beag” forms part of the knowledge base. The student can use the rules for explanation, suggestions for the correct form and verification. The language is interpreted.

1. Introduction

Grammar books exist for many languages, for example, [1, 2, 3]. They give examples and ‘rules’ about inflection and so on. They may also give some pointers on sentence structure but not in a formal style. Here an attempt is being made to generate something similar but with the advantage that rules can be applied by the system. It is a development from earlier work[4].

The objective is two-fold. One to provide a mechanism for the grammarian to specify the grammatical points in a language and a style close to that found in grammar books and two, to make the grammar rules more accessible to the language student

2. Some Grammatical Points

Irish uses inflection to some extent and also has changes to the beginning of a word.[3]. The initial changes include lenition and eclipse. Nouns and adjectives are inflected and agree in number, gender and case, when used attributively. The verb is inflected for person, number, tense and mood.

3. Language Constructs

The language being suggested here includes a basic knowledge base, augmented by the dictionary “An Foclóir Beag”[5] and, in effect, a facility for defining domains as well as the rule specification mechanism.

3.1 Predefined Functions

Included are some predefined functions, for example,

```
first_letter
second_letter
length
```

which work on a “word” (i.e. a string).

3.2 Set Enumeration

The structure in the language for set enumeration is shown below.

```
set set_name {elements}
member element_name
subset of set_names
elements      element
elements      element, elements
set_names     set_name
set_names     set_name, set_names
element_name  name
element       name | symbol
name          letter(letter|digit|_|-)*
symbol        apc+
```

where apc is any printable character other than space.

3.3 Rules

```
rule
rule title:grammar_elements;action.
grammar_elements  grammar_pt
grammar_pt        pos:form features.
grammar_pt        pos:form features.
                  grammar_pt
features          feature
features          feature, features
```

The features used are a subset of the particular part of speech under consideration. Only the features pertinent to the action being specified are necessary and similarly for the *pos* and the *form*.

```
condition        action
```

where action is the implementation of some grammatical point, for example, lenition.

4. Knowledge Base

The knowledge base includes the parts of speech, their properties and common grammatical functions, such as eclipse. Some

common grammatical rules are pre-defined. Along with a rule is generated a quasi-natural language explanation. The grammarian can modify the explanation.

5. Sample Entries

The alphabet is based on the Latin alphabet, although some of the letters are rarely used. The individual items which make up the alphabet are known as letters. The pronunciation of the letters is included.

set	Alphabet
	{a,,á, b, c, d, e, é, f, g, h, i, í, j, k, l, m, n, o, ó, p, q, r, s, t, u, ú, v, w, x, y, z}
member	Letter
note	j k q v w x y z rarely used

set	Consonants
	{b, c, d, f, g, h, j, k, l, m, n, p, q, r, s, t, v, w, x, y, z}
subset of	Alphabet
member	Consonant

5.1 Sample definitions

Nouns and adjectives may be inflected for case, number and gender.

set:	Case
member	{nominative, vocative, genitive, dative}
Note	accusative same form as nominative
Note	dative rare
Note	vocative same ending as either genitive or nominative
Variation	common := nominative accusative dative
name	Number
set	[singular, plural]
name	Gender
set	{feminine masculine}

The rules can be used in a variety of ways. They can be accessed as text, either by name or by searching on a word. They can be used to return the appropriate form of a word or words.

5.2 Queries

The queries include requesting all the forms of groups of words, the form for a particular condition or to verify a certain form

The commands are find all forms, find forms for specified condition or verify.

Sample:

Find all forms

an bean noun

The system responds with

an bhean	nominative singular feminine lenite noun after the article
na mná	genitive singular feminine
na mná	nominative plural feminine
na mban	genitive plural feminine eclipse noun after the article

6 Speech –Synthesized and Recorded

Some work has done on speech synthesis for Irish[7] and the text to phoneme system has been rewritten[6]. The speech synthesis is being further developed. With the speech synthesis, phrases, such as the above, which are not already recorded, can be synthesized.

A recording can be attached to the rules. The alphabet is recorded, for example. Along with recording sets and grammar samples, the phonetic representation of a word is also available.

A computerized version[8] of the dictionary “An Foclóir Beag” forms part of the system and includes part of speech information and the inflected forms of the words.

References

- [1] Calder G (1990). *A Gaelic Grammar*, Gairm Publications, Glasgow
- [2] Faverau F (1997). *Grammaire du Breton*, Skol Vreizh, Mantroules
- [3] Christian Brothers (1994). *New Irish Grammar*, C.J. Fallon, Dublin
- [4] McElligott A and Ó Néill G (1995). CALL with Methodical Explanations, *Journal of Celtic Language Learning*, 1: - 52, NAACLT.
- [5] *An Foclóir Beag*, An Gúm (1991) Dublin.
- [6] Ó Néill G (2003). *A Guide to Irish Pronunciation with Various Applications*, 5th Digitalstream Conference, CSUMB, Ca
- [7] Charonnat L (1998). *Synthèse de la Parole*, Université de Limerick
- [8] McElligott A and Ó Néill G (1993). *A Dictionary for CALL* EuroCALL, Hull.