Syntactic Alternations in Learner Language

Julia Krivanek
Universität Tübingen
SFB 833, Project A4

Second Tübingen-Berlin Meeting on Analyzing Learner Language
Tübingen, 5./6. Dezember 2011
Overview

Introduction

Levin Classification

The Setup
  Corpora
  TGrep2

Research Steps

First Results
Correspondence between verbal meaning and syntax

- *Diathesis alternations* - alternations in the realization of verbs’ argument structure (Levin 1993)
- Verbs with the same diathesis alternations
  - share certain meaning components
  - ⇒ form a *semantically coherent class*
- The knowledge about **verbs’ meaning** enables *speakers of English* to make subtle judgements about *syntactical behaviour of a verb*
Examples

(1) a. *Martha carved a toy for the baby.
   b. *Martha carved baby a toy.

(2) a. *Carmen bought a dress for Mary.
   b. Carmen bought Mary a dress.

(3) a. *She was named press secretary.
   b. *She was named as press secretary.

(4) a. *The president appointed her Head of the Committee.
   b. *The president appointed her as Head of the Committee.
The meaning of a verb enables **native speakers of English** to build grammatically correct utterances with the given verb.

What about **non-native English speakers**?

Native speakers vs. learner English:
- Distribution of alternations
- Ungrammatical sentences
Motivation

- Error detection in learner texts
- Recognition of native and non-native learner texts
- Recognition of learner texts’ levels
- Recognition of the mother tongue
Levin Classification

- Alternations
- Verb classes
International Corpus of Learner English (ICLE)

- Argumentative essays each 500-1000 words long
- Higher intermediate to advanced learners of English
- 3.7 million words
- 16 mother tongue backgrounds
Louvain Corpus of Native English Essays (LOCNESS)

- Essays written by English and American university students
- 300000 words
- Topics: transport, the parliamentary system, fox hunting, boxing, the National Lottery, etc.
TGrep2

- Unix-based tool to search syntactically and POS-annotated corpora
- Query Example:
  \[ VP << (^{VB^*}/$.)(NP$.)(PP < (IN < as) < NP)) \]
Research Steps

1. Parse ICLE and LOCNESS with Berkeley Parser
2. Add Lemma Information to the parsed corpora (from TreeTagger)
3. Choose verbs and their alternations
4. Create TGrep2-queries
5. Count alternations for each verb in ICLE and LOCNESS
Verb Selection

- Frequency - more than 100 occurrences
- Ambiguity - verbs belonging only to one class
- Relationship between selected verbs
  - *buy* vs. *sell*
  - *talk* vs. *say*
- Selected verbs: agree, assume, name, sell, buy, search, talk, say
Alternation Selection

- Patterns are limited to those which can be found by using syntactic information
- Selected alternations: dative, benefactive, as alternation, etc.
Agree

- Class: **Correspond Verbs**
- Frequency: 1119 times in the ICLEv2 corpus

<table>
<thead>
<tr>
<th>Num.</th>
<th>Pattern</th>
<th>Occur.</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Bill agreed that ...</td>
<td>256</td>
<td>Even a brilliant scientist will agree that this is the key to the mystery of life.</td>
</tr>
<tr>
<td>2.</td>
<td>Bill agreed with Kathy.</td>
<td>201</td>
<td>Yeah, I agree with her.</td>
</tr>
<tr>
<td>3.</td>
<td>Bill and Kathy agreed</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>*Bill agreed Kathy</td>
<td>3</td>
<td>Also, women who are under 16 and their parents agree them to take abortion too.</td>
</tr>
<tr>
<td>5.</td>
<td>*Bill agreed</td>
<td>1</td>
<td>You may not agree but technology has to do with that apathy in people, with that total alienation.</td>
</tr>
</tbody>
</table>

Table: Syntactic patterns for the verb *agree* and their frequency in the ICLEv2 corpus
Agree: Alternations

- Collective NP Subject (patterns 3 and 5)
- Simple Reciprocal Alternation: allow both ’NP1 V [P NP2]’ and [NP1 and NP2] V’ frames (patterns 2 and 3)
- *Understood Reciprocal Object Alternation:
  - The intransitive verb agree (pattern 3) cannot be paraphrased by its transitive variant when it takes the reciprocal each other as object (pattern 4).
- *With Preposition Drop Alternation (patterns 2 and 4)
Name

- Class: **Dub Verbs**
- Frequency: 145 times in the ICLE corpus

<table>
<thead>
<tr>
<th>Num.</th>
<th>Pattern</th>
<th>Occur.</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>He named the ship Seafarer</td>
<td>7</td>
<td>We name it multimedia.</td>
</tr>
<tr>
<td>2.</td>
<td>*He named the ship as Seafarer</td>
<td>9</td>
<td>We might name the other class as the inventions in communication.</td>
</tr>
<tr>
<td>3.</td>
<td>*He named him to be Seafahrer</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Table:** Syntactic patterns for the verb *name* and their frequency in the ICLEv2 corpus

- Alternations:
  - *As Alternation
  - *Infinitival Copular Clause
Difficulties

- The size of LOCNESS
- Parsing errors → Difficult to get precise matches
  ⇒ Search for patterns, not for alternations
References I