Scuola Normale Superiore

Piazza dei Cavalieri 7

Pisa

4 – 5 November 2010

Aula Dini

Palazzo del Castelletto

Interdisciplinary Workshop on Verbs

The Identification and Representation of Verb Features

Scuola Normale Superiore – Laboratorio di Linguistica

Università di Pisa – Dipartimento di Linguistica

EDITED BY

Pier Marco Bertinetto

Anna Korhonen

Alessandro Lenci

Alissa Melinger

Sabine Schulte im Walde

Aline Villavicencio

Scuola Normale Superiore

Normale 200 1810 2010

Università di Pisa

Proceedings of Verb 2010

Interdisciplinary Workshop on Verbs

The Identification and Representation of Verb Features

Scuola Normale Superiore – Laboratorio di Linguistica
Università di Pisa – Dipartimento di Linguistica

edited by

Pier Marco Bertinetto
Anna Korhonen
Alessandro Lenci
Alissa Melinger
Sabine Schulte im Walde
Aline Villavicencio

Scuola Normale Superiore
Pisa, 4-5 November 2010

Preface

Verbs and their features have always received wide attention in various disciplines concerned with linguistic research, since their contribution is essential to the structure and the interpretation of language. In recent years, the availability of new lexical resources and increasingly large corpora, the application of empirical methods and statistical algorithms and the development of technical devices such as eye-trackers and magnetic resonance imaging has led to advances in several linguistic areas.

Their great interest and relevance notwithstanding, verbs still defy attempts by linguists and cognitive scientists to achieve a clear understanding of their organizational principles, as well as of the features entering into their constitution. Verb complexity derives not only from their notoriously high polysemy, but also and especially from the fact that verbs are crucially the cornerstone of the syntax-semantics interface. The semantic behaviour of verbs is therefore strongly intertwined with the syntagmatic constraints governing their distributions. As a consequence, while there is a consensus on the multifarious nature of verb semantic representations, however, the different types of verb features analysed in the literature (e.g., event properties, argument structure, aspect, etc.) still lie as separate pieces of a puzzle which is far to be completed.

Success in this type of research is brought about by close collaboration between (computational) linguists and cognitive scientists. To this end, interdisciplinary workshops can play a key role in advancing existing and initiating new research. This was demonstrated by the interest generated on the Verb Workshop 2005, which received 33 submissions and was held as a standalone event at Saarland University over 2 days. A more clear understanding of the (computational) linguistic and cognitive properties of verbs will bring a positive reflect on the results of the research done within these communities. Therefore there is a real need to provide a forum where researchers can meet across disciplines.

In the call for papers, we solicited papers focusing on the following issues:

Empirical studies and formal descriptions of verb features and verb senses: these are some of the key fundamental factors in verb treatment, and are relevant for representing and distinguishing verbs across disciplines.

Representation of verbs by verb classes: generalisation is crucial to the acquisition of verbs and categorisation in cognitive linguistics, and for many computational linguistic tasks; computational learning of verb classes and properties provides insights into argument alternations, verb polysemy, selectional preferences, etc.

Cognitively motivated models of verbs: the definition of verb semantics according to human perception, the collection of human judgments on verb senses and verb properties, and psycholinguistic studies and experiments on verbs are important interdisciplinary contributions to verb characterisation.

Evidence from cognitive neuroscience and neuropsychology on verb features. **Corpus-based methods to extract empirical features:** the distributional account of verb senses and verb features provides essential contributions to verb analysis. We also welcome contributions on the use of distributional data to model (neuro)cognitive evidence on verb representation.

Data resources and tools: the definition of verb senses and verb properties are important for basic and task-oriented research; especially the annotation of lexical verb information provides valuable data to computational learning procedures and evaluation methods.

Language-specific and cross-linguistic aspects of verbs: which verb features are specific to a language, and which are universal?

Most of these topics lie at the heart of the papers accepted to the workshop. Verb 2010 received 69 submissions, and acceptance was very competitive; the acceptance rate was 23% for oral presentations (16/69 submissions) and 35% for poster presentations (24/69).

We would like to thank all the authors who submitted papers, as well as the members of the Program Committee for the time and effort they contributed in reviewing the papers.

Verb 2010 is sponsored by the EU-Project PANACEA (http://www.panacea-lr.eu).

Pisa, November 2010

Pier Marco Bertinetto
Anna Korhonen
Alessandro Lenci
Alissa Melinger
Sabine Schulte im Walde
Aline Villavicencio

Workshop organization

Workshop Organizers

Pier Marco **Bertinetto**
Scuola Normale Superiore - Pisa - Italy

Anna **Korhonen**
University of Cambridge – UK

Alessandro **Lenci**
University of Pisa - Italy

Alissa **Melinger**
University of Dundee - UK

Sabine **Schulte im Walde**
University of Stuttgart - Germany

Aline **Villavicencio**
Federal University of Rio Grande do Sul – Brazil
University of Bath – UK

Program Committee

Artemis **Alexiadou**
Department of Linguistics, University of Stuttgart, Germany
Afra **Alishashi**
Department of Computational Linguistics and Phonetics, Saarland University – Germany
Tim **Baldwin**
Department of Computer Science and Software Engineering, University of Melbourne – Australia
Colin **Bannard**
Max Planck Institute for Evolutionary Anthropology, Leipzig – Germany
Roberto **Basili**
Department of Computer Science, University of Roma Tor Vergata – Italy
Nuria **Bel**
University Pompeu Fabra, Barcelona – Spain
Gemma **Boleda**
Universitat Politècnica de Catalunya – Spain
Chris **Brew**
Department of Linguistics, Ohio State University – USA
Miriam **Butt**
Department of Linguistics, University of Konstanz – Germany
Dick **Crouch**
Microsoft, USA
Barry **Devereux**
University of Cambridge

Amit **Dubey**
School of Informatics, University of Edinburgh – UK
Magali **Duran**
USP/SC, Brazil
Sonja **Eisenbeiß**
Department of Language and Linguistics, University of Essex – UK
Katrin **Erk**
Department of Linguistics, University of Texas at Austin - USA
Afsaneh **Fazly**
Department of Electrical and Computer Engineering, Shiraz University – Iran
Sam **Featherston**
Department of General and Theoretical Linguistics, University of Tübingen
Pablo **Gamallo**
Universidade de Santiago de Compostela – Spain
Yuki **Kamide**
School of Psychology, University of Dundee
Tracy **King**
Microsoft – USA
Beata Beigman **Klebanov**
Northwestern Institute for Complex Systems, Northwestern University, USA
Jean-Pierre **Koenig**
Department of Linguistics, University at Buffalo – USA
Valia **Kordoni**
Saarland University and DFKI GmbH
Beth **Levin**
Department of Linguistics, Stanford University – USA
Bernardo **Magnini**
Fondazione Bruno Kessler, Trento – Italy
Daniela **Marzo**
Department of Linguistics, University of Stuttgart – Germany
Gail **Mauner**
Department of Psychology, University of Buffalo
Diana **McCarthy**
Lexical Computing Ltd., Brighton – UK
Ken **McRae**
Department of Psychology, University of Western Ontario – Canada
Paola **Merlo**
University of Geneva – Switzerland
Brian **Murphy**
CIMeC, University of Trento
Rainer **Osswald**
Department of Linguistics, University of Duesseldorf, Germany
Sebastian **Padó**
Institute for Natural Language Processing, University of Stuttgart – Germany
Martha **Palmer**
Department of Linguistics, University of Colorado at Boulder – USA
Chris **Parisien**
Department of Computer Science, University of Toronto, Canada
Massimo **Poesio**
Center for Mind/Brain Sciences, University of Trento – Italy

James **Pustejovsky**
Department of Computer Science, Brandeis University - USA
Carlos **Ramisch**
University of Grenoble, France
Billi **Randall**
University of Cambridge
Anna **Rumshisky**
Department of Computer Science, Brandeis University – USA
Christoph **Scheepers**
Department of Psychology, University of Glasgow
Tatiana **Schnur**
Department of Psychology, Rice University
Ekaterina **Shutova**
Computer Laboratory, University of Cambridge – UK
Maity **Siqueira**
Federal University of Rio Grande do Sul
Achim **Stein**
Department of Linguistics, University of Stuttgart, Germany
Suzanne **Stevenson**
Department of Computer Science, University of Toronto – Canada
Britta **Stolterfoht**
Department of German Linguistics, University of Tübingen, Germany
Lauren **Tonietto**
Institute of Psychology, Federal University of Rio Grande do Sul, Brazil
Beata **Trawinski**
Universitaet Wien, Austria
Gabriella **Vigliocco**
Department of Psychology, University College London – UK
David **Vinson**
Deafness Cognition and Language Research Center, University College London – UK
Alessandra **Zarcone**
Department of Computational Linguistics, University of Stuttgart, Germany
Yi **Zhang**
Saarland University

Local Organization

Pier Marco **Bertinetto**
Scuola Normale Superiore, Pisa – Italy
Alessandro **Lenci**
University of Pisa – Italy
Martina **Johnson**
University of Pisa – Italy
Irene **Ricci**
Scuola Normale Superiore, Pisa – Italy

Table of Contents

Invited speakers

Oral presentations

Poster presentations

DANTE: a New Resource for Research at the Syntax-Semantics Interface

Diana McCarthy
Lexical Computing Ltd,
Brighton, UK
diana@dianamccarthy.co.uk

Abstract

Since Levin's seminal work (Levin, 1993) there has been a rising interest in computational linguistics research which aims to examine the relationship between the syntax and semantics of verbs. A substantial portion of the work comprises efforts to discover semantic classes from syntactic behaviour and also from selectional preferences. There is also some work on directly examining related phenomena, such as detecting subcategorisation frames and diathesis alternations. Work in this area is typically corpus based, although many manually constructed resources have also been used as start points and for evaluation. In this paper, we present an English lexical database (being finalised at the time of writing, and to be released late 2010) which we believe will be a major catalyst for work of this nature, both as a starting point for automatic methods and as a gold standard for evaluation

1 Introduction

There has been a growing interest in computational linguistics in the semantics-syntactic interface, particularly as regards verbs. A trigger for this was Levin's work (Levin, 1993) on verbs which, following her predecessors e.g (Fillmore, 1967), demonstrated that given that a verb's meaning is related to its syntactic behaviour, we can group verbs into semantic classes by virtue of their shared syntactic behaviour. A key issue in any research on this relationship is identifying what the key syntactic behaviour and semantic components are since there are a great many possibilities and it is a non trivial task to identify the appropriate features. Diathesis alternations are different surface realisations of a verbs arguments. Levin's work demonstrated that diathesis alternations are extremely useful in classifying verbs.

Levin's alternation inventory, whilst the first of its kind and providing a broader and more thorough manual analysis than anything that had been been available before, was restricted to a subset of subcategorisation frames (SCFs) involving NPs and PPs, i.e excluding sentential complements. The resource was produced manually and not from corpus examples. Baker and Ruppenhofer (2002) point out that many examples of syntactic behaviour Levin provides, are not attested in the corpus data (the BNC (Leech, 1992)) that they used for the FrameNet project. Furthermore, actual use of alternations for verb classification would give rise to a finer granularity than is present in Levin's classification; many of Levin's classes are semantically motivated, rather than being totally determined by the alternation behaviour. Despite these limitations, the book has triggered a large amount of research in computational linguistics in automatically identifying the links between syntactic behaviour and verb meaning.

Prior to the work on automatic classification, there was research on automatic acquisition of verbal information from corpora that would in turn be exploited for subsequent work on classification. Acquisition of SCFs (Brent, 1991; Manning, 1993) was conducted with a view to improving results in parsing (Carroll et al., 1998). Selectional preference acquisition (Resnik, 1993) was performed to help with structural and lexical ambiguity resolution (Li and Abe, 1998; Resnik, 1997; McCarthy and Carroll, 2003). Levin's work spurred further research using automatically acquired lexical information for diathesis alternation identification (McCarthy, 2000; McCarthy and Korhonen, 1998; Lapata, 1999) and for verb classification (Schulte im Walde, 2006; Sun and Korhonen, 2009; Stevenson and Merlo, 1999; Merlo and Stevenson, 2001).

In this paper we will give a very brief overview of the lexical acquisition work in this direction [1], and a summary of some of the key existing lexical resources that can be used as input to the work or for evaluation purposes. We then describe DANTE (Atkins et al., 2010) a recently released lexical database produced by a team of lexicographers scrutinising a 1.7 billion word corpus of English. The database includes over 6,300 headword verbs with just under 3000 phrasal verbs with just under 300,000 examples of the various features of these verb and phrasal verb entries. [2] We expand on the potential of this resource for lexical research and we end by highlighting the possibilities for integration of DANTE with existing lexical resources to further its potential yet still.

While there is interesting related work in other languages (Schulte im Walde and Brew, 2002) the bulk of the resources and lexical acquisition work in this area has been with regard to English. DANTE presented here is also an English resource. For this reason, this paper will focus on how DANTE relates to English resources. Fully automatic methods that simply use such resources for evaluation are in many cases applicable to languages other than English.

2 Background: automatic acquisition of verbal subcategorisation, selectional preferences, diathesis alternations and semantic class

We will highlight some key contributions, but unfortunately have not been able to include all due to lack of space.

[1]Related topics of semantic role labelling, word sense induction and word sense disambiguation are outside the scope of this paper.

[2]There is likewise a wealth of information and examples for other PoS, but we do not go into those details here.

2.1 Automatic Acquisition of SCF and Selectional Preferences

There have been many works on automatic acquisition of SCFs. The earliest is due to Brent (1991) who proposed a system capable of recognising five frames, using information from unambiguous cases, for example using pronouns for detecting noun phrases. Following this pioneering work there has been increasing attention paid to a more comprehensive classification, and coverage of more data using statistical techniques to filter parser errors. Briscoe and Carroll (1997) developed a system distinguishing 161 SCFs and, because it is not restricted to unambiguous input, can output relative frequencies of these frames for a given verb. Korhonen (2002) made various refinements of the system, including use of Levin style verb classes to improve statistical filtering to distinguish genuine frames from parser noise. Preiss et al. (2007) extended this approach to adjective and nominal frames.

Alongside the acquisition of SCFs, work has been conducted on selectional preference acquisition using data in the argument heads of these frames (McCarthy, 2000), or directly on parser output (Resnik, 1993; Li and Abe, 1998). Erk (2007) used example sentences from FrameNet as input to selectional preference acquisition. Early work used WordNet to provide classes for generalisation of the preferences (Resnik, 1993; Li and Abe, 1998; Clark and Weir, 2002), but more recently there has been work using distributional similarity for generalisation (Erk, 2007; McCarthy et al., 2007)

2.2 Automatic Identification of Verbal Participation in Diathesis Alternations

Resnik (1993) demonstrated a link between selectional preference strength and participation in alternations where the direct object can be omitted. e.g. *The boy ate the popcorn.* ↔ *The boy ate*.
Lapata (1999) identified participation in the dative and benefactive alternations using a shallow parser and various linguistic and semantic cues, which are specified manually for these two alternations. Another approach is to use cues for syntactic frames, coupled with the overlap of lexical fillers between the alternating slots. McCarthy and Korhonen (1998) carried out preliminary experiments which were extended by McCarthy (2000) on detecting 'role switching al-

ternations'. Role switching alternations are defined as those where an argument appears in different slots in different frames, examples are the causative and conative alternations. McCarthy and Korhonen (1998) and McCarthy (2000) used WordNet as a means of generalising the lexical fillers to semantic classes and used Li and Abe (1998)'s selectional preference models to find semantic classes with an appropriate level of generalisation. Tsang and Stevenson (2010) extended this work by a graphical method which compares the probability of the lexical items at the alternating slots in the WordNet hypernym structure as a whole rather than at a set of individual classes cutting across that structure. Using this method they demonstrated an improvement on (McCarthy, 2000), particularly with regard to low frequency verbs.

2.3 Automatic Identification of Verb Classes

In this subsection, we describe approaches which classify verbs according to evidence often also used for diathesis alternation detection, however alternation participation is not overtly detected in these methods. Merlo and Stevenson (2001) detected three major classes of optionally intransitive verbs (unergative unaccusative and object drop) verbs based on argument structure using corpus evidence of transitivity, causativity and animacy of the arguments as well as other surface features such as passivisation. Schulte im Walde (2006) demonstrated that SCF can be used for clustering German verbs. She also experimented with selectional preferences using GermaNet (Kunze and Lemnitzer, 2002) but without finding a significant improvement over syntactic information alone. More recently, (Sun and Korhonen, 2009) demonstrated that unsupervised clustering of the argument heads themselves can be used as selectional preference features which in turn improved the clustering of the verbs when used alongside SCFs in contrast to the SCFs features alone.

3 Lexical Resources Available for Research

The focus here is on verbal information. Note that DANTE and FrameNet also provide a wealth of information on other PoS.

Levin's classification A classification of 3100 verbs into 193 classes based on verbal participation in 80 diathesis alternations, involving mainly NP and PP constituents. This classification was produced manually and examples were obtained from introspection rather than corpus evidence.

VerbNet (Kipper-Schuler, 2005) (Now extended VerbVet) A verbal lexicon comprising 3769 lemmas with 5257 senses organised in hierarchical WordNet classes but supplemented with valuable syntactic information as well as thematic roles and selectional preferences

Propbank (Palmer et al., 2005) A one million word corpus which supplements the Penn Tree Bank (Marcus et al., 1993) and has been annotated with predicate-argument information. The semantic role labels assigned to arguments have meanings that are specific to each verb. This resource is particularly useful for research in semantic role labelling (Màrquez et al., 2008). Although the corpus is currently limited to Wall Street Journal News text, there is work underway to annotate further corpus data.

Valex (Korhonen et al., 2006) This is an automatically produced SCF lexicon of 6397 verbs using the system of Korhonen (2002) on a corpus of 900 million words. A portion of the output has been evaluated but the lexicon is automatically produced and each individual corpus occurrence has not been validated.

FrameNet (Ruppenhofer et al., 2010) is a lexicon produced from analysed texts that places lexical units (senses) in semantic frames, for example **removing** or **emptying** which classify verbs (and nouns and adjectives) according to the semantic frames that they participate in. Examples are provided from the BNC and an American newswire corpus. The database currently includes 135,000 corpus sentences for over 10,000 lexical entries (nouns, verbs and adjectives) in approximately 800 frames.

WordNet (Fellbaum, 1998) A list of 11529 verbs [3] (including multiword expressions marked as verbs) with synonyms and semantic relations marked. Although there is some information on derived forms and some domain tags, the resource is focused on senses

[3] Here we refer to the latest version of WordNet: version 3.0

and semantic relationships e.g. troponymy and entailment, and does not include syntactic, grammatical and collocational behaviour.

In the following section we describe DANTE, a new lexical database built from inspection of 1.7 billion word corpus.

4 DANTE

DANTE (Database of ANalysed Texts of English) [4] was produced during the first stage of production of a New English Irish Dictionary, and is funded by Foras na Gaeilge, the official body for the (Gaelic) Irish language. DANTE is a target-language-neutral monolingual analysis of the source language listing all the phenomena that might possibly have an unexpected translation. DANTE is a collection of lexical entries with information and examples on every variety of lexical information that the lexicographers have deemed potentially relevant for a thorough and accurate description of English. DANTE relates to the Corpus Pattern Analysis approach of Hanks (Forthcoming) in that a major focus is the prototypical syntagmatic patterns of words in use.

The project team combined expertise in corpora, computational linguistics and lexicography, and from the very outset the project has been solidly corpus-based The corpus used comprised 1.7 billion words from the UKWaC (Ferraresi et al., 2008), some contemporary American newspaper text and Irish English data from the NCI (Kilgarriff et al., 2006). This data was then part-of-speech tagged with TreeTagger [5] and loaded into the Sketch Engine corpus query system (Kilgarriff et al., 2004).

The distinctive feature of the Sketch Engine is 'word sketches': one-page, corpus-driven summaries of a word's grammatical and collocational behaviour. The corpus is parsed using a simple tag sequence grammar and a table of collocations is extracted for each grammatical relation. For DANTE, the set of grammatical relations was defined to give an exact match to the grammatical patterns that the lexicographers were to record. The word sketch for the word would, in so far as the PoS-tagging, parsing, and statistics worked correctly, identify precisely the grammatical patterns and collocations that the lexicographer needed to note in the dictionary. Figure 1 shows a smallish portion of the word sketch for the verb *blend*. The interface allows for seamless switching between specific collocations in the word sketch and a concordance containing those collocations. This switching from the word sketch to concordance is extremely useful for finding examples of significant phenomena. A key feature of DANTE, is that all lexical information is supplemented with example sentences from the corpus. The examples were not edited making them ideal for building and evaluating robust computational linguistics systems which can cope with real language. In order to help the lexicographers find good examples for the phenomena under scrutiny an automatic program (GDEX) that is part of the sketch engine suite of tools was used for sorting the examples so that the 'best' (according to a set of heuristics) are shown to the lexicographer first (Kilgarriff et al., 2008).

4.1 Lexical Information within DANTE

For a full description of the contents of DANTE, refer to the web site [6] and (Atkins et al., 2010). Here we provide a summary of information pertinent to automatic lexical acquisition of verbs. [7] Note that all the subsequent categories of information are associated with word senses.

senses Lexicographers break headwords into senses based on corpus evidence and provide examples of each, along with brief definitions. The definitions are designed to differentiate one sense from another within the same entry for a given lemma and are not as polished as they would be in a conventional dictionary. The focus in DANTE is on comprehensive corpus citations as examples of all lexical information. Extensive exemplification of senses are potentially more useful to computational approaches compared to definitions which are produced for human readers.

subcategorisation frames There are 42 frames in total for verbs, with additional specification of preposition (see figure 2). These are based on the work of Charles Fillmore and are described in (Atkins et al., 2003).

[4] DANTE is described at www.danteweb.com where you can also find a interface for querying the database.

[5] www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/

[6] http://www.webdante.net/

[7] In this paper we provide our own labels for information provided in DANTE.

blend *(verb)* **LEXMCI freq = 13521**

Constructions		
Ving	528	5.8
that_0	195	0.9
NP_Vinf_to	192	2.9
Vinf_to	159	0.6

NP	6083	5.3
learning	345	5.94
approach	212	3.63
solution	146	3.52
oil	114	3.84
sound	98	3.84

Part	63	1.4
in	40	4.53

NP_Part	37	1.1
down	16	0.11
in	5	1.53

Part_PP	26	1.1
down_to	6	2.72

NP_PP	2208	5.3
with	975	1.58
into	217	1.63
along	7	0.11

AJP	107	0.4
busy	6	0.87

AVP	3049	4.4
together	480	5.35
well	473	3.7
seamlessly	132	9.13
perfectly	115	6.19
then	114	1.69

subj_NP	2101	2.6
voice	43	2.82
colour	27	1.68
style	26	1.27
hand	22	0.62
vocal	19	4.0

PP_X	2604	
PP_with-i	1347	16.8
PP_into-i	655	36.2
PP_in-i	259	1.1
PP_for-i	66	0.5
PP_from-i	55	1.0

PP_PP	543	1.5
with_of	166	2.96
into_of	58	3.73
with_in	25	1.82
with_from	20	3.23
with_for	18	2.43

Figure 1: A portion of the word sketch for *blend*.

inherent grammar e.g. *rain* impersonal

multiword expressions including idioms, support verbs, phrasal verbs, compounds, chunks

collocations e.g. *fire* (discharge a weapon) NP collocations *shot*, *round*, *gun* . . .

corpus patterns tendencies e.g. plural noun as object

usage markers include:

- evaluative e.g. *meddle* (pejorative)
- regional variety e.g. *nick* (British) as in *you're nicked*
- domain e.g. *multiply* (maths)

4.2 DANTE as a Resource for Research at the Syntax-Semantics Interface

DANTE is being released without charge for research purposes. For computational linguistics, and perhaps also other linguistics research it is the combination of syntactic, semantic and usage information alongside numerous examples that makes DANTE stand out in contrast to previously available resources. While some existing resources do have corpus examples (Ruppenhofer et al., 2010; Palmer et al., 2005), DANTE provides a far greater number (300,000 for verb and phrasal verb entries alone) and from a far larger and more and varied source (in contrast to previous resources with examples from the BNC (FrameNet) or the Wall Street Journal (Propbank)) with manual verification of the data (in contrast to automatically produced resources such as valex (Korhonen et al., 2006)). This makes it a perfect resource for systems which experiment with data exhibiting specific phenomena e.g. particular SCFs for diathesis alternation detection contrasting argument fillers at different slots. For example, the PP slots in the two NP_PP_X frames with prepositions *with* and *into* as exemplified in figure 2. While it is of course possible to use automatic resources as a start point (McCarthy, 2000) use of DANTE would enable researchers to isolate PoS, parser error and other sources of noise that are difficult to avoid (Korhonen et al., 2000) when using fully automatic methods.

In addition to the 300,000 verbal manually verified corpus examples [8] it is possible to obtain further examples direct from the 1.7 billion word corpus using the SCF and collocation information. Indeed, this information is already being used in a preliminary word sense disambiguation project. [9] Computational linguistic approaches for selectional preference and diathesis alternation acquisition could use the data to gather argument heads in specific slots of SCFs. Since all the data is assigned to word senses, and the word senses

[8] There are 622,000 examples over all PoS.

[9] See http://www.webdante.com/disambiguation_project.html.

blend: (PoS: v)
`meaning:` combine
`SCF:` NP
`corpus pattern:` with plural noun as object
`example:` *I have very little idea of how to **blend** colour.*
`corpus pattern:` blend sth and sth
`example:` *High Points : The attempt to **blend** melodrama comedy and horror is a worthy if failed effort.*
`SCF:` NP_PP_X with
`example:` *Kazakhstan was interested in **blending** palm oil with its own cotton seed and sunflower seed oils for industrial application , officials said.*
...
`SCF:` NP_PP_X into
example: *I **blend** different colours into the background of my paintings to evoke sections of light .*

Figure 2: A portion of the entry for *blend*. The portion has been simplified and shortened for presentation here, with only a couple of examples and features shown. Further examples are provided at http://www.webdante.net/.

have associated usage information, there is scope for doing experiments linking sense to syntactic behaviour. Moreover, as well as a start point for acquisition, the resource can be used as a gold standard for evaluation of automatic acquisition of information contained therein such as SCF, sense induction, sense disambiguation and usage, for example domain.

5 Conclusion

In this paper we have presented the DANTE lexical resource which we believe will prove a useful resource for computational linguistics, particularly at the syntax-semantics interface but elsewhere also. We have suggested ways in which the data therein could be used as a starting point for research at the syntax-semantics interface, for example alternation detection and selectional preference acquisition, and also as a resource for lexical acquisition evaluation.

There are a multitude of resources for English dealing with predicate-argument structure and word sense. No one resource is a panacea and researchers have already highlighted the merits of combining resources (Merlo and van der Plas, 2009). SemLink [10] is a great initiative in this direction with mappings between VerbNet and propbank and VerbNet and FrameNet. Atkins (2010) proposes possibilities in this direction for combining DANTE with FrameNet using syntactic information common to both and distributional thesauruses (such as those in Sketch Engine) for relating lexical units. We believe that interesting research will result from such endeavours and that, as well as automatic approaches for linking these resources should prove interesting in their own right.

[10] http://verbs.colorado.edu/semlink/

Acknowledgments

Thanks to my colleagues on the DANTE project, particularly Sue Atkins, Adam Kilgarriff, Cathal Convery, Michael Rundell, Diana Rawlinson and Valerie Grundy. I am indebted to Sue Atkins, Adam Kilgarriff and Anna Korhonen for comments on an earlier draft. Any mistakes in this article are solely my own responsibility.

References

Sue Atkins, Charles Fillmore, and Christopher Johnson. 2003. Lexicographic relevance: selecting information from corpus evidence. *International Journal of Lexicography*, 16(3):251–280.

Sue Atkins, Michael Rundell, and Adam Kilgarriff. 2010. Database of ANalysed Texts of English (DANTE). In *Proceedings of Euralex*.

Sue Atkins. 2010. The DANTE database: Its contribution to English lexical research, and in particular to complementing the FrameNet data. In Gilles-Maurice de Schryver, editor, *A Way with Words: Recent Advances in Lexical Theory and Analysis. A Festschrift for Patrick Hanks*. Menha.

Collin F. Baker and Josef Ruppenhofer. 2002. FrameNet's frames vs. Levin's verb classes. In J. Larson and M. Paster, editors, *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*, pages 27–38.

Michael R. Brent. 1991. Automatic acquisition of subcategorization frames from untagged text. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pages 209–214.

Ted Briscoe and John Carroll. 1997. Automatic extraction of subcategorization from corpora. In *Proceedings of the Fifth Applied Natural Language Processing Conference*, pages 356–363.

John Carroll, Guido Minnen, and E. Briscoe. 1998. Can subcategorisation probabilities help a statistical parser? In *Proceedings of the 6th ACL/SIGDAT Workshop on Very Large Corpora*, pages 118–126, Bergen, Norway.

Stephen Clark and David Weir. 2002. Class-based probability estimation using a semantic hierarchy. *Computational Linguistics*, 28(2):187–206.

Katrin Erk. 2007. A simple, similarity-based model for selectional preferences. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 216–223, Prague, Czech Republic, June. Association for Computational Linguistics.

Christiane Fellbaum, editor. 1998. *WordNet, An Electronic Lexical Database*. The MIT Press, Cambridge, MA.

Adriano Ferraresi, Eros Zanchetta, Marco Baroni, and Silvia Bernardini. 2008. Introducing and evaluating ukwac, a very large web-derived corpus of english. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008)*, Marrakech, Morocco.

Charles Fillmore. 1967. The grammar of *hitting* and *breaking*. In Roderick A. Jacobs and Peter S. Rosenbaum, editors, *Readings in English Transformational Grammar*, pages 120–133. Ginn and Company: a Xerox Company, Waltham M.A.

Patrick Hanks. Forthcoming. *Lexical Analysis: Norms and Exploitations*. MIT Press.

Adam Kilgarriff, Pavel Rychly, Pavel Smrz, and David Tugwell. 2004. The sketch engine. In *Proceedings of Euralex*, pages 105–116, Lorient, France.

Adam Kilgarriff, Michael Rundell, and Elaine Uí Dhonnchadha. 2006. Efficient corpus development for lexicography: building the new corpus for Ireland. *Language Resources and Evaluation Journal*, 40(2):127–152.

Aadam Kilgarriff, Milos Husák, Katy McAdam, Michael Rundell, and Pavel Rychly. 2008. GDEX: Automatically finding good dictionary examples in a corpus. In *Euralex Proceedings*, Barcelona.

Karin Kipper-Schuler. 2005. *VerbNet: A broad-coverage, comprehensive verb lexicon*. Ph.D. thesis, Computer and Information Science Dept., University of Pennsylvania. Philadelphia, PA.

Anna Korhonen, Genevieve Gorrell, and Diana McCarthy. 2000. Statistical filtering and subcategorization frame acquisition. In *Proceedings of the 2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora.*, pages 199–206, Hong Kong. ACL.

Anna Korhonen, Yuval Krymolowski, and Ted Briscoe. 2006. A large subcategorization lexicon for natural language processing applications. In *Proceedings of the 5th international conference on Language Resources and Evaluation*, Genova, Italy.

Anna Korhonen. 2002. *Subcategorization Acquisition*. Ph.D. thesis, University of Cambridge.

Claudia Kunze and Lothar Lemnitzer. 2002. Germanetrepresentation, visualization, application. In *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC)*, Las Palmas, Canary Islands, Spain.

Maria Lapata. 1999. Acquiring lexical generalizations from corpora: a case study for diathesis alternations. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, pages 397–404.

Geoffrey Leech. 1992. 100 million words of English: the British National Corpus. *Language Research*, 28(1):1–13.

Beth Levin. 1993. *English Verb Classes and Alternations: a Preliminary Investigation*. University of Chicago Press, Chicago and London.

Hang Li and Naoki Abe. 1998. Generalizing case frames using a thesaurus and the MDL principle. *Computational Linguistics*, 24(2):217–244.

Christopher Manning. 1993. Automatic acquisition of a large subcategorization dictionary from corpora. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, pages 235–242.

Mitchell Marcus, Beatrice Santorini, and Mary Marcinkiewicz. 1993. Building a large annotated corpus of english: the penn treebank. *Computational Linguistics*, 19(2).

Lluís Màrquez, Xavier Carreras, Ken Litkowski, and Suzanne Stevenson. 2008. Semantic role labelling: an introduction to the special issue. *Computational Linguistics*, 34(2):145–159.

Diana McCarthy and John Carroll. 2003. Disambiguating nouns, verbs and adjectives using automatically acquired selectional preferences. *Computational Linguistics*, 29(4):639–654.

Diana McCarthy and Anna Korhonen. 1998. Detecting verbal participation in diathesis alternations. In *Proceedings of the 17th International Conference on Computational Linguistics and the 36th Annual Meeting of the Association for Computational Linguistics*, volume 2, pages 1493–1495.

Diana McCarthy, Sriram Venkatapathy, and Aravind Joshi. 2007. Detecting compositionality of verb-object combinations using selectional preferences. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 369–379.

Diana McCarthy. 2000. Using semantic preferences to identify verbal participation in role switching alternations. In *Proceedings of the First Conference of the North American Chapter of the Association for Computational Linguistics. (NAACL)*, pages 256–263, Seattle,WA.

Paola Merlo and Suzanne Stevenson. 2001. Automatic verb classification based on statistical distribution of argument structure. *Computational Linguistics*, 27(3):373–408.

Paola Merlo and Lonneke van der Plas. 2009. Abstraction and generalisation in semantic role labels: PropBank, VerbNet or both? In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 288–296, Suntec, Singapore, August. Association for Computational Linguistics.

Martha Palmer, Dan Gildea, and Paul Kingsbury. 2005. The proposition bank: A corpus annotated with semantic roles. *Computational Linguistics*, 31(1):71–106, sept.

Judita Preiss, Ted Briscoe, and Anna Korhonen. 2007. A system for large-scale acquisition of verbal, nominal and adjectival subcategorization frames from corpora. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 912–919, Prague, Czech Republic, June. Association for Computational Linguistics.

Philip Resnik. 1993. *Selection and Information: A Class-Based Approach to Lexical Relationships*. Ph.D. thesis, University of Pennsylvania.

Philip Resnik. 1997. Selectional preference and sense disambiguation. In *Proceedings of the* SIGLEX *Workshop on Tagging Text with Lexical Semantics: Why What and How?*, pages 52–57, Washington, DC.

Josef Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. 2010. FrameNet II: Extended theory and practice. Technical report, International Computer Science Institute, Berkeley, June. http://framenet.icsi.berkeley.edu/.

Sabine Schulte im Walde and Chris Brew. 2002. Inducing german semantic verb classes from purely syntactic subcategorization information. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL-2002)*.

Sabine Schulte im Walde. 2006. Experiments on the automatic induction of german semantic verb classes. *Computational Linguistics*, 32(2):159–194.

Suzanne Stevenson and Paola Merlo. 1999. Automatic verb classification using distributions of grammatical features. In *Proceedings of the Nineth Conference of the European Chapter of the Association for Computational Linguistics*, pages 45–52.

Lin Sun and Anna Korhonen. 2009. Improving verb clustering with automatically acquired selectional preferences. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 638–647, Singapore, August. Association for Computational Linguistics.

Vivian Tsang and Suzanne Stevenson. 2010. A graph-theoretic framework for semantic distance. *Computational Linguistics*, 36(1), March.

Event-based Thematic Role Concepts

Ken McRae

University of Western Ontario, Canada

There are considerable differences among researchers' conceptions of the semantic content of verbs' thematic roles. In many theories, semantic content is limited to (grammatically-relevant) binary selectional restrictions such as ±animacy. In some theories, researchers allow for selectional restrictions of many types, such as ±mailable, ±cookable, or ±inflatable. The notion of selectional restrictions has played an important role in linguistics and psycholinguistics. One reason is that selectional restrictions are viewed as lexically-based, and therefore are assumed to be available to influence on-line language processing more rapidly than is knowledge about real-word events. In contrast, my colleagues and I have been investigating the idea that thematic roles are event-based concepts, and that thematic role assignment during on-line language comprehension reflects this. An event-based conceptual view of thematic processing has a number of implications. For example, it entails a dynamic view of thematic role assignment in which the fit between a verb's thematic role and a particular noun concept depends not only on the specific verb, but also on verb sense (or the class of events to which the verb refers given a particular context). As another example, from an expectancy generation point of view, verb aspect can influence expectancies for upcoming roles. I will present psycholinguistic studies that provide evidence for this view, including word-word priming, self-paced reading, eyetracking, and ERP experiments. These studies demonstrate that although there may be a distinction between lexical constraints on the one hand, and conceptual event-based knowledge on the other, this distinction has no relevance for the time course of the activation and use of these types of knowledge. Thus, there is no architecturally-determined delay of knowledge about real-world events during language comprehension. Furthermore, it appears that selectional restrictions, which are often considered to be lexical-grammatical constraints, and event-based knowledge, which is conceptual, may, in fact, be the same thing.

The interaction of light verbs and verb classes of Urdu

Tafseer Ahmed
Universität Konstanz
Germany
tafseer.khan@uni-konstanz.de

Abstract

The paper describes an attempt of identifying Urdu verb classes on the basis of the distribution of light verbs with different main verbs. We started with a frequency analysis of main + light verb sequences. The analysis of that data lead us to a thorough manual analysis of main + light verb sequences by using native speaker judgments. We focused on the three most frequent light verbs *dE* 'give', *lE* 'take' and *jA* 'go'. The verb classes were identified by considering acceptability/unacceptability of these light verbs with the main verbs. We identified some new classes. For example, we found that mental gain verbs e.g. *samajH* 'understand' are different from mental state verbs e.g. *Dar* 'fear'. The verb classes can also be used to disambiguate different senses of polysemous main verbs and different syntactical usages of the light verb forms.

1 Introduction

Urdu is an Indo-Aryan language. It is closely related to Hindi with a similar grammatical structure, but differences in script and vocabulary.

There is no comprehensive work on the verb classes of Urdu. Some authors have identified interesting classes and syntactic patterns for Urdu verbs. One of these classes consists of bodily expression verbs that allow an optional ergative marker (Butt 1995, Davison 1999). Another interesting class is of ingestive verbs (Saksena 1982, Butt 2006, Ramchand 2008). Khan (2009) identified six classes on the basis of non-canonical second argument of the verb. Ahmed (2010) clustered 184 Urdu verbs on the basis of related light verbs and aspectual auxiliaries and found four major classes. However much work remains to be done and the findings to date need to be verified and integrated with one another.

We follow Levin's (1993) classic assumption that the verb classes can be identified by their syntactic properties. She presented classes of English verbs using alternations related to English verbs. According to her, verbs that have similar syntactic properties also share semantic properties.

All of the alternations presented by Levin are not present in Urdu. An example is the beneficiary alternation that distinguishes some English verb classes. On the other hand, there are some other syntactic patterns that have not been discussed for Urdu. These can be used to classify Urdu verbs. Each light verb is acceptable with some main verbs. There exists a set of main verbs that are not acceptable with that light verbs. Hence the acceptability of certain light verbs with main verbs can be a criterion for identifying verb classes of Urdu.

Section 2 introduces Urdu light verbs. Section 3 explains how we identify the proposed verb classes for Urdu. Section 4 discusses the semantic properties of these classes and their relation to the semantic properties of the light verbs that are allowed/not allowed with these classes.

2 Light Verbs in Urdu/Hindi

In Urdu, we find sequence of verbs in which the main verb is followed by another verb (Schmidt 1999). The second verb of the sequence (that follows the main verb) can be an aspectual auxiliary, a modal or a light verb. A light verb is used to show completeness, suddenness or similar properties. In (1b), the light verb *paR* 'fall' is used that shows suddenness.

(1) a. gARI cal-I
vehicle move-Perf.F.Sg
'The vehicle moved.'

b. gARI cal paR-I
vehicle move fall-Perf.F.Sg
'The vehicle suddenly moved.'

Siddiqui (1971), McGregor (1972) and Hook (1974) provide lists of such verbs. Some of these verbs are: *dE* 'give', *lE* 'take', *A* 'come', *jA* 'go', *DAl* 'insert', *paR* 'fall', *beTH* 'sit', *uTH* 'rise', *dE* 'give', *rakH* 'put', *ban* 'get make', *lag* 'touch/hit', *nikal* 'come out', *Tahar* 'stop' and *cal* 'move'.

Butt and Geuder (2001) used the term light verbs for these verbs.[1] They argue that the light verbs are different from aspectual markers.

Most of the light verbs are not acceptable with all of the main verbs (Hook 1974, Butt and Geuder 2001). Every light verb is acceptable with a set of compatible verbs. Consider the example of the light verb *dE* 'give'. It is not acceptable with the verb *ruk* 'stop' as shown in (2b), however it is acceptable with the verb *cal* 'move'.

(2) a. gARI cal dI
vehicle.F.Sg move give.Perf.F.Sg
'The vehicle moved.'

b. *gARI ruk dI
vehicle.F.Sg stop give.Perf.F.Sg
'The vehicle stopped.'

However, the same verb *ruk* 'stop' is acceptable with the light verb *jA/ga* 'go'.

(3) gARI ruk ga-yI[2]
vehicle.F.Sg move give-Perf.F.Sg
'The vehicle moved.'

In the previous literature, the semantic reasons for the use of these light verbs are mentioned, but there is no mention of their relation to verb classes.

An important issue with the light verbs is that these can be polysemous. McGregor (1972) pointed out that *jA* 'go' has a light verb usage to depict completion. However, it can also occur as a main verb (in conjunction) after another main verb.[3] Similarly, we find main verb + main verb sequences for *dE* 'give' and *A* 'come'. Consider the following examples.

(4) a. cAnd nikal ga-yA
moon emerge go-Perf.M.Sg
'The moon emerged.'

b. vuh [draxt kAT (kar)] ga-yA
3SG tree cut having go-Perf
'Having cut the trees, he went.'

While in (4a) *jA/ga* is used as a light verb, in (4b) it is a main verb. As *kar* 'having' can be dropped from the conjunctive clause, both sequences (verb + verb and verb + light verb) become form identical.

We find a similar ambiguity problems related to the light verb *dE* 'give'. Beside the light and main verb usages similar to (4a-b), the verb *dE* has another syntactic pattern. It introduces an additional dative or benefactive marked argument when it is used with certain verbs. The verb *xarId* 'buy' does not occur with the light verb *dE*. However *xarId* + *dE* has a dative marked beneficiary in the following example.

(5) a. *us=nE mujHE kitAb xarId-I
3SG=Erg 1SG.Da book win-Perf.F.Sg
'He bought me a book.'

b. us=nE mujHE kitAb
3SG=Erg 1SG.Dat book
xarId dI
buy give-Perf.F.Sg
'He bought me a book.'

Polysemy and identification of the correct/preferred sense is also concerned with the main verb. There are Urdu verbs that have more than one sense. Many of these senses are compatible with different light verbs. For example, the form *paRH* is used for both 'read/study' and 'read out' senses.

When *paRH* is followed by *dE*, it is used in 'read out' sense, as in (6b).

(6) a. us=nE xat paRH li-yA
3SG=Erg letter read take-Perf.M,Sg
'He read the letter.'

b. us=nE xat paRH di-yA
3SG=Erg letter read give-Perf.M.Sg
'He read out the letter.'

Moreover, most of the Urdu verbs have morphological causative counterparts. For example, *gir* 'fall' and *paRH* 'study' have causatives *gir-A* 'make fall' and *paRH-A* 'teach' respectively. However, there are some verbs like *badal* 'change' where the same form is used for both root and causative usages. The monovalent *badal* '(get) change' allows *jA* and rejects *dE*. On the other hand, divalent *badal* '(make) change' allows *dE* and rejects *jA*.

[1] In Urdu, noun + light verb and adjective + light verb are also used as complex predicate. However, we focus only on main verb + light verb sequences in this paper.

[2] The verb *jA* has the irregular form *ga* when used in perfective form. So the forms containing *ga* are the examples of the verb *jA*. Similarly, *dE/di/dI* 'give' and *lE/li/lI* 'take' are variants of the same form in other examples.

[3] There are other usages/senses of *jA* 'go' after the main verb. For example when *jA* comes after the perfective form of the verb, it is considered as a passive marker. However, the light verb *jA* that is used to represent completion is always used after the root form of the main verb. Hence it is ambiguous with the conjunction sequence only.

Similarly, all the light verbs are used only with specific form of the main verb preceding them.

In summary, we know that different light verbs are used with different kinds of verb, hence these can be used to identify classes of Urdu verbs.

3 Verb classes based on light verbs

Our analysis as to the interaction of main verbs and light verbs started with the shallow processing of a corpus. We collected data related to main verb + light verb combinations by processing a (raw) corpus consisting of seven thousand documents containing 14 million tokens. The documents were obtained from CRULP's (www.crulp.org) Urdu corpus and websites www.urduweb.org and www.kitaabghar.com.

The manual inspection of this data suggests patterns and verb classes related to different light verbs. However, the data has some noise/unwanted results because of polysemous verbs and light verbs, as explained above. Other reasons were homophonous/homographic words and data sparseness for some verbs.

The frequency data and the polysemy problems were the motivation for the manual identification of verb classes. The frequency analysis helped in finding the major patterns, but the final decisions were made on the basis of native speaker's judgments. These judgments are crosschecked by Google search.

The frequency analysis shows that the light verbs *jA* 'go', *lE* 'take' and *dE* 'give' occurred with 127, 97 and 95 main verbs respectively. (There were 184 high frequency main verbs in our analysis.) The fourth most frequent light verb was *A* 'come' that occurred with 48 (out of 184) main verbs. Hence, we used acceptability of frequently used *jA*, *lE* and *dE* in the analysis of verb classes.

In Table 1, we display verb classes on the basis of their acceptability/preference with the light verbs *jA*, *lE* and *dE*. An acceptable sequence is marked as '+', an unacceptable sequence is marked with '-' and a semantically odd combination is marked as '?'.

The classes listed in Table 1 do not cover all the verbs of Urdu. We find that most of the divalent/transitives do not show special syntactic patterns with respect to the light verbs *dE* and *lE*. They accept both *dE* and *lE*.

4 Discussion

Table 1 shows that we find different groups of verb classes on the basis of acceptability of light verbs *dE*, *lE* and *jA*. In the following discussion, we describe the semantic reasons of compatibility of the verb classes with these light verbs. We also discuss interesting verb classes found in this analysis.

Verb Class	**Valency**	***jA*** 'go'	***lE*** 'take'	***dE*** 'give'
Change of state *gir* 'fall', *kaT* '(get)	1	+	-	-
Ingestive *kHA* 'eat', *nigal* 'swallow'	2	+	+	-
Mental Gain *mAn* 'accept', *jAn* 'know'	2	+	+	-
Mental State *Dar* 'fear'	2	+	-	-
Perception *dEkH* 'see', *jHAnk* 'peep'	2,1	-	+	-
Grab *pakaR* 'grab', tHAm 'hold'	2	-	+	-
Send Away *pHEnK* 'throw', *bHEj* 'send'	2	-	-	+
Bodily expressions *hans* 'laugh', *cIx* 'scream'	1	-	?	+
Manner of Motion *ter* 'swim'	1	-	?	-
Manner of Displacement *uR* 'fly', *bHAg* 'run'	1	+	?	-
Sparkle *camak* 'shine', *mahak* 'smell (fragrantly)'	1	+	-	-

Table 1: Acceptability of some verb class and light verb sequences

In this analysis, we borrow the terms undergoer, resultee and rheme used by Ramchand (2008). However our analysis is not exactly similar to her analysis, therefore we use the terms undergoer', resultee' and rheme'. We consider the recipients to be a type of resultee' and the received entity as a type of rheme'.

4.1 Verb classes related to *jA* 'go'

The verbs compatible with *jA* 'go' are those whose subject is an undergoer' i.e. it undergoes a change. One example of *jA* compatible verbs is the monovalent change of state verbs like *kaT* '(get) cut'. The state of the subject of these verbs is changed.

The other classes of *jA* compatible verbs are more interesting. The ingestive and mental gain/state verbs are divalent. The peculiar behavior of ingestive verbs in causativization and their unusual event structure have already discussed in Saksena (1982), Butt (2006) and Ramchand (2008). According to Ramchand (2008), the subject of ingestives is an undergoer. For this reason, the ingestives allow the light verb *jA*.

The subject of mental gain and mental state verbs undergoes a change. Hence, these verbs also allow the light verb *jA*. The verbs that do not accept *jA* are the ones whose subject is not an undergoer'.

There are two other interesting verb classes that accept *jA*. For *sparkle* verbs, the meaning/sense of the verb is changed when these are used with *jA* 'go' light verb. If the verb *camak* 'shine' is used in a sentence without any light verb, it means that the subject shines. However, when it is used with *jA* then it means that the subject becomes shiny.

Table 1 has a class *manner of displacement* that is different from the *manner of motion* class. Traditionally, the verbs *uR* 'fly' and *bHAg* 'run' are considered as manner of motion verbs. However these verbs allow *jA* 'go' which is related to change of state. For this reason, we introduce a special class 'manner of displacement' for these verbs.

The identification of *jA* accepting verb classes enables us to disambiguate (or find preferred reading) for the ambiguous verb + *jA* sequences as discussed in section 2. For example, since the verb *kAT* 'cut' does not belong to a *jA* accepting class, the sequence *kAT gayA* in (4b) must be a conjunctive clause.

4.2 Verb classes related to *lE* 'take'

Almost all of the verbs allowing *lE* 'take' are divalent. This light verb comes with the verbs whose subject can be a receiver/endpoint of an action. In other words, the subject can be a resultee' having a rheme'.

Table 1 shows that there are three kinds of syntactic patterns with respect to *lE* 'take' and *dE* 'give' light verbs.

There are verbs that allow *lE* and disallow *dE*. These are the verbs whose subject gets something and acts as resultee' having rheme'. Table 1 shows that *grab*, perception, mental gain and ingestive verbs belong to this kind of verbs. These verbs are semantically similar.

The ingestive and mental gain verb classes allow both *lE* and *jA* 'go' light verbs. The subject of these verbs is/can be a resultee' as well as an undergoer'. Beside these, there are some other verbs that show the same pattern because of the same semantic reasons. The verb *jIt* 'win' behaves like ingestive and mental gain verbs (allow *lE* and *jA*). Similarly, *pahan* 'put on' behaves like grab verbs (allow *lE* only).

There are many verbs that allow both *lE* and *dE*. These verbs do not have any special requirement about receiving/giving of the subject. The subject can be a resultee' but it is not mandatory. These verbs can presumably be classified into finer classes on some other basis.

There are other verbs that do not allow *lE*. The subject of these verbs cannot act as a receiver or endpoint of a theme/result i.e. the subject cannot be a resultee'. The *send away* verbs are the example of these verbs that do not allow *lE*. As the subject of *bHEj* 'send' sends the object to some other place, it cannot be considered as the recipient or end point of the object that has been sent. Hence, the light verb *lE* that shows the reception/end point at the subject cannot be used with this verb.

The light verb *lE* distinguishes two different classes of mental or pysch verbs. As shown in table 1, mental gain verbs e.g., *samajH* 'understand' allow both *lE* 'take' and *jA* 'go'. When someone understands some fact, he/she goes through a change of mental state (hence *jA* is allowed) by gaining the fact (hence *lE* is allowed). However the verb *Dar* 'fear' behaves differently.

(7) a. vuh sANp=sE Dar ga-yA
3SG snake=Abl fear go-Perf.M.Sg
'He feared a/the snake.'

b. *us=nE sANp=sE Dar li-yA
3SG=Erg snake=Abl fear take-Perf.M.Sg
'He feared a/the snake.'

As the stimulus in (7a-b) i.e. *sANp* 'snake' is not gained by the subject, the verb *Dar* 'fear' cannot be classified as a mental gain verb. We classify it in a class that is different from the class of *samajH* 'understand' and *jAn* 'know' verbs. In Urdu, we have an independent evidence for this classification. The stimulus of *Dar* 'fear' is marked by the ablative marker. It shows that stimulus is a potential source and not the rheme of result. See Khan (2009) for more details.

4.3 Verb classes related to *dE* 'give'

The light verb *dE* 'give' is acceptable with the verbs that can have a receiver/endpoint that is different from the subject. It means the subject of these verbs cannot be a resultee' with rheme'.

The syntactic patterns of *dE* are similar to the patterns of *lE* 'take'. There are many divalent verbs that allow both *dE* and *lE*. These are the verbs whose subject is not necessarily a sender or receiver.

Beside these, there are verbs that allow *dE*, but does not allow *lE*. The *send away* verbs e.g. *bHEj* 'send' and *pHENk* 'throw' have a subject that cannot receive the theme i.e. it cannot be a resultee'. Hence, these verbs disallow *lE* and allow *dE.*

For a similar reason, *grab* verbs are not allowed with *dE*. The subject of these verbs is the receiver/endpoint and hence it is in conflict with the semantics of the light verb *dE*.

A similar observation for English light verb *give* was made in Newman (1996). He noted that *give*, in its extended meaning, is related to the emission. One can say *give a throw*, but *give a catch* is not acceptable. The reason is that the act of catching does not involve emission.

The monovalent verbs whose subject is not an undergoer' allow *dE* e.g., bodily expressions. However, the incompatibility of *dE* with manner of motion and displacement verbs e.g. *ter* 'swim' and *uR* 'fly' needs explanation in future work.

The verb classes related to *dE* help us in the disambiguation of some polysemous verbs. As described in section 2 and examples (6a-b), the verb *paRH* has two different senses. When *paRH* is used in 'read/study' sense, it acts as an ingestive verb and disallows *dE*. The other sense of *paRH* i.e. 'read out' is somewhat similar to bodily expression verbs, and hence it allows *dE.* Therefore, if we find a sequence of *paRH* and *dE*, we will consider it as an instance of 'read out' sense.

4.4 Verb classes related to *A* 'come'

Although we did not consider the light verb *A* 'come' as part of our analysis in Table 1, a consideration of its patterns of use bring out another interesting point. Rather than being sensitive to event structure components such as resultee', rheme' and undergoer', it provides a sense of directionality of the action.

The verbs which accepts *A* 'come' turn out to be a subset of the ones which accept *jA* 'go'. However, only those verbs which have inherent potential directionality in their lexical semantics can be used with *A*. For example, the verbs *nikal* 'emerge', *ug* 'grow' and *baRH* 'increase' are related to direction.

(8)	cAnd	nikal	ga-yA/A-yA
	moon.M.Sg	emerge	go-Perf/come-Perf
	'The moon emerged.'		

The other direction-less change of state verbs e.g. kaT '(get) cut' does not allow *A* 'come' light verb.

(9)	daraxt	kaT	gayA/*A-yA
	tree.M.Sg	cut	go-Perf/come-Perf
	'The tree got cut.'		

5 Conclusion:

The study presented some classes of Urdu verbs on the basis of allowing combinations with the light verbs *jA* 'go', *lE* 'take' and *dE* 'give'. The identified classes show that light verbs are related to specific semantic classes. This work is an important step towards identifying Urdu specific (Levin-style) alternations that can give a comprehensive list of Urdu verb classes.

In this study we found that ingestive, mental gain and perception verbs behave similarly. Moreover, we found that mental/pysch verbs can be classified into (at least) two classes on the basis of light verb acceptability. The verbs of these two classes (mental gain and mental state) are semantically different form each other. Hence, the syntactic difference correctly determined the difference in semantics. Similarly, we identified two classes of manner of motion verbs.

The study needs further refinement especially in terms of semantic constructs explaining verb classes. However, the classes presented in Table 1 and the rough sketch of the semantic model that enable us to understand the problem and future directions for a complete solution.

Acknowledgment:

This research was carried out at Urdu ParGram project supported by the DFG (Deutsche Forschungsgemeinschaft).

References:

Abul Lais Siddiqui. 1971. *Jamaul Qawaid (Comprehensive Grammar)*. Karachi.

Alice Davison. 1999. Ergativity: Functional and Formal Issues. In Darnell et al. (eds.) *Functionalism and Formalism in Linguistics, Volume I: General Papers*, 177-208. John Benjamins.

Anuradaha Saksena. 1982. *Topics in the Analysis of Causatives with an Account of Hindi Paradigms*. Berkeley: University of California Press.

Beth Levin. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. , Chicago: University of Chicago Press.

Gillian Ramchand. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*, Cambridge: Cambridge University Press.

John Newman. 1996. *Give: A Cognitive Linguistic Study*. Berlin: Mouton de Gruyter.

Miriam Butt. 1995. *The Structure of Complex Predicates in Urdu*. Stanford: CSLI Publications.

Miriam Butt. 2006. *Theories of Case*. Cambridge: Cambridge University Press.

Miriam Butt and Wilhelm Geuder. 2001. On the (semi)lexical status of light verbs. In *Semi-lexical categories: the function of content words and the content of function words*. Berlin: Mouton de Gruyter.

Peter Edwin Hook. 1974. *The compound verb in Hindi*. Ann Arbor: University of Michigan.

R. S. McGregor. 1972. *Outline of Hindi grammar.* Oxford: Oxford University Press.

Ruth Laila Schimdt. 1999. *Urdu: An Essential Grammar*, London: Routledge.

Tafseer Ahmed. 2010. Clustering Urdu verbs on the basis of related aspectual auxiliaries and light verbs. In *Proceedings of KONVENS (Conference on Natural Language Processing) 2010*. Saarbrücken.

Tafseer Ahmed Khan. 2009. *Spatial Expressions and Case in South Asian languages*. PhD dissertation. Konstanz.

Telicity and event culmination in Hindi perfectives

Sudha Arunachalam
Northwestern University
Evanston, IL, USA
s-arunachalam@northwestern.edu

Anubha Kothari
Stanford University
Stanford, CA, USA
anubha@stanford.edu

Abstract

We report the results of an experimental study of Hindi speakers' judgments of telic perfective predicates describing events that either do or do not come to completion. We discuss the results in terms of a semantic vs. pragmatic treatment of telicity, as well as cross-linguistic differences in verb representation.

1 Introduction

The proper treatment of telicity has long been debated. Whether or not a predicate is telic apparently depends on a combination of factors. One important factor is the presence or absence of a feature (quantity) on the verb's complement. But while this may be a necessary condition (cf. *She walked to the store in an hour*), it is not sufficient: in, e.g., *push the cart*, the object is quantized but the predicate is nevertheless atelic. Other properties, such as whether the predicate involves a process component, are also relevant.

But context is also important in the calculation of telicity. Folli & Harley (2006), for example, note the contrast between (1a-b):

(1) a. John lengthened a rope
(*in 2 minutes / for 2 minutes).
b. The tailor lengthened a pair of trousers
(in 2 minutes / for 2 minutes).

See also Borer (2005) for examples in which telicity results not from reaching a natural endpoint, but rather meeting a certain threshold.

These facts raise the question of how to understand the interplay between featural properties of the predicate (e.g., Vendler classes, quantization of object) and contextual/pragmatic factors for calculating telicity.

Adding to the complexity is a related puzzle, the one we pursue in the current study. This is the phenomenon whereby languages differ in whether telicity seems to co-occur with completion of the event. In a range of languages including Japanese (Ikegami, 1985), Tamil (Pederson, 2007), and Hindi (Kothari, 2008; Singh, 1998), a verb does not entail completion of the event it describes. In the Hindi sentence (2), for example, the verb appears in the perfective, but the event can end at some arbitrary endpoint before the apple is completely eaten. This is true even though (a) the verb's complement is quantized, and (b) the event is one, unlike *push the cart*, which has a natural endpoint (i.e. when the apple is completely eaten). (The English counterpart is infelicitous.) Note, however, that the default interpretation, without the cancellation, is still that the event has arrived at its natural endpoint.

(2) Maya-ne biskuT-ko khaa-yaa
(par use puuraa nahiin khaa-yaa)
Maya-ERG cookie-ACC eat-PERF
but it-ACC full not eat-PERF
Maya ate a cookie #(but not completely)

In (3), a light verb appears on the main verb; it is infelicitous unless the event ends at its natural endpoint.

(3) Maya-ne biskuT-ko khaa-li-yaa
#(par use puuraa nahiin khaa-yaa)
Maya-ERG cookie-ACC eat-take-PERF
but it-ACC full not eat-PERF
Maya ate a cookie #(but not completely)

One account for these facts would be to say that simple verb (SV) predicates as in (2) are atelic, and that a telicity feature is contributed by the light verb. Standard telicity tests asking whether SV predicates are atelic show mixed results, but do support this hypothesis. But if SV predicates are atelic, it remains to be explained why the default interpretation—if event culmination is not explicitly cancelled—involves full completion. Event completion appears to be implicated, but not entailed, by the SV, and entailed by the complex verb construction (CV) in (3).

This phenomenon raises important questions. First, if the difference between the Hindi SV and CV is one of implication vs. entailment, rather than, e.g., the presence/absence of a quantity feature on the verb's complement, is a semantic (rather than pragmatic) treatment necessary?

Second, is there a parametric difference between Hindi-type and English-type languages? Syntactically, of course, English uses the SV sentence type for describing eventualities like these, but the unavailability of cancellation of event culmination suggests a meaning more like the Hindi CV. Are we to say that verbs in Hindi have a different meaning from their translation-equivalent English counterparts? (Ikegami (1985), for example, proposes that an English accomplishment or achievement is interpreted more like an activity in Japanese.)

The only full treatment of this phenomenon in Hindi that we are aware of is from Singh (1998). Singh posits a new thematic relation relating the event and the affected object, couched in a homomorphism approach; the difference between the SV and CV constructions lies in how much of the "theme" object is affected. This account makes several predictions, among them: (1) Only accomplishment predicates with incremental themes should show the SV-CV distinction, (2) *All* accomplishment predicates with incremental themes should show the distinction.

To test these predictions, we undertook an experimental study of Hindi speakers' interpretations of predicates that in English are construed as telic and entail completion of their endstates. We included accomplishment predicates with incremental themes, as well as achievement predicates, in a variety of contexts.

An experimental method served two functions. First, the judgments in question are often subtle, and experiments allow us to obtain a large number of judgments without speakers being aware of our theoretical interests. Second, the experimental method allowed us to carefully control the real-world context surrounding the events, such that only the relevant variables (whether events completed, and SV vs. CV syntax) varied.

2 Experimental Study

We showed Hindi speakers video clips of actions that either fully completed (e.g., woman eating a cookie), or partially completed (e.g., woman eating most of a cookie). At the conclusion of each video clip participants heard an SV or CV sentence describing the video and were asked to provide a true/false judgment.

Methods

Participants. Twenty-four adults participated.

Materials. For each of 8 predicates, we filmed pairs of short video clips. One video of each pair depicted a fully-completed event and the other depicted a partially-completed event.

At the end of each clip, participants heard a recording of a native speaker describing the event. Participants heard either an SV sentence (e.g., *us-ne biskuT-ko* ***khaa-yaa***, "She ate the cookie"), or a CV sentence (e.g., *us-ne biskuT-ko* ***khaa li-yaa***). They were asked to give a true/false judgment as to whether the sentence described the event they had viewed.

Predictions

We predicted that if Hindi speakers are sensitive to the SV-CV distinction, participants would show different responses for partially-completed events depending on syntactic condition. Because CV sentences entail completion of the event they describe, we expected 0% acceptance of CVs as descriptions of partially-completed events. SV sentences were expected to have a high acceptance rate, though perhaps not 100%, given that the default interpretation for SVs is still full completion. Because both SV and CV sentences are felicitous descriptions of fully-completed events, we expected 100% acceptance, regardless of syntactic condition.

We made further predictions about the range of predicates to which the SV-CV distinction should apply. If partial completion interpretations arise via a homomorphism between the measuring out of the event and the theme object, then only accomplishments with incremental themes (*cover*, *draw*, *eat*, *fill*) should show the pattern. For all other predicates, both SVs and CVs should only be acceptable for fully-completed events, receiving an acceptance rate of 0% for partially-completed events.

Results and Discussion

These predictions partially held. For fully-completed events, participants accepted both SV and CV sentences (99.5%). For partially-completed events, participants' responses differed by syntactic condition, with a higher acceptance rate for SV sentences (53%) than CV sentences (29%). An ANOVA on participant means revealed main effects of Event Completion ($F(1,23) = 134.1$, $p < .001$), and Syntax ($F(1,23) = 9.6$, $p < .01$), and a significant interaction (F(1,

23) = 11.3, $p < .005$). The same effects are evident in an analysis on predicate means instead of participant means (Event Completion: $F(1, 7) = 83.9$, $p < .001$); Syntax: $F(1, 7) = 5.8$, $p < .05$), Interaction: $F(1, 7) = 7.4$, $p < .05$).

These results support the distinction described in the literature whereby SV sentences can describe events with arbitrary endpoints, while CV sentences can only describe events that reach their natural endpoints.

But our predictions about the range of predicates which should show this distinction did not entirely hold. Of the four canonical incremental theme predicates (*cover*, *draw*, *eat*, *fill*), all but *draw* showed the pattern in the expected direction. *Draw* (*a circle* / *a flower*) showed no difference between the two syntactic conditions, although in both conditions acceptance rates were relatively high (40%), suggesting that partial completion interpretations *are* available.

For the achievement predicates (*extinguish*, and *pluck*), SV sentences were accepted more often than CV sentences as descriptions of partially-completed events, though the differences are not statistically significant. However, all three trials had very low acceptance rates, even for SV sentences. *Pluck*, for example, yielded just a 17% acceptance rate for SV sentences describing partially-completed events, suggesting that most speakers require a plucking event to be fully-completed to be describable with this predicate, regardless of syntactic condition.

For *wake* and *extinguish*, there were trials in which, for partially-completed events, the event reached its natural endpoint, but then retracted to its initial state. For example, in the *wake* trials, the partially-completed video showed a man jostled into some state of wakefulness, slightly opening his eyes, but quickly closing them again and returning to an apparent sleep state. For both of these predicates, the SV-CV distinction manifested, in the predicted direction. This is contrary to our prediction that only predicates with a process component and/or incremental theme should show the SV-CV distinction.

This finding lends further support to the idea that the conditions for the SV-CV distinction are heavily context-based, and not dependent on the amount of the theme object which has been affected, nor on how much of the process has been achieved, but rather on a perception of whether the action has been functionally completed. For CVs, the object must be in the relevant endstate at the time of evaluation (here, when the video ends and the sentence is uttered), even if the endstate was achieved at some point.

This striking result requires a rethinking of the importance of quantization, process components, incremental themes, and other features with respect to the SV-CV distinction, and has consequences for our understanding of telicity.

3 Conclusions

The data confirm that while SV and CV perfectives are equally compatible with natural endpoints, they differ with regards to their relative compatibility with arbitrary endpoints. SV perfectives can be used to describe events with arbitrary endpoints, while CVs cannot. However, the distinction is graded rather than categorical, with SVs only accepted half the time. Our results also show that the SV-CV distinction is not limited to events that involve an incremental theme. Predicates like ‘wake up’, for example, showed the expected SV-CV difference in the Partial condition, even though it is an achievement and in fact involves no incremental theme. Rather, whether or not an event arrived at its intended, pragmatically-determined ending point appears to underlie the SV-CV distinction (Kothari, 2008).

Within and across languages, context mediates interpretation of event completion. Of course, the grammatical distinction between SV and CV contributes to interpretation as well. But what role is there for a semantic/featural approach to telicity? These results support a pragmatic approach in at least two areas. First, we have demonstrated that incremental affectedness of a theme object is not the primary semantic criterion affecting interpretation; a semantic homomorphism treatment is not the whole story. Second, the relatively low acceptance of SVs for partially-completed events supports our hypothesis that the default interpretation is one of full completion. This can be explained pragmatically as well. Because full completion (telic) interpretations *entail* partial completion interpretations, the full completion interpretation is stronger, and therefore speakers may prefer it (acting on Gricean quantity) unless context strongly drives a partial completion interpretation.

The pragmatic approach provides a clear reason why languages like English and languages like Hindi should differ; because Hindi has the syntactic availability of the CV construction, the SV takes over a different function. Pederson (2007) argues that while English has a number of ways to express incompleteness (e.g., *almost*, *halfway*), this is not universal; languages may use other devices, here the CV, to achieve this semantic function. Translation, then, need not be radical; Hindi and English verbs pick out similar concepts, but the availability of different linguistic and pragmatic factors conspire to make event completion more or less strongly implicated.

References

Ikegami, Y. 1985. 'Activity'—'Accomplishment'—'Achievement'—A language that can't Say 'I burned it but it did not burn' and one that can. In: A. Makkai and A. K. Melby, eds., *Linguistics and Philosophy: Essays in Honor of Rulon S. Wells*, 265-304. Amsterdam: John Benjamins.

Kothari, Anubha. 2008. Event Culmination as Implicature in Hindi Perfectives. Ms., Stanford University.

Raffaela Folli, and Heidi Harley. 2006. On the licensing of causatives of directed motion: Waltzing Matilda all over. *Studia Linguistica*, 60:1-35.

Pederson, Eric. 2007. Event realization in Tamil. P. Brown and M. Bowerman, eds., *Crosslinguistic Perspectives on Argument Structure: Implications for Learnability*, 331-355. Lawrence Erlbaum, New York City, NY.

Singh, Mona. 1998. On the semantics of the perfective Aspect. *Natural Language Semantics*, 6:171-199.

The motion component is preserved in metaphorical sentences. A TMS study.

Cristina Cacciari
University of Modena, Italy
cacciari.cristina@unimore.it
Maria Concetta Pellicciari
IRCCS San Giovanni di Dio Fatebenefratelli, Brescia, Italy
pellicciari@med.unibs.it
Arianna Fogliata
University of Milano-Bicocca, Italy
arianna.fogliata@unimib.it
Costanza Papagno
University of Milano-Bicocca, Italy
costanza.papagno@unimib.it
Carlo Miniussi
University of Brescia, Italy
miniussi@med.unibs.it

Abstract

We used Transcranial magnetic stimulation (TMS) to assess whether reading literal and non-literal (i.e., fictive, metaphorical, idiomatic) motion sentences modulates the activity of the motor system. In Experiment 1, TMS was delivered immediately after the verb when participants were not yet aware of the literal or figurative nature of the conveyed motion. These sentence fragments elicited a significant change in the MEPs area only when the agent was animate. In Experiment 2, TMS was delivered at the end of the sentence. The MEP response was larger when participants were presented with metaphorical than with idiomatic or fictive motion sentences. These results suggest that the excitability of the motor system is modulated by: a) the animate vs. inanimate nature of the sentential subject, and b) the motor component of the verb that is preserved in metaphorical motion sentences. We showed that the activation of motor representations is influenced by the linguistic context and only appears when the use of the verb preserves the basic semantic components of the verb meaning.

1 Introduction

In recent years, the hypothesis that the neural circuitries associated with action are recruited when processing action-related words and sentences has opposed traditional *amodal/disembodied* models of conceptual knowledge to *embodied* models. The first models posit that conceptual knowledge is fundamentally amodal and abstract and represented separately from modality-specific systems recruited for perception and actions. In contrast, *embodied* models of cognition posit that sensory-motor processes are a fundamental part of the mental representation of abstract and concrete concepts. The neural architecture of language-induced *motor resonance* would therefore comprise regions encoding information that is not purely linguistic or conceptual but reflects the sensory-motor properties associated with the underlying concept.

2. Aim of the study

The relationships between language and action has been investigated by an impressive amount of studies employing behavioral, neuropsychological and neuroscientific methodologies (for overviews, see Fischer & Zwaan, 2008; Glenberg et al., 2008; Vigliocco et al., 2004). Notwithstanding, important problems remain (Mahon & Caramazza, 2008). One of them concerns the extent to which motor areas are activated during figurative sentence processing. When someone says *The road turned left suddenly*, it is evident that she or he does not refer to a physical entity moving (this expresses a *fictive motion*, Talmy 2000, see below). These properties are instead implied in literal sentences as *The man turned left suddenly*. But what happens when the verb *turn* is used in a metaphorical context as *The lady turned her thought away from sorrow*? We hypothesized that the activation of the motor system reflects how much the motion component of the verb meaning is preserved. Following the claim of a behavioral study on metaphorical verb meaning (Torreano et al., 2005), we hypothesized that metaphorical sentences might preserve the motion component. Specifically, we assumed that in metaphorical sentences the semantic component of a verb is abstracted out and employed to predicate a type of movement of whatever subject can change direction, regardless from its literal or figurative nature. We verified this claim in two experiments that employed a Transcranial Magnetic Stimulation (TMS) protocol with literal and figurative sentences that differed in the extent to which the motion component of the verb was preserved.

3. Method

Participants. Eight and twelve right-handed Italian participants were enrolled in the first and second experiment, respectively**.**
Materials. We selected twenty-seven common Italian verbs expressing a movement that involved the legs (e.g., *follow*, *cross*, *run*) and created four types of sentence for each verb: 1. Literal sentences (e.g., *The policeman follows the thief*); 2. Metaphorical sentences (e.g., *The girl follows her instinct always*); 3. Idiomatic sentences (e.g., *Giuseppe follows the footsteps of his father*); 4. Fictive motion sentences (e.g., *The railway follows the stream of the river*). Twenty-seven sentences of similar length and syntactic structure containing a mental verb were created as control sentences (e.g., *Cristina considers the idea very interesting*).
Procedure. We recorded motor evoked potentials (MEPs) from right inferior limb muscles while delivering single-pulse TMS on the left primary motor cortex. Variations of the motor cortex excitability indexed by MEPs provided a measure of the involvement of the motor system. The sentences were divided into three segments (the noun phrase, the verb and the final part of the sentence) presented on the screen one at a time. The participants' task was to read for comprehension. In the first experiment, the TMS pulse was delivered immediately after the verb (*The policeman follows*, *The railway follows*) namely when participants were not yet aware of the literal or figurative nature of the full sentence. The rationale was to verify whether the mere presence of a motion verb activated motion areas, regardless of the animate or inanimate nature of the agent. In the second experiment, TMS was delivered at the end of the full sentence. Readers were presented with figurative motion sentences (idiomatic and metaphoric), fictive motion and mental (control) sentences. We did not include literal sentences since it was clear from the results of the first experiment and from previous studies (e.g., Oliveri et al., 2004; Buccino et al., 2005) that literal motion sentences indeed activated the motor cortex. In both experiments, the effect of the sentence types on motor cortical excitability was evaluated by means of MEP changes expressed in terms of the ratio (Δ) between motion and mental sentences.

4. Results

Motion sentence fragments significantly modulated the MEPs evoked in the GCM muscle but only when the sentential subject was animate (Experiment 1) [animate motion fragments vs. inanimate motion fragments: $t(7) = -2.76$; $p = .03$]. When idiomatic, metaphorical, fictive motion sentences and mental sentences were presented in their full form (Experiment 2), the highest motor cortical excitability occurred in metaphorical motion sentences [Sentence Type factor: $F(2, 18) = 3.92$, $p < .04$; pairwise comparisons: metaphorical vs. idiomatic motion sentences $p < .036$; idiomatic vs. fictive motion sentences and fictive vs. metaphorical sentences: n.s.]. Fictive motion sentences triggered very low motor excitability, and even less so idiomatic sentences. In sum, we found that language-induced motor resonance was largest in metaphorical motion sentences than in fictive and idiomatic motion sentences.

5. Conclusion

The aim of this study was to determine the impact of literal and non-literal motion sentences on motor excitability as reflected by MEP changes during TMS stimulation. The high motor excitability induced by metaphorical sentences is consistent with the behavioral claim that the metaphorical use of a verb preserves the basic semantic components of the verb meaning. The difference between metaphorical and literal motion sentences lies in the fact that in the metaphorical sentence the motion verb did not take its default arguments, for instance a physical entity. The level of abstractness of the motion component conveyed by literal and metaphorical sentences differs since in metaphorical sentences the motion verb is used at higher level of abstraction to refer to any instance of goal-driven conjoint motion.

Differently from metaphorical motion sentences, our results showed that the motion component of the verb was almost lost when it was embedded in idiomatic sentences. Why metaphor and idiom differ in the extent to which their meaning can resonate with the motor system? We believe that this is due to the different structure of these figurative expressions: in fact the relationship between an idiom's constituent words and the idiomatic meaning generally is arbitrary and learned and the idiomatic meaning overlearned (Azizh-Zadeh et al., 2006; Boulenger et al., 2009). Idioms typically convey abstract meanings and not concrete motor acts. Even though many idioms originate from metaphors, this origin can be totally unperceived by readers. Lastly, it should be mentioned that the absence of activity of the motor system in fictive sentences contrasts with what was found in some previous studies (e.g., Wallentin et al., 2005). However, we believe that this lack of modulation might depend on the inanimate nature of the agent typical of fictive motion sentences. In fact, as the results of Experiment 1 showed, the motor system did not activate when the action agent was inanimate as in *The railway follows the stream of the river*, for instance.

In sum, our findings indicate that the semantic representations grounded in the sensory-motor system indeed play a role in processing sentential meaning. However, the activation of motor representations is strongly influenced by the linguistic context and only appears when the verb preserves its basic semantic components, as in literal and metaphorical sentences.

References

Aziz-Zadeh, L., Wilson, S.M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, *16*, 1818-1823.

Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping the ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebral Cortex, 19*, 1905-1914.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, *24*, 355-363.

Fischer, M. H., & Zwaan, R. A. ((2008). Embodied language: A reviw of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*, *61*, 825-850.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palombo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *The Quarterly Journal of Experimental Psychology*, *61*, 905-919

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris, 102*, 59-70.

Oliveri, M., Finocchiaro, C., Shapiro, K., Gangitano, M., Caramazza, A., & Pascual-Leone, A.(2004). All talk and no action: A Transcranial Magnetic Stimulation study of motor cortex activation during action word production. *Journal of Cognitive Neuroscience, 16*, 374-381.

Talmy, L. (2000). *Toward a cognitive semantics* (2 voll). Cambridge, MA: MIT Press.

Torreano, L, Cacciari, C., & Glucksberg, S. (2005). When dogs can fly: Level of abstraction as a cue to metaphorical use of verbs. *Metaphor & Symbol*, 20, 259-274.

Vigliocco, G., Vinson, D. P, Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words. The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*, 422-488.

Wallentin, M., Ostergaard, S., Lund, T.E., Østergaard, S., Østergaard, L., & Roepstorff, A. (2005). Concrete spatial language: See what I mean? *Brain and Language, 92*, 221-233.

Force dynamics in verbal semantics: Verbs of maintaining

Bridget Copley
Structures Formelles du Langage
CNRS / Université Paris 8
Paris, France
bridget.copley@sfl.cnrs.fr

Heidi Harley
Department of Linguistics
University of Arizona
Tucson, Arizona, USA
hharley@email.arizona.edu

A long-standing puzzle in the event-structure literature concerns the status of "maintaining" predicates like *keep* and *stay* (Jackendoff, 1972). They are clearly eventive, as diagnosed by the usual eventuality tests; for example, the progressive gets an 'ongoing-now' reading, as in *John is keeping the door open* or The door is staying open, while present tense is interpreted as habitual, as in *John keeps the door open* or *The door stays open.* However, it is difficult at first glance to understand what distinguishes these predicates from stative predicates such as *The door is open* and *John has the door open.* Both kinds of predicates refer to a situation in which the door's being open endures over time, so there is no obvious formal rationale for their distinct Aktionsart types. Yet at the same time, there is an intuition that the dynamicity of maintaining predicates is not an accident, as there is energy being put into the situation. It is not immediately obvious how to formally characterize a dynamic eventuality in which energy is put into the situation but nothing changes.

Within the framework of Montagovian formal semantics, we can imagine several neo-Davidsonian analyses for *keep*, but they are all unsatisfactory. We assume that *keep* and *stay* take a small clause complement p; in the case we are examining, p would be [the door open]). The problem with (1a) ("cause to be") and (1b) ("cause to become") is that it is possible to keep something in a location without strictly being the cause of its being there or coming to be there. On the other hand, *keep* might instead be "cause to stay," as in (1c). But for *stay* we run out of options: there can be no external argument or causing event, and there is no obvious way to combine the caused event e_2 and the small clause predicate p in such a way as to reflect the fact that *stay* is not the same as *be*.

(1) a. keep =? $\textsc{Agent}(x,e_1)$ & e_1 $\textsc{Cause}$ e_2 & $\textsc{Be}(e_2, p)$
b. keep =? $\textsc{Agent}(x,e_1)$ & e_1 $\textsc{Cause}$ e_2 & $\textsc{Become}(e_2, p)$
c. keep =? $\textsc{Agent}(x,e_1)$ & e_1 $\textsc{Cause}$ e_2 & $\mathrm{stay}(e_2,p)$
d. stay $\neq$ $\textsc{Be}(e_2, p)$

What is needed is some way to represent the idea that maintaining events involve the input of energy into a situation. Other kinds of events, of course, should also involve the input of energy; however, with these other kinds of events, the input of energy results in a different situation from the initial situation, while with maintaining events, the result of inputting energy is the same as the initial situation.

We propose to alter the neo-Davidsonian framework to view events—intuitively speaking—as inputs of energy into situations (Talmy, 1988, 2000), and—formally speaking—as forces that are functions from one situation to another, where the latter situation is the one that results provided that no other force intervenes. A situation s is a collection of individuals and their properties, a notion compatible with DRT-like theories (Kamp and Reyle, 1993) but also compatible with treatments of situations as partial worlds (Barwise and Perry, 1983; Kratzer, 1989). A force is a function f from situations to situations; i.e., it is type $\langle s,s \rangle$, which we will abbreviate as type f. The theorem in (2a) connects forces with situations, the definition of successor in (2b) links them into causal chains, and the terminology introduced in (2c) allows us to recover initial and final situations from a force.

(2) a. For any situation s_n, there is a

force f_n such that f_n is the net force of s_n.

b. For any situation s_n, its successor s_{n+1} is defined as $f_n(s_n)$.

c. For any force f_n which is a net force of a situation s_n, init(f)=: s_n and fin(f) =: s_{n+1}.

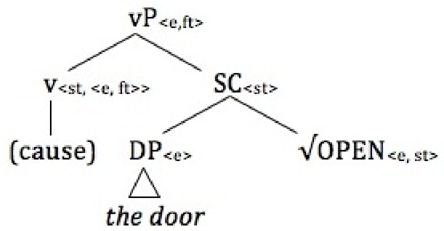
By the formal object we call a "force," we mean to include not just contact forces that result in a change in the spatiotemporal properties of an object, i.e., where it is, whether it is moving or at rest, etc. In these cases, the situations init(f) and fin(f) differ only in the the spatiotemporal properties of objects. But in fact, any change could be represented abstractly as a function from one situation to another.[1] One robust category of such abstract forces is the category of what we may think of as "psychological forces." For example, just as we can speak of pushing or putting pressure on an object, we can also speak of pushing or putting pressure on someone, in a psychological sense, to accept an idea or to do an action. The idea that the conception of the physical world is co-opted for use in the psychological or psychosocial domain is present in Jackendoff (1987) and Lakoff and Johnson (1999), among many others (see, e.g., Bloom et al. (1999) for a representative sample). Talmy (1988, 2000) has extensively championed this view that force dynamics is the way to understand this link between the physical and the psychological. For example, while the sentence in (3a) (Talmy, 2000, (vol 1): 412) is "force-dynamically neutral," the sentence in (3b) conveys that some other force, whether physical or psychosocial, prevents John from going out of the house if he wants to.

(3) a. John doesn't go out of the house.
b. John can't go out of the house.

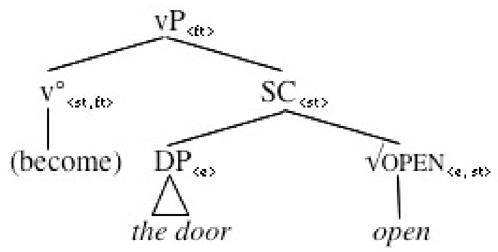
Wolff (2007), for one, has tested this idea experimentally, showing subjects a scene in which a pedestrian wants to go in a certain direction and a policeman directs her to go in a certain (possibly different) direction, and asking if the policeman *caused* the pedestrian to reach her destination, *helped* her reach her destination, or *prevented* her from reaching her destination. The results exactly parallel the results he obtains with inanimate objects exerting forces on each other. Based on such findings, it is not controversial to treat even non-spatiotemporal events as forces. So John can keep the door closed by pressing it closed, or he can keep Mary home by forbidding her to go out (and having the authority to make sure she obeys); in either case, he will be applying a maintaining force, whether physical or psychological in origin.[2]

To analyze *keep* and *stay* under a force-based framework, we first present logical forms for *cause* and *become*. *Cause* introduces an external argument (with a "source" role, similar to an agent role), while *become* does not. The initial situation is one where the small clause predicate p does not hold, and the final situation is one where p does hold (Jackendoff, 1972; Dowty, 1979; Pustejovsky, 1991). The event of someone opening the door (for *cause*) or the door opening (for *become*) is represented by the force f that effects the transition from door-not-open to door-open.

(4) $[\![\text{cause}]\!] = \lambda p\ \lambda x\ \lambda f\ .\ \neg p(\text{init}(f))\ \&\ p(\text{fin}(f))\ \&\ \text{source}(x,f)$



(5) $[\![\text{become}]\!] = \lambda p\ \lambda f\ .\ \neg p(\text{init}(f))\ \&\ p(\text{fin}(f))$



[1]This abstraction is already present in Aristotle's *Physics* (V:1), although he does not extend the analysis to verbs of creation and destruction.

[2]We recognize that the notion of a psychological force is more complex than the notion of a physical force, in that it involves a belief of the entity desiring the outcome. We address the intensional character of psychological forces in Copley and Harley (2010).

For *keep* and *stay*, the intuition is that energy must be added to an initial situation to maintain identity between it and the final situation. This is true when the net force of an initial situation would, without the additional maintaining force, normally produce a transition to a different final situation. (Stative predicates such as *The door is open* characterize situations with a zero net force, where no energy need be inputted to maintain the situation). The predicates *keep* and *stay*, then, are very similar to the predicates *cause* and *become*. Both take a predicate-of-situations (that is, type $\langle$ s, t$\rangle$) complement. They require that this type $\langle$s, t$\rangle$ complement be true of both the initial situation and the final situation. *Keep* and *stay* will also be differentiated in the same way as *cause* and *become* in that *keep* introduces an external argument and *stay* does not. Thus, the logical forms of *keep* and *stay* are as follows:[3]

(6) a. $[\![\text{keep}]\!] = \lambda\text{p}\ \lambda\text{x}\ \lambda\text{f}$. p(init(f)) and p(fin(f)) and source(x,f)
b. $[\![\text{stay}]\!] = \lambda\text{p}\ \lambda\text{f}$. p(init(f)) and p(fin(f))

We assume that when *keep* or *stay* takes another eventive predicate as its complement, as in *John kept Bill running around all day*, the aspect represented by *-ing* has applied to map the type $\langle$f,t$\rangle$ constitutent [run around all day] to an appropriate $\langle$s,t$\rangle$ predicate. (This suggests, perhaps, that *Bill kept running around all day* involves a control structure with a PRO subject of the gerund in the lower constituent.) Predicates like *endure* and *preserve* consist of the *stay* and *keep* functions with null existence predicates in their complements.

Two issues that arise with verbs of maintaining deserve further attention. Firstly, certain uses of verbs of maintaining seem to involve "maintenance by prevention" as in the example in (7) (due to an anonymous reviewer):

(7) The cattle grid kept the road clear of animals.

In this case, the cattle grid prevents the animals' actions that would normally cause the road to not be clear of animals. We propose an analysis inspired by Wolff's (to appear) force-dynamic analysis of "causation by omission", in which the force-dynamic configurations for *A prevents B* and *B prevents C*, taken together, result in the force-dynamic configuration for *A causes C*.

The second issue that arises has to do with cases that seem to involve not physical forces, but behavior that is out of the ordinary for the agent. For example, (8) can indeed be uttered when John is not physically preventing the door from closing:

(8) John is keeping his door open.

In that case, however, the hearer accommodates the idea that John does not typically have his door open. We suggest that the force being opposed in (8) is a force of John's typical tendency to close his door. This tendency can be compared to the tendency of an object to fall in the gravitational field of the earth; cf. the Aristotelian explanation (Physics, VIII:4) for gravity, in which heavy things (earth, etc.) have a tendency to descend, while light things (smoke, fire) have a tendency to ascend; Talmy (2000) as well uses this notion of tendency to understand forces. Just as the force of gravity on an object can be understood as the object's tendency to fall, so can John's tendency to close his door be understood as a force on John.

We will discuss some further implications of this proposal for argument structure, including for activities. Activity predicates have no associated result state; we treat such predicates (*sing,* etc.) as pure predicates of forces. Within this framework, their special ability to function as manner predicates in Accomplishment constructions such as *John whistled his way to the store* is unsurprising.

To the extent that this proposal captures the argument structure of verbs other than verbs of maintaining, but also captures other verbs such as *cause*, *become*, and Activity verbs, the import of this proposal goes beyond merely accounting for a backwater of ver-

[3]A reviewer points out that there is a grammatical difference between *keep* and *cause*, namely that the former takes a bare VP (e.g., *John kept the door open*) and the latter takes an infinitival clause (e.g., *John caused the door to open*). We believe this difference to be orthogonal to the difference between *keep* and *cause*, as a verb very similar to *cause*, namely *make*, also takes a bare VP complement (e.g., *John made the door open*).

bal semantics. The understanding of events as forces could clarify the interface with the cognitive system, since its ontology–situations as arrangements of individuals with the forces on them—may be preferable to that of the event-based framework with its concatenated events that somehow cause one another. It should also be preferable to treatments of situations as partial worlds (Barwise and Perry, 1983; Kratzer, 1989), since it is not at all clear how to make cognitively plausible sense out of possible worlds.

Another advantage of our approach has to do with how arguments of the verb are composed in syntax. The particulars of the force-situation framework suggest that it is more straightforwardly compositional than is the event-based framework. In the latter, the constituents denoting subevents are related to each other by means of a stipulated "CAUSE" interpretive relation, imposed when a type mismatch is detected between the event-denoting subparts of the vP. In the force-situation framework, however, all components of the vP are composed via function application, just as other nodes in the structure are; the lower VP in John opened the door, for example, which we take to be a small clause [the door open] with denotation $\langle s,t \rangle$, is selected by a v^0 head of type $\langle \langle s, t \rangle, \langle e, \langle f,t \rangle \rangle \rangle$, such that the $\langle s, t \rangle$ predicate denoted by the VP is interpreted as the final state of the force introduced by the v^0 head.

References

Barwise, J. and J. Perry (1983). *Situations and Attitudes*. Cambridge, Mass.: MIT Press.

Bloom, P., M. Peterson, L. Nadel, and M. Garrett (Eds.) (1999). *Language and Space (Language, Speech, and Communication)*. MIT Press.

Copley, B. and H. Harley (2010). An ontology for a force-dynamic treatment of events. semanticsarchive.net; to appear in B. Copley and F. Martin (Eds.) *Forces in Grammatical Structures: Causality from linguistics to philosophy*.

Dowty, D. (1979). *Word meaning and Montague Grammar*. Dordrecht: Reidel.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Jackendoff, R. (1987). On beyond zebra: The relation of linguistic and visual information. *Cognition 26*, 89–114.

Kamp, H. and U. Reyle (1993). *From Discourse to Logic*. Dordrecht: Kluwer.

Kratzer, A. (1989). An investigation of the lumps of thought. *Linguistics and Philosophy 12*, 607–653.

Lakoff, G. and M. Johnson (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books.

Pustejovsky, J. (1991). The syntax of event structure. *Cognition 41*, 47–81.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science 12*(1), 49–100.

Talmy, L. (2000). *Towards a Cognitive Semantics*. Cambridge, Mass.: MIT Press.

Wolff, P. (2007, February). Representing causation. *Journal of Experimental Psychology. General 136*(1), 82–111.

Wolff, P. (to appear). Force dynamics in causal meaning and reasoning. In B. Copley and F. Martin (Eds.), *Forces in Grammatical Structures: Causality from linguistics to philosophy*.

Light verbs features in European Portuguese

Inês Duarte[†*] **Anabela Gonçalves**[†*] **Matilde Miguel**[†*] **Amália Mendes**[*]
Iris Hendrickx[*] **Fátima Oliveira**[α]
Luís Filipe Cunha[α] **Fátima Silva**[α] **Purificação Silvano**[α]

Faculty of Letters[†]
University of Lisbon
Lisbon, Portugal

Centre of Linguistics[*]
University of Lisbon
Lisbon, Portugal

Faculty of Letters/ Centre of Linguistics[α]
University of Oporto
Oporto, Portugal

isduarte@sapo.pt anajop@netcabo.pt tilde@mail.telepac.pt
{amalia.mendes,iris}@clul.ul.pt foliveira@netcabo.pt luisfilipecunha@gmail.com
mhenri@letras.up.pt puri_silvano@hotmail.com

Abstract

We present a study of constructions of the type <light verb + noun> in European Portuguese. We analyse these constructions as complex predicates where both the light verb and the noun share an important role in the predication and we focus on the aspectual combinatory properties of the two elements of the complex predicate. We propose that light verbs inherit the feature specification of the corresponding main verbs, but that the light verbs are underspecified regarding (some of) the event structure features.

1 Introduction

Constructions where a verb and a noun with predicative properties combine have been studied from different theoretical approaches. In these constructions, verbs have been considered i) light verbs as defined in Jespersen's (1949) seminal work, ii) support verbs (Gross, 1981), in the sense that they have lost part or all of their meaning and have no predicative value in the construction, iii) auxiliary verbs with aspectual properties (Abeillé *et al.*, 1998) or iv) a specific subclass of verbs that play a relevant role in the predication (see Rosen, 1990; Butt and Geuder, 2001; Butt, 2003; Samek-Lodovici, 2003, a.o.), also referred to as light verbs. In this paper, we will argue for the latter predicate-like approach on the basis of evidence from European Portuguese. We will concentrate on the properties of sequences of the form <light verb + noun> headed by light verbs *dar* 'to give', *fazer* 'to make/do' and *ter* 'to have' (See examples 1).

(1) (a) O presidente deu algumas orientações ao governo.
'The president gave some orientations to the government.'
(b) O primeiro-ministro fez uma apresentação da nova lei no Parlamento.
'The prime-minister made a presentation of the new law at the Parliament.'
(c) O presidente teve uma conversa com o primeiro-ministro.
'The president had a talk with the prime-minister.'

We will take into account the interaction between these light verbs, the aspectual classes of the verbs from which the nouns are derived and the final interpretation of the resulting complex predicates.

2 Properties of complex predicates of the type <light verb + noun>

These light verbs behave like predicates insofar as they have their own argument structure, preserve the core lexical meaning of their corresponding main verb and exhibit some syntactic alternations of the same kind as the ones main verbs exhibit (as shown in Duarte *et al.*, 2009). They are also sensitive to the aspectual class of the noun they combine with. Taking into account Vendler's (1967) and Moens's (1987) aspectual verb classes and extending them to deverbal nouns (Filip, 1999, a.o.), we observe that the light verb *dar* combines with nouns derived from predicates denoting points, processes, and culminated processes, but not states nor culminations (cf. examples 2):

(2) (a) *O João deu uma estada no Brasil. STATE
the John gave a stay in Brazil
(b) O João deu um passeio. PROCESS
the John gave a walk
'John took a walk.'
(c) O trabalhador deu uma pintura à casa. CULMINATED PROCESS
the worker gave a painting to the house
'The worker painted the house.'
(d) *A Maria deu um nascimento / um assalto à casa. CULMINATION
the Mary gave a birth / a holdup to the house
(e) Cristiano Ronaldo deu um espirro / deu um toque na bola. POINT
Cristiano Ronaldo gave a sneeze / gave a touch to the ball
'Cristiano Ronaldo sneezed / kicked the ball'

The light verb *fazer* combines with nouns derived from predicates denoting culminated processes and processes, but not points or states (cf. examples 3).

(3) (a) *O João fez uma vida. STATE
the John made a life
(b) A Maria fez uma caminhada. PROCESS
the Maria made a walk
'Maria took a walk.'
(c) O João fez uma leitura do artigo. CULMINATED PROCESS
the John made a reading of the paper
'John read the paper.'
(d) *O João fez um espirro. POINT
the John made a sneeze
'John sneezed.'

And finally, the light verb *ter* combines with deverbal nouns denoting processes, culminated processes, culminations, points and states (cf. examples 4):

(4) (a) A Maria teve uma vida fascinante. STATE
'Mary had a fascinating life.'
(b) Os turistas tiveram uma viagem agradável. PROCESS
'The tourists had a nice trip.'
(c) O edifício teve uma construção difícil. CULMINATED PROCESS
the building had a difficult construction
'The building was constructed with difficulty.'
(d) O atleta teve uma chegada triunfal. CULMINATION
'The athlete had a triumphal arrival.'
(e) Cristiano Ronaldo teve um toque genial. POINT
Cristiano Ronaldo had a touch ingenious
'Cristiano Ronaldo kicked the ball with genious.'

Several properties interact to the characterisation of the sequence <light verb+noun> as a complex predicate. On the one hand, its aspectual class is crucially determined by the noun (accepting the hypothesis about the preservation of the aspectual value of the noun as in Marín and McNally, 2009); on the other hand, the external argument of the light verb controls the external argument of the noun it combines with (compare 5a with 5b):[1]

(5) (a) Os professores deram uma grande motivação aos alunos.
the teachers gave a great motivation to the students
'The teachers motivated greatly the students.'
(b) *Os professores deram uma grande motivação dos examinadores aos alunos.
*The teachers gave a great motivation of the examinee to the students.'

3 Event structure of complex predicates

In the spirit of the long trend which describes Vendler's (1967) verb classes in terms of feature clusters (Dowty, 1979; Smith, 1991; Scher, 2005; a.o.) and adopting Moens' (1987) verb classes, we will adapt Harley's (2009: 333) feature specification for verbalizers that form main verbs (see examples 6) to account for the aspectual properties of main verbs themselves.

(6) (a) V_{BE}: [-dynamic], [-change of state], [-cause]
(b) V_{CAUSE}: [+dynamic], [+change of state], [+cause]
(c) V_{BECOME}: [+dynamic], [+change of state], [-cause]

[1] For further arguments in favour of considering the sequence as a complex predicate, see Butt (2003), Duarte *et al*. (2006) and Duarte *et al*. (2009).

(d) V_{DO}: [+dynamic], [-change of state], [-cause]

We propose (i) to keep the features [± dynamic], [± cause] from Haley's feature specification (ii) to use [± change] (ranging over change of state, of location and of possession) (iii) to introduce the feature [±durative], to distinguish culminated processes, processes and states from culminations and points (iv) to introduce [±instant(aneous)], in order to distinguish points from all the other classes (see Smith, 1991) when the verb combines with the noun. The results are presented in (7), with the following association: BE (state), CAUSE (culminated process), BECOME (culmination), DO (process) and DO_INSTANT (points).

(7) (a) V_{BE}: [-dynamic], [-change], [-cause], [+durative] [-instant]
(b) V_{CAUSE}: [+dynamic], [+change], [+cause], [+durative] [-instant]
(c) V_{BECOME}: [+dynamic], [+change], [+cause], [-durative] [-instant]
(d) V_{DO}: [+dynamic], [-change], [-cause], [+durative] [-instant]
(e) $V_{DO\_INSTANT}$: [+dynamic], [-change], [-cause], [-durative] [+instant]

The main verb *dar* is a culmination, *fazer* a culminated process and *ter* a state. We compare these aspectual features to those of the light verbs *dar*, *fazer* and *ter* in (8), arguing for the underspecification of features for this class of predicates.

(8) (a) $\text{dar}_{\text{light}}$: [+dynamic], [±change], [±cause], [±durative] [±instant]
(b) $\text{fazer}_{\text{light}}$: [+dynamic], [±change], [±cause], [±durative] [-instant]
(c) $\text{ter}_{\text{light}}$: [±dynamic], [±change], [±cause], [±durative] [±instant]

The feature specification proposed for these light verbs captures the different combinatory properties and interpretations of the complex predicate headed by each of these verbs. The light verbs *dar* and *fazer* have the [+dynamic] feature, preventing them to combine with nouns denoting states, which are [-dynamic]. As they are underspecified for the [change] feature, they may combine with nouns denoting processes and culminated processes, which will value one of these features: [-change] in the case of a process noun, and [+change], in the case of culminated process nouns. Whereas *dar* is underspecified for the [instant] feature, allowing nouns denoting points, *fazer* keeps the [-instant] value of its homonymous main verb, thus excluding nouns denoting points. In the case of the light verb *dar*, the exclusion of culminations must be accounted for post-syntactically, in the C-I interface. The light verb *ter* is the most defective one, allowing all classes of predicative nouns, which derives from the fact that all its features are underspecified.

We present in (9) and (10) two different results of a combination between the light verb *dar* and two nouns, *passeio* 'walk' and *estada* 'stay'.

(9) (a) dar um passeio
to give a walk
'to take a walk'
(b) $\textit{dar}_{\text{light}}$: [**+dynamic**], [±change], [±cause], [± durative]
(c) *passeio*: [**+dynamic**], [-change], [-cause], [+ durative]
(d) *dar um passeio:* [+dynamic], [-change], [-cause], [+durative]

(10)(a) *dar uma estada no Brasil
to give a stay in Brazil
'to stay in Brazil'
(b) $\textit{dar}_{\text{light}}$: [**+dynamic**], [± change], [±cause], [± durative]
(c) *estada*: **[-dynamic**], [-change], [-cause], [+ durative]

Combining the light verb *dar* with a process denoting noun (*passeio*) do not result in a conflict between the feature specification (cf. 9b-c), while combining the same verb with a state denoting noun (*estada*) creates a conflict between the values for the [dynamic] feature. In (9), the other light verb features are underspecified and the entire sequence inherits the values specified in the noun features.
The complex predicate of the type <light verb + noun> is formed in Syntax: the interpretable underspecified features of the light verb are valued by the specified features of the noun when the verb moves to a functional head, whose Specifier is occupied by the moved DP object.

4 Corpus Annotation

In the scope of our work on complex predicates, we have established guidelines for corpus annotation, taking into consideration a larger set of constructions than the one described here: (i) two main verbs, forming a restructuring construction, like *querer estudar* 'to want to study' (ii) two main verbs in a causative construction, like *fazer rir* 'to make laugh'; (iii) a light verb followed by a deverbal noun: *dar um passeio* 'to take a walk', or by a psych-noun: *ter medo* 'to have fear'. We will here discuss the annotation results regarding the latter type <light verb+noun>, which we restricted to the three verbs *ter*, *fazer* and *dar*, discussed in the preceding sections. (See Hendrickx *et al.* (2010) for the complete discussion of the guidelines, methodology, evaluation and results). The noun+verb constructions are denoted with the tag [CN] where a determiner precedes the noun, and with the tag [CNB] in contexts with bare nouns (cf. examples 11).

(11) (a) Facto que leva a CGD a considerar que não [CNB]tem obrigações em relação aos trabalhadores.
'A fact that leads the CGD to believe that it doesn't have obligations towards the workers.'
(b) o erro de [CN]fazer uma interpretação literal
'the error of making a literal interpretation'

We consider that there is a typical ordering of the elements of the CP in the sense that the canonical form will be <verb (determiner) noun>. However, in the corpus, this is not always the order in which the elements occur (consider, for example, the case of relative clauses) and this is taken into consideration in the annotation system. Consequently, each element of the CP is tagged with information on its typical position inside the CP (position 1, 2, etc.), as well as on its contextual position in the corpus (B=Beginning, I=Intermediate, E=End). The elements forming the CP may not be contiguous in the corpus and in that case only the elements pertaining to the CP are annotated. For example, the adverb *logo* 'immediately' in (12) is not annotated.

(12) dar[CN1_B] logo uma[CN2_I] ajuda[CN3_E]
give immediately an help
'give help immediately'

These guidelines were applied to the manual annotation of the CINTIL[2] corpus, a 1 million words corpus of European Portuguese, tagged and manually revised (Barreto *et al.*, 2006), composed of transcribed spoken materials, both formal and informal (one third of the corpus) and of written materials (remaining two thirds). We present in Table 1 information on the frequency of this type of complex predicates in the written and spoken subpart of the corpus and the partial frequencies of bare nouns and nouns preceded by a determiner.

label	written	spoken	total
CN total	706	586	1292
CNB	353	213	566
CN_	353	373	726
total	**1176**	**805**	**1981**

Table 1: Frequency of the CPs of the type <light verb + noun> in the CINTIL corpus

In Table 2 we observe the frequencies of the three verbs in the corpus. We notice significant differences according to the presence or not of a determiner with the noun: the verb *fazer* is clearly dominant when followed by a noun preceded by a determiner, while the verb *ter* is the more frequent light verb with bare nouns. Further studies of these frequencies will have to take into consideration whether the noun is singular or plural.

CNB	**written**	**spoken**
dar	69	27
fazer	87	52
ter	197	134
total	**353**	**213**
CN	**written**	**spoken**
dar	79	34
fazer	193	231
ter	81	108
total	**353**	**373**

Table 2: Frequency of *dar*, *fazer* and *ter* with determiner +noun [CN] and with bare nouns [CNB]

We plan to contrast our proposals of aspectual restrictions holding between light verbs and nouns presented in Section 3 to the list of nouns occurring in the corpus with each of the three verbs. Another objective is to partially include in the annotation system some of our findings regarding the aspectual information conveyed by

[2] The CINTIL corpus is a joint project of the Natural Language and Speech group of the Sciences Faculty of the University of Lisbon (NLX-FCUL) and the Centre of Linguistics of the University of Lisbon (CLUL).

both light verb and noun and on the aspectual restrictions that hold between the two elements.

5 Conclusion

We have argued for the predicate status of the sequence <light verb + noun> and for the specificity of light verbs with regard to the corresponding main verbs in terms of the underspecification of (some of) their event structure features. We have proposed that complex predicate formation takes place in the Syntax, through the checking/agree operation of the light verb unvalued interpretable aspectual features. Due to this operation, the complex predicate gets its aspectual features (compositional semantics may assign derived aspectual readings to the whole sentence).

Acknowledgments

We would like to thank the anonymous reviewers for their helpful comments and valuable suggestions. The work presented in this paper was undertaken as part of the project *Complex Predicates: typology and corpus annotation*, supported by the National Foundation for Science and Technology (FCT) under the grant PTDC/LIN/68241/2006.

References

Abeillé, A., D. Godard and I. Sag. 1998. Complex Predicates in Nonderivational Syntax. *Two Kinds of Composition in French Complex predicates*, *Syntax and Semantics* vol. 30. San Diego Academic Press, San Diego.

Barreto, F., A. Branco, E. Ferreira, A. Mendes, M. F. P., Bacelar do Nascimento, F. Nunes, and J. Silva. 2006. Open resources and tools for the shallow processing of portuguese. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC2006)*, Genoa, Italy.

Butt, M. and W. Geuder. 2001. On the (Semi)Lexical Status of Light Verbs. In Corver, N. and H. van Riemsdijk (eds.), *Semi-lexical Categories: On the Content of function words and the function of content words*: 323-370. Berlin: Mouton de Gruyter.

Butt, M. 2003. The Light Verb Jungle. *Harvard Working Papers in Linguistics,* Vol 9: 1-49.

Dowty, D. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

Duarte, I., A. Gonçalves and M. Miguel. 2006. Verbos leves com nomes deverbais em português europeu. In *Textos seleccionados do XXI Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa: APL, pp. 315-328.

Duarte, I., M. Colaço, A. Gonçalves, A. Mendes and M. Miguel. 2009. Lexical and syntactic properties of complex predicates of the type <light verb+deverbal noun>, *Arena Romanistica*, 2009:4, pp. 48-57.

Gross, M. 1981. Les bases empiriques de la notion de prédicat sémantique. *Langages*, 63:7–52.

Harley, H. 2009. The morphology of nominalizations and the syntax of vP*. In Giannakidou, A. and M. Rathert (eds), *Quantification, Definiteness, and Nominalization.* Oxford: Oxford University Press.

Hendrickx, I., A. Mendes, S. Pereira, A. Gonçalves, I. Duarte. 2010. Complex Predicates annotation in a corpus of Portuguese. In Proceedings of *the fourth Linguistic Annotation Workshop (*LAW IV), pp 152-156, Association for Computational Linguistics, Uppsala, Sweden, 2010.

Jespersen, O. 1949. *A Modern English Grammar on Historical Principles*. London: George Allen & Unwin; Copenhagen: Ejnar Munksgaard.

Marín, R. and L. McNally. 2009. *From Psych Verbs to Nouns* (ms). International Workshop 'Events across Categories', Madrid.

Moens, M. 1987. *Tense, Aspect and Temporal Reference*. University of Edinburgh: PhD. Dissertation.

Rosen, S. 1990. *Argument Structure and Complex Predicates.* New York, Garland.

Samek-Lodovici, V. 2003. The Internal Structure of Arguments and its Role in Complex Predicate Formation. *Natural Language & Linguistic Theory,* 21: 835-881.

Filip, H. 1999. *Aspect, Eventuality Types and Noun Phrase Semantics*. Garland Publishing, New York.

Scher, A.P. 2005. As Categorias Aspectuais e a Formação de Construções com o Verbo Leve *Dar. Revista GEL,* v. 2: 9-38.

Smith, C.S. 1991. *The Parameter of Aspect.* Dordrecht: Kluwer Academic Publishers

Vendler, Z. 1967. *Linguistics in Philosophy.* Ithaca / London: Cornell University Press.

Polish Sign Language Verbs

Sylwia Fabisiak
Section for Sign Lingustics
University of Warsaw
Poland
S.fabsiak@uw.edu.pl

Paweł Rutkowski
Section for Sign Lingustics
University of Warsaw
Poland
p.rutkowski@uw.edu.pl

Abstract

The aim of this paper is to present an overview of the verbal system of Polish Sign Language (*Polski Język Migowy*, *PJM*) – the native language of the Polish Deaf community. Our analysis derives from corpus data and focuses on certain grammatical characteristics of PJM verbs. The classification of PJM verb signs outlined in this paper is based on the criterion of sign modifiability. We also report the results of an experiment that shows to what extent PJM predicates are interpretable when extracted from the discourse context in which they occur. The classification we propose could be of use in sign language lexicography, especially in bilingual dictionaries (Polish to PJM and PJM to Polish).

1 Polish Sign Language

1.1 Polish Sign Language and signed Polish

Polish Sign Language (*Polski Język Migowy*, *PJM*) is a natural visual-spatial language used by the Polish Deaf community. PJM emerged around 1817, with the foundation of the first school for the deaf in Poland – the Warsaw Institute of the Deaf-Mute and the Blind (cf. Świdziński, 2003). This language is genetically independent of the Polish language, and should not be confused with an artificial manually coded language created by hearing people, the so called Language-Sign System (*System Językowo-Migowy*, *SJM*), a kind of signed Polish, which combines signs from PJM, artificial signs and signs borrowed from other sign languages with the grammar of spoken Polish (cf. Szczepankowski, 1999). As a signed sub-code of Polish, SJM is a foreign language to the Deaf. Up until recently the research on signing in Poland focused exclusively on SJM, perceived as a tool of communication between the hearing majority and the hearing-impaired minority. The research into the natural sign language started only less than 20 years ago (Farris, 1994), becoming part of the dynamically developing field of sign language linguistics (see e.g. Sandler and Lillo-Martin, 2006).

1.2 Articulation in sign language

Sign language is articulated with the whole body of a signing person, especially their hands and face, but also their torso. According to the tradition of sign language research, a number of sub-lexical parameters may be distinguished in a sign – perceived by some scholars as equivalent to phonemes, whereas by others solely as diacritic features of a sign. These parameters include the position of one or both hands, place of articulation of a sign, movement involved in sign formation, palm orientation relative to the body, and facial expression (cf. Stokoe, 2005).

1.3 Visual-spatial grammar

The grammatical system of PJM, as in any other visual-spatial language, includes components which are not considered part of grammar in spoken languages – body movements (e.g. leaning forward as one of the elements constituting a question) and facial expressions (e.g. raising eyebrows and frowning, wrinkling one's nose, mouth movements, direction of gaze), which mark, among others, the intensity of verbs, comparison of adjectives, interrogative and imperative sentences etc. Grammar of sign languages is characterized by three-dimensionality (spatiality) and the possibility of simultaneous articulation of signs (which results from the simultaneous use of different articulators, e.g. the manual and non-manual ones). PJM is an analytic language, although some phenomena present in the language could be interpreted as examples of inflection, such as the behavior of directional verbs, i.e. signs whose meaning is modified depending on the direction (agent-patient, patient-agent) in which they are produced (e.g. in the sentence 'you give me', the sign GIVE is directed towards the signer, whereas in the sentence 'I give you', it is directed towards the addressee, cf. Grzesiak, Chrzanowska, 2007).

2 Verb signs

2.1 Distinguishing verb signs

The issue of dividing PJM signs into parts of speech is still open. It is often the case that the same sign plays different grammatical roles, depending on the context. This applies, among others, to PJM verbs, whose shape is often indistinguishable from the shape of semantically related nouns. Researchers attempting to identify the category of sign language verbs run up against numerous difficulties. For the needs of this paper we will tentatively define PJM verbs as a group of signs, which differ from nouns with respect to syntax and semantics – namely, through their predicative function. One of the most characteristic traits of PJM verb signs is the semantic incorporation of the agent/patient of the predication – we will therefore find signs such as WASH_WINDOW, WASH_WHOLE_BODY, WASH_FACE, WASH_DISHES, but no general sign for WASH, one that could be combined with any object.

2.2 Classifiers and classifier predicates

When discussing sign language verbs, one cannot omit the related category of classifiers, i.e. signs that fulfil an anaphoric function. A classifier is a defined position of one or two hands, which replaces a previously indicated noun (Klima and Bellugi, 1979). The form of a classifier is unchangeable and reflects such qualities as shape, size, position in space; hence the division into person classifiers, animal classifiers (different for small and big animals, as well as those that move in a characteristic way, e.g. bears), vehicle classifiers (car, plane, bike, bus) and inanimate object classifiers (e.g. books). The use of a classifier makes it often possible to produce a sentence devoid of any lexicalized verb sign. In such cases, the so-called classifier predicates function as quasi-verbs. Such predicates are based on a classifier handshape combined with a movement imitating the three-dimensional movement of the entity denoted (coding its speed, path and character). There is a lot of flexibility in what classifier predicates may look like. Their spatial and dynamic properties are often very complex since they are dependent on the real-world movement that is being mimicked. Therefore, it is usually difficult to describe the dynamic characteristics of a classifier predicate in terms of discrete linguistic features. The class of classifier predicates consists of a virtually unlimited number of possible combinations of various classifiers and movement types. Such predicates may imitate all kinds of real-world situations. For instance, a classifier predicate may represent a human being that moves ("walks") in a certain direction, but also an instrument that is being handled by a human being in a certain way. Since classifiers do not denote specific entities but rather allude to their general physical properties, the exact interpretation of classifier predicates depends on the context in which they are used. To give an example, the classifier imitating a hand holding a pipe-like object combined with a back-and-forth movement may as easily refer to digging with a spade as to vacuum-cleaning.

2.3 Verb-to-object adaptation and imitation

Another important issue related to sign language verbs is what we somewhat generally label as verb-to-object adaptation. It is a quality of a sign, expressed by the modification of a verb's shape, depending on the noun it takes. The character of these modifications is usually strictly iconic, i.e. they imitate the actual movements of an object (e.g. the notion of 'swimming' will be expressed with a different shape when referring to a person, a different one when talking about a frog, a still different one for a fish, and so on). Verb-to-object adaptation will therefore be linked to incorporation and the use of classifiers. In certain works the verb-to-object adaptation is described as a form of verb inflection (cf. Klima, Bellugi, 1979); however, it seems that in PJM the sign structure of the verbs in question does not contain any easily definable morphological sub-components comparable to inflection. The existence of such verbs seems to indicate the importance of the iconic aspect of sign language, which is reflected in many dimensions of language structure, and above all on the level of denotation mechanisms (cf. Taub, 2001).

Imitation, understood as the way in which the sign relates to reality, i.e. the similarity between the sign and its denotation, is a common phenomenon in sign languages.

3 Classifying PJM verb signs

3.1 Other scholars' classifications

Previous research on sign languages has produced several classifications of verb signs, some more detailed than others. The basic distinction found in the literature is between plain verbs, spatial verbs, and agreement verbs (cf. Padden, 1988). This classification is based on the role the hands play in encoding the arguments of a verb. Plain verbs (e.g. LIKE in American Sign Language, ASL) resemble typical verbs in spoken languages, as they do not incorporate any grammatical features of their arguments, i.e. they have to be linearly combined with separate nominal arguments in order to form sentences. Spatial verbs (e.g. PUT in ASL) convey information related to the motion and location in space of their arguments (usually, the locative source and goal of an action), i.e. they often involve hand movement, whose path in space reflects the real-world movement related to the predication in question. Agreement verbs (e.g. GIVE in ASL) denote transfer, i.e. they encode the syntactic role of their arguments, e.g. by directing the movement of the hands from the subject to the object. Other criteria have also been used to analyze the syntactic properties of verbs in sign languages. For instance, Zeshan (2000) distinguishes a class of closed (i.e. unmodifiable) signs and a class of modifiable signs, which includes verbs. The latter may be modified according to the place of articulation, movement, and location in space. In the following section we would like to propose a classification of PJM verbs, based not on how the arguments of verbs are encoded, but rather on the criterion of modifiability.

3.2 Three classes of PJM predicates

Our research is based on 2 hours of video material selected from the corpus of PJM that is being compiled at the University of Warsaw. In the PJM corpus project, data is collected from signers who either have Deaf parents or have used PJM since early childhood. The informants are asked to react to certain visual stimuli, e.g. by describing a picture or discussing a video recording. The signers are all adults and come from different regions of Poland. For the purposes of the present study, we have analyzed the semantic and syntactic properties of all the predicates that occur in signed utterances produced by two PJM signers (one hour of recorded material per person). A detailed inspection of this set of data has allowed us to distinguish three types of PJM predicates. Their main characteristics are discussed below.

3.2.1 Context-free predicates

This class consists of plain verbs that do not undergo any contextual modifications. It includes signs like SEEK, SLEEP, MEET. From the syntactic point of view these predicates behave like verbs in spoken languages, i.e. they are linearly combined with their arguments. Most importantly, they do not adapt in any way to their arguments, which means that their shape is independent of the sentential context in which they occur.

3.2.2 Context-modifiable predicates

This class includes verbal signs that are usually modified in a certain way, i.e. they adjust to their sentential context. The modification in question may be related to one of the following parameters:
- direction and path of movement (resulting from the presence and location of the agent and patient of an action); this kind of modification is possible in the case of predicates like CRITICIZE, HAVE, LOOK, e.g. the sentences 'I criticize you' and 'you criticize me' have to be signed in the opposite directions;
- manner (speed, intensity, emotional attitude towards an action etc.); e.g. EXERCISE, THINK, QUARREL; this kind of modification is usually obtained not only by intensifying the movement parameter, but also through the use of non-manual elements, e.g. in the sentence 'I exercised hard' the predicate is accompanied by a frown that functions as an intensifier;
- aspect (expressed through the opposition of one vs. reduplicated movement or created with the aid of analytical constructions, i.e. by using auxiliary signs, such as ALREADY or WAS which mark perfectivity).

3.2.3 Context-dependent predicates

This group includes predicates that are essentially imitational/mimetic. Their shape is strictly dependent on the action that they refer to.

Most of them are classifier predicates, i.e. they include one of many classifier handshapes available in PJM. However, the classifier seems to be the only conventional element in such signs. The rest of the semantic content is conveyed by an iconic movement, which can be freely modified. For instance, when a signer wants to refer to 'sliding down a slide', they will use a classifier representing a person and move it in three-dimensional space, following the path of the real-world slide that is being referred to. Needless to say, the exact interpretation (i.e. the fact that such a sign refers to sliding and not to, for instance, ski jumping) derives from the context.

3.3 An experiment: semantic interpretation of context-dependent predicates

In order to judge whether the above classification corresponds to psychological reality, we have carried out an additional questionnaire investigation aimed at testing the degree to which context-dependent predicates are interpretable out of the context in which they were originally produced. We presented a number of short video clips extracted from the PJM corpus material to 15 subjects: five native signers, five L2 learners of PJM, and five hearing speakers of Polish with no knowledge of PJM. They were all asked to interpret a set of classifier predicates, which consisted of the following 12 signs:

Clip No.	Original meaning
1.	to dabble in water
2.	to pillow fight
3.	to yell at each other
4.	to gnaw at leaves
5.	to sniff at each other
6.	to vacuum clean
7.	to vacuum clean (with the carpet sucked into the vacuum cleaner)
8.	to swim (synchronized, forming a heart-like shape)
9.	to handle a surfboard
10.	to ballet dance
11.	to climb a ladder
12.	to slide down a slide

The subjects were not told anything about the grammatical status of the signs (i.e. whether they were nouns or verbs) – their only task was to explain (or, as in the case of the subjects with no knowledge of PJM, to guess) the meaning of the 12 signs. The informants filled out the same questionnaire in which they were asked to give all the possible meanings of each of the signs presented to them. Our initial expectation was that the native signers should be able to interpret the predicates more easily than the other subjects. However, the results showed that, when devoid of contextual information, context-dependent predicates are equally difficult to interpret for the three groups of subjects. The table below shows the responses that we obtained.

Clip No.	Responses in the experiment		
	Deaf	PJM Learners	Hearing
1.	to drum [2 subjects], bicycle, to run, bull	running animal, dog to drum, to laugh, to play the piano	to drum [2 subjects], to run [2 subjects], to swing legs
2.	to fight [2 subjects], arrival [2 subjects], adventure, meeting, to whip	meeting, assembly, to kill, to slap, to pack, to load	to beat [2 subjects], to fight, to hit, to explain
3.	to bite each other [2 subjects], to shout [2 subjects], to quarrel	to gossip, to quarrel, to bite, to bark, to growl, to shout	to quarrel [3 subjects], to shout, to eat
4.	to scratch [3 subjects], to fawn, to bite	to scratch [2 subjects], predator, to knock, to dial	to shout [2 subjects], to eat, to talk, animal

5.	to jump [2 subjects], to have sex [2 subjects], request	to prepare for a fight, to analyze, to plug in, to yelp, two animals	to jump, to pick, to cry, to look at each other, difficulty
6.	to pump [3 subjects], to vacuum clean, to fish	to pump, to drag, to drop, to play the double bass, to pull on	to drag [2 subjects], to pull [2 subjects], to vacuum clean
7.	to dig [3 subjects], spade, to bury	to dig [3 subjects], spade [2 subjects]	to dig [4 subjects], to broom
8.	sky, heart [2 subjects], necklace of beads, carousel, amusement park, bundle	to swim, a heart-like shape made by two planes, aerial stunts, love, swans	love [2 subjects], dolphins, lovers, change
9.	to rock [4 subjects], to tear off	to rock [3 subjects], baby carriage [2 subjects]	to swing [2 subjects], to bike, to move, to shake
10.	ballet [2 subjects], dancing [2 subjects], metronome	metronome, legs, to bounce, pendulum, Charlie Chaplin	dancing [2 subjects], legs [2 subjects], cutlery
11.	squirrel [2 subjects], to climb [2 subjects], ladder	to climb [4 subjects], stairs	to climb [2 subjects], to write, to seek, to go up
12.	to jump [2 subjects], slide, going down, ski jump	to land [2 subjects], to go down and land, ski jump, going down	to land [3 subjects], ski jump, to jump

Our experiment showed no significant differences between the three groups of subjects in terms of the accuracy of semantic interpretation. Interestingly, most of the subjects correctly identified the signs presented to them as verbs, however the exact meaning was much more difficult to define. This suggests that context-dependent predicates are far more iconic and far less conventionalized than other PJM signs. They also allow for a virtually unlimited number of modifications. Although we do not assume that these results can be interpreted as conclusive evidence for the necessity of distinguishing context-dependent predicates from other types of PJM verbs, we can clearly see that, unlike their counterparts in spoken languages, such predicates are underspecified with respect to their semantic value, which makes them difficult to interpret out of the discourse context in which they occur. This kind of underspecification does not seem to have obvious parallels in spoken languages. We treat this experiment as a pilot study for further research.

4 Summary

We hope that this short review of some important issues related to the distinguishing and classifying of PJM verb signs has allowed us to demonstrate the problems that may be encountered by a researcher of a visual-spatial language, especially one that still remains a terra incognita, like PJM. The proposed classification of PJM verbs is of course a tentative one; it may however be the basis for further, more detailed research. Such research could result, among other things, in facilitating the process of compiling PJM dictionaries. Sign lexicography in Poland is practically non-existent – up until today the only dictionaries published have been those between Polish and signed Polish (SJM), which are usually collections of photographs presenting signs accompanied by glosses in Polish (cf. Hendzel, 1995). Obviously, one of the most problematic issues related to PJM lexicography is that of assigning a grammatical interpretation to sign lexemes. The classification of verbs proposed in this paper could be used in lexicography by providing verb signs with information on what types they belong to and what their possible modifications are. This information could be of great use in the sign language learning/teaching process.

References

Farris Michael. 1994. Sign language research and Polish Sign Language. *Lingua Posnaniesis*, 36: 13-36.

Grzesiak Iwona and Milena Chrzanowska. 2007. O znakach kierunkowych w Polskim Języku Migowym. In: Iwona Grzesiak (ed.) *Język migowy we współczesnym szkolnictwie w Polsce i na świecie. Materiały z sympozjum naukowego w Malborku w dn. 31.03.2007.* Fundacja na Rzecz Osób Niesłyszących i Języka Migowego, Malbork, pp.46-61.

Grzesiak Iwona. 2007. *Strukturalna klasyfikacja i systematyzacja znaków Polskiego Języka Migowego dla potrzeb leksykografii dwujęzycznej*. Fundacja na Rzecz Osób Niesłyszących i Języka Migowego, Olsztyn.

Hendzel Józef. 1995. *Słownik polskiego języka miganego*. OFFER, Olsztyn.

Johnston Trevor and Adam Schembri. 2007. *Australian Sign Language. An introduction to sign language linguistics*. Cambridge University Press, Cambridge, UK.

Sandler Wendy and Diane Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge University Press, Cambridge, New York.

Klima Edward and Ursula Bellugi. 1979. *The signs of language*. Cambridge University Press, Cambridge.

Padden Carol. 1988. *Interaction of morphology and syntax in American Sign Language*. Garland Press, New York.

Stokoe William. 2005. Sign language structure: An outline of the visual communication systems of the American Deaf. *Journal of Deaf Studies and Deaf Education*, 10 (1): 3-37.

Świdziński Marek. 2003. Bilingual education of the Deaf in Poland: the past and the future. In: Marek Świdziński, Tadeusz Gałkowski (ed.) *Studia nad kompetencją językową i komunikacją niesłyszących*. Uniwersytet Warszawski: Wydział Polonistyki, Wydział Psychologii, Polski Komitet Audiofonologii, Instytut Głuchoniemych im. ks. Jakuba Falkowskiego, Warszawa, pp. 265-270.

Szczepankowski Bogdan. 1999. *Niesłyszący, głusi, głuchoniemi – wyrównywanie szans.* WSiP, Warszawa.

Taub Sarah. 2001. *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge University Press, Cambridge, New York.

Zeshan Ulrike. 2000. *Sign Language in Indo-Pakistan. A description of a signed language.* John Benjamins Publishing Company, Philadelphia.

Bootstrapping a Classification of French Verbs Using Formal Concept Analysis.

Ingrid Falk
INRIA/Nancy 2,
Nancy, France

Claire Gardent
CNRS/LORIA,
Nancy, France

Abstract

We use Formal Concept Analysis (FCA) to bootstrap a classification of French verbs. We show that the resulting classification has good factorisation power, compare it with the English Verbnet and report on a partial qualitative evaluation.

1 Introduction

Verb classifications have often been proposed which group together verbs with similar syntactic and/or semantic behaviour. On the practical side, verb classes permit capturing generalisations about verb behaviour thus reducing both the effort needed to construct a verb lexicon and the likelihood that errors are introduced when adding new entries. On the theoretical side, (Levin, 1993) has shown that syntax reflects semantics and consequently, that verbs that belong to a syntactic class can be shown to often share a semantic component.

In this paper, we explore the use of Formal Concept Analysis (FCA) to acquire classes for French verbs from available lexical resources. We start by outlining the intuition behind the proposal and describing the lexical resources used. We then show how FCA can be used to produce a verb classification and compare it with the English Verbnet[1].

2 Formal concept analysis

FCA (Ganter and Wille, 1999) is a classification technique which permits creating, from a so-called formal context, a concept lattice where concepts associate sets of objects with sets of attributes. Here, the concept objects will be verbs while the attributes will be syntactic frames and semantic features. Intuitively, a concept is a pair $\langle O, A\rangle$ such that all the objects in O have exactly the attributes in A and vice versa, all attributes in A are true of exactly all the objects in O. That is, our concepts will group together sets of verbs which share exactly the same set of syntactic and semantic features.

More formally, a formal context $\mathcal{K}$ is a triple $\langle \mathcal{O}, \mathcal{A}, R\rangle$ such that $\mathcal{O}$ is a set of objects, $\mathcal{A}$ a set of attributes and R a relation on $\mathcal{O} \times \mathcal{A}$. Given such a context, a concept is a pair $\langle O, A\rangle$ such that $O = \{o \in \mathcal{O} \mid \forall a \in A.\ (o, a) \in R\}$ and vice versa $A = \{a \in \mathcal{A} \mid \forall o \in O.\ (o, a) \in R\}$. Two operators, both denoted by $'$, connect the power sets of objects $2^{\mathcal{O}}$ and attributes $2^{\mathcal{A}}$ as follows: $' : 2^{\mathcal{O}} \to 2^{\mathcal{A}}, X' = \{a \in \mathcal{A} \mid \forall o \in X.\ (o, a) \in R\}$. The operator $'$ is dually defined on attributes. For a formal concept $\langle O, A\rangle \in O \times A$ we have $O' = A$ and $A' = O$. O is called the *extent* or *extension* and A the *intent* or *intension* of the formal concept.

A concept $C1 = \langle O1, A1\rangle$ is smaller than another concept $C2 = \langle O2, A2\rangle$ (written $C1 \leq C2$) iff $O1 \subseteq O2$ and $A1 \supseteq A2$. The set of all formal concepts of a context $\mathcal{K}$ together with the order relation $\leq$ form a complete lattice called $\mathbb{K}$, the concept lattice of $\mathcal{K}$. That is, for each subset of concepts there is always a unique greatest common subconcept and a unique least common superconcept.

3 Lexical resources

We now present the linguistic resources used to build and evaluate a classification of French verbs namely, Dicovalence, the LADL tables and VerbNet.

Dicovalence (van den Eynde and Mertens, 2003) is a syntactic lexicon for French verbs which lists among other things the valency frames of 3 936 French verbs. We use here a version of Dicovalence converted (Gardent, 2009) as follows. Each verb is associated with one or more valency

[1]Other applications of FCA to linguistics and lexical resources are presented for eg. in (Priss, 2005) and (Valverde-Albacete, 2008).

frame characterising the number and type of the syntactic arguments expected by this verb. Further, each frame describes a set of syntactic arguments and each argument is characterised by a grammatical function[2] and a syntactic category[3]. For instance, the frame of *Jean maintient ouvert le robinet / Jean maintains the tap open* will be SUJ:NP, OBJ:NP, ATO:XP.

The LADL tables (Gross, 1975), (Guillet and Leclère, 1992) were specified manually over several years by a large team of expert linguists and contain syntactic and semantic information about French verbs. For instance, a table might state that the subject of all verbs in that table must be human; or that the object is a destination, etc. The LADL tables group 5076 verbs into 61 distinct tables each table being associated with a defining valency frame and an informal description of the properties shared by verbs in that table[4].

VerbNet (Schuler, 2006) is a verb classification for English which was created manually and classifies 3 626 verbs using 411 classes. Each VerbNet class includes among other things a set of verbs and a set of valency frames.

4 Acquiring verb classes

Our ultimate aim is to create a classification which facilitates the maintenance and verification of lexical verbal information such as in particular, valency frames and thematic grids. In the present paper however, we take an intermediate step towards that goal and aim at finding a method for producing verb classifications which display the following properties.

Factorisation: the number of classes remains relatively small (no more than a few hundred) and in average, classes are balanced and well populated. That is, there are not too many classes with either very few frames or very few verbs.

Coverage: The classification covers most of the verbs and (verb, frame) pairs present in Dicovalence.

[2]SUJ refers to the subject grammatical function, OBJ to the object, P-OBJ, A-OBJ and DE-OBJ describe prepositional objects introduced by any preposition, *à* or *de* respectively and ATO indicates an object attribute.

[3]NP indicates a noun phrase, PP a prepositional phrase, CL a clitic and XP any major constituent

[4]The columns of the table give further more detailed information about each verb in the table but we do not use this information here.

Similarity: The classes group together verbs sharing both a syntactic (frames) and a semantic (selectional restrictions, event type, argument structure) component

The FCA lattice. To create verb classes which capture both a shared syntactic behaviour (a shared set of valency frames) and a shared meaning component, we first build a concept lattice[5] based on the formal context $\langle V, F, R \rangle$ such that the set of objects V is the set of verbs contained in the intersection of Dicovalence and the LADL tables, the set of attributes F is the union of the set of valency frames used in Dicovalence with the set of LADL table identifiers and the relation R the mapping such that $(v, f) \in R$ if either Dicovalence or the LADL tables associates the verb v with the frame/table f.

Filtering. The resulting lattice contains 36065 concepts. Not all these concepts are interesting verb classes however. In particular, many concepts only have 1 or 2 verbs and can hardly be viewed as classes. Similarly, concepts with few frames are less interesting especially if many of the verb subclasses of the extension of these concepts have more frames than there are in their intension. To select from this lattice those concepts which are most likely to provide appropriate verb classes, we consider only concepts (i) whose attribute set contains at least one table identifier and one valency frame that is, which share both a syntactic and a semantic feature and (ii) that are intensionally stable (Kuznetsov, 2007). Intensional stability is a measure which helps discriminating potentially interesting patterns from irrelevant information in a concept lattice based on possibly noisy data. The intensional stability of a concept (V, F) is defined as $\sigma_i((V, F)) = \frac{|\{A \subseteq V | A' = F\}|}{2^{|V|}}$. In words, the stability of a concept (V, F) is defined as the number of those object subsets of V which have the same set of attributes as V divided by the total number of subsets of V. Intuitively, a more stable concept is less dependant on any individual object in its extent and is therefore more resistant to outliers or other noisy data items. Selecting concepts with high intensional stability yields classes which provide a good level of generalisation (their frame set is true of many verb sets).

[5]We used the Galicia Lattice Builder software (`http://www.iro.umontreal.ca/~galicia/`) to build the lattices

Coverage. One drawback with our filtering method is that since not all concepts are kept, some verbs and some frames might not be covered by the classification. In practice however, taking the 430 concepts with stability threshold 0.9995 (*Class430* in the following) and whose attribute set obey the set constraints (i.e., at least one table and one frame) yields a classification which covers 98.41% of the verbs, 25% of the frames and 83.17% of the (verb, frame) pairs. That is, the resulting classification covers most of the input data except for frames that have a rather low coverage due to many frames (in particular VPinf subject frames) with low frequency.

5 Quantitative evaluation

We first comment on the classification obtained based on a quantitative comparison with Verbnet.

5.1 Comparison with Verbnet.

Table 1 gives a more detailed presentation of the impact of the stability threshold on the obtained classification. A threshold of 0.9995 yields a number of classes closest to that observed in Verbnet (430 against 411 in Verbnet). Fig. 1, a comparison of the distributions of verbs and frames in classes for VerbNet and *Class430*, shows that the distributions are similar, although VerbNet has more classes with a small number of verbs. The main difference between Verbnet and our classification stems from the inventories of frames used. Although Dicovalence and Verbnet use approximately the same number of frames (116 and 117 respectively), many frames have a low frequency in Dicovalence so that our classification only retains 29 of the 116 initial Dicovalence frames. As a result, Verbnet has classes with a higher number of frames (average and maximum) and relatedly a lower number of verbs. Interestingly, finer grained classes are used in Verbnet where in particular, NP and PP categories are sometimes specialised with thematic roles (e.g., NP.patient vs NP.topic) and sentential arguments are differentiated into whether/how/what sentences. In future work, we intend to extend the classes and frames with thematic roles which might result in a classification distribution closer to that of Verbnet.

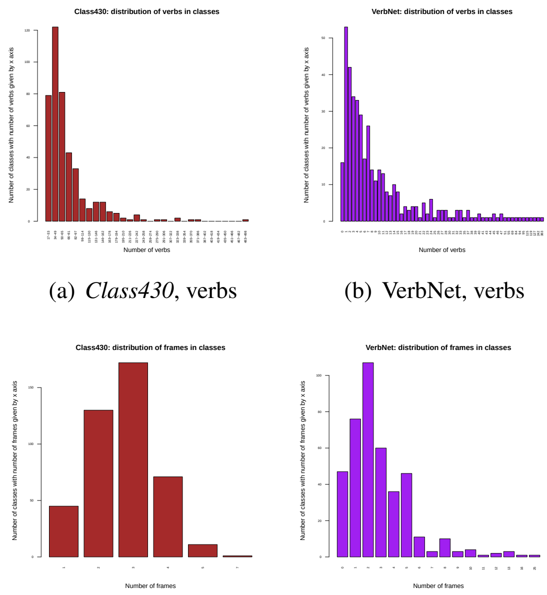
5.2 Factorisation.

Each class is associated with one or more semantic label (i.e., LADL table) and between 1 and 7 valency frames. Furthermore, the resulting classes each contain between 18 and 498 verbs. Overall thus, the classification obtained associates verb sets with an informative syntactico-semantic characterisation; groups together a satisfactory number of verbs and frames; and permits covering a majority of verbs and (verb, frame) pairs present in Dicovalence.

We also plotted the LADL tables against the number of classes they include. For most tables (61%), less than 5 classes are identified. There are 5 tables which are assigned no class – these are all relatively small tables (around 20 verbs) for which no class could be found whose verbs were included in the set of verbs contained by the table.

5.3 Example class.

An example class extracted by this method associates the LADL tables 32RA (Make Adj$_v$), 8 (Verbs with sentential complement in *de*) and the frames SUJ:NP; SUJ:NP,OBJ:NP; SUJ:NP,DE-OBJ:PP with the verb set { blanchir (*to whiten*), bleuir (*to turn blue*), blêmir (*to turn pale*), pâlir (*to turn white*), rajeunir (*to become younger*), rosir (*to turn pink*), rougir (*to blush*), verdir (*to turn green*), vieillir (*to age*)}. That is, the class groups together verbs which indicate a change of state (mainly colour and age) and which can be used with and without object as well as with a senten-



(a) *Class430*, verbs (b) VerbNet, verbs

(c) *Class430*, frames (d) VerbNet, frames

Figure 1: Number of classes with the number of verbs/frames given on the x-axis for VerbNet and *Class430*.

Minimal stability	0.9999	0.9995	0.9990	VerbNet
Nb. of classes	340	430	500	411
Min. verbs	20	18	18	1
Max. verbs	498	498	498	383
Min. frames	1	1	1	1
Max. frames	5	7	7	25
Min. depth		2		1
Max. depth		6		4
Classes with 1 verb	0	0	1	29
Classes with 1 frame	41	45	49	44
Avg. class size (verbs)	78.5	70.13	66.16	14.96
Avg. class size (frames)	2.61	2.71	2.76	4.02
Avg. class size (harm. mean)	6.87	7.02	7.09	4.67
Verb coverage (%)	97.99	98.41	98.70	
Frame coverage (%)	17.74	18.28	18.28	
Verb-frame pairs coverage (%)	80.81	83.17	84.19	
Total number of verbs			3536	3626
Total number of frames			116	117

Table 1: Some features of the verb classification depending on the chosen stability threshold.

tial *de*-object.

6 Qualitative evaluation

To explore the extent to which our classes group together verbs with identical thematic grids, we focus on psychological verbs i.e., verbs which, in the LADL tables, are described by table 4. Figure 2 shows the subgraph of our classification (*Class430*) rooted in the class with *table 4* as semantic feature. By inheritance all classes in this subgraph also have *table 4* as semantic feature.

Table 4 contains 616 verbs describing emotion or psychological verbs. All verbs in this table enter a transitive construction where the object is always human and the subject may be clausal (eg. *Que Luc agisse ainsi amuse Max /That Luc behaves this way amuses Max*). Because the subject of table 4 verbs may be phrasal, the EXPERIENCER is always the object (not the subject). Furthermore, the subject may accept both a (non-agentive) CAUSE and a (volitional) AGENT reading. Consequently, the thematic grid of verbs in this subclassification is [(CAUSE or AGENT), EXPERIENCER].

We now consider the subclassification in more detail and point out to several interesting ways in which the FCA approach interacts with polysemy and linking i.e., the mapping between syntactic arguments and thematic roles.

6.1 Polysemy

In Fig. 2, we outlined in blue the classes which have an additional table identifier in their attribute set and therefore may have an additional meaning[6]. For instance, class 4562 is associated not only with table 4 but also with table $32C$ which contains transitive verbs with a concrete object (Eg. *toucher le mur/touch the wall*). This suggests that the verbs in this class have both a psychological reading (Table 4, *toucher le public/move the audience*) and a concrete object reading (Table 32C, *toucher le mur/touch the wall*). More specifically, this suggest that verbs in class 4562 accept not one but 2 thematic grids and linkings namely:

Table 32C	Table 4
NP.AGENT NP.PATIENT	NP.CAUSE/AGT NP.EXP.
Jean touche le mur	*Jean touche le public*
Jean touches the wall	*Jean moves the audience*

6.2 Linking

Next we considered those classes (marked with a red font in Fig. 2) which are not characterised by an additional table and have at least 3 frames (marked with a red font in Fig. 2). That is, we consider classes which are semantically homogeneous (Table 4 only) and syntactically varied (several frames). For these classes, we examined whether it was possible to consistently determine linking i.e. to consistently assign thematic roles to syntactic arguments in the various frames.

This worked well for most of the 20 classes fulfilling our selection criterion (table 4, more than 3 frames). For instance, class 14650 (28 verbs) could be assigned the following linking information:

NP.CAUSE OR AGENT NP.EXPERIENCER
Jean irrite Marie/Jean irritates Mary
NP.CAUSE OR AGENT
Jean irrite/Jean irritates.
NP.EXPERIENCER, reflexive
Marie s'irrite/Marie irritates herself
NP.EXPERIENCER, PP.CAUSE OR AGENT, reflexive
Marie s'irrite contre Jean/Marie irritates herself against Jean.

Interestingly, Class 15856 departs from Class 14650 in that it groups together verbs for which the thematic role of the subject is ambiguous (Cause or Experiencer) when the verb is used in the intransitive form:

NP.CAUSE OR AGENT NP.EXPERIENCER
La douleur étouffe Marie/The pain suffocates Marie.
NP.CAUSE
La douleur étouffe/Pain suffocates.
NP.EXPERIENCER
Marie étouffe./ Marie suffocates
NP.EXPERIENCER, reflexive
Marie s'étouffe/Marie suffocates.

[6] In the LADL tables, the same verb occurring in different tables usually indicate that the verb has several possible meanings.

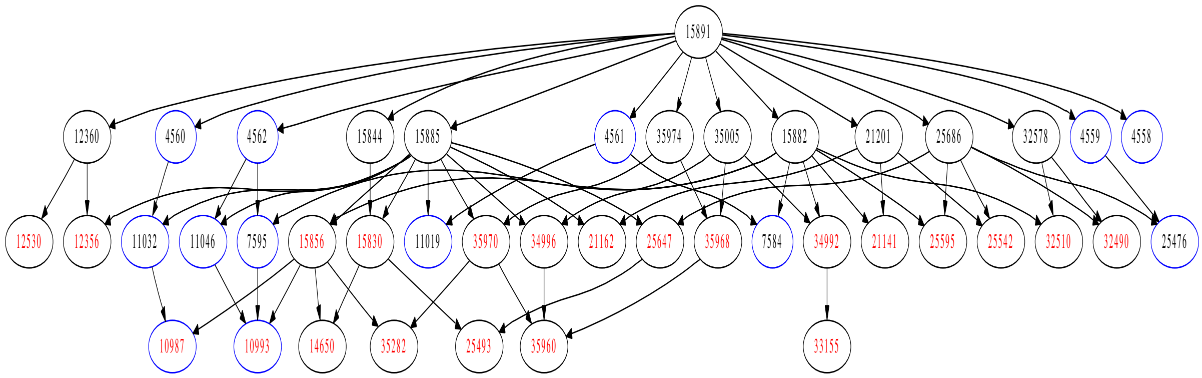


Figure 2: Hierarchic representation of verb/frame classes for LADL table 4

For the other eight classes with intransitive irreflexive frames this ambiguity did not appear, in particular the linking between syntactic argument and thematic role is straightforward in the three cases where the subject of the intransitive is clausal.

Finally, class 25647 (with 37 verbs) suggests that a more fine-grained representation of prepositional objects is needed to correctly determine linking. More specifically, information about preposition type (locative vs. beneficiary) is required to determine whether the EXPERIENCER role is realised by the object NP (*les jeunes*) or the prepositional object (*en moi*). Here, taking into account prepositions may help at separating the verbs of this class according to the syntactic realisation of EXPERIENCER.

NP.CAUSE NP.EXPERIENCER
Ceci exaspère/anime Marie.
NP.EXPERIENCER, reflexive
Marie s'exaspère/s'anime.
NP.CAUSE/AGENT, NP1, P-OBJ:PP
Elle exaspère [en moi]$_{\text{P-OBJ:EXPERIENCER}}$ *ce désir.*
Paul anime [les jeunes]$_{\text{NP1:EXPERIENCER}}$ *contre moi.*

To sum up, this case study shows that the proposed classification scheme permits associating thematic role with syntactic arguments for a large majority of classes.

7 Conclusion

Developing a verb classification by hand is time consuming and error prone. It also makes it difficult to ensure consistency within and across classes. The results presented in this paper suggest that FCA is an appropriate framework for bootstraping a verb classification for French from existing lexical resources. First, concepts naturally model the association between object (verbs) and attributes (syntactic and/or semantic features). Second, like fuzzy clustering, FCA permits "soft clustering" in that a data element may belong to several classes – a property of the produced classifications which is essential for our task since verbs are highly polysemous and may belong to several syntactic and/or semantic classes. Third, stable concepts and symbolic filtering on the attribute sets permit creating classes with good factorisation power (e.g., a few hundred syntactic classes to cover roughly 3 500 verbs) and linguistically sound, empirical content (good average number of verbs and frames within the classes). Fourth, a preliminary and partial qualitative evaluation suggests that the classes built adequately describe the association between verb sets, syntactic frames and thematic grids.

Ongoing work concentrates on enriching the classification with additional features such as passivisation, reflexivisation, middle voice, etc.; and on further evaluating the classes obtained in particular, wrt their ability to group together verbs with identical thematic grids.

Acknowledgments

The research reported in this paper was partially supported by the French National Research Agency (ANR) in the context of the Passage project (ANR-06-MDCA-013). We would like to thank Yannick Toussaint for his suggestion to use stability as a filter as well as for general feedback and help on the topics addressed in this paper and Corinna Anderson for help and suggestions on several linguistic issues.

References

[Ganter and Wille1999] Bernhard Ganter and Rudolph Wille. 1999. *Formal concept analysis: Mathematical foundations*. Springer, Berlin-Heidelberg.

[Gardent2009] Claire Gardent. 2009. Evaluating an Automatically Extracted Lexicon. In *4th Language & Technology Conference*, Poznan, Poland.

[Gross1975] Maurice Gross. 1975. *Méthodes en syntaxe*. Hermann, Paris.

[Guillet and Leclère1992] A. Guillet and Ch. Leclère. 1992. *La structure des phrases simples en français. 2 : Constructions transitives locatives*. Droz, Geneva.

[Kuznetsov2007] Sergei O. Kuznetsov. 2007. On stability of a formal concept. *Annals of Mathematics and Artificial Intelligence*, 49(1-4):101–115.

[Levin1993] Beth Levin. 1993. *English Verb Classes and Alternations: a preliminary investigation*. University of Chicago Press, Chicago and London.

[Priss2005] Uta Priss. 2005. Linguistic Applications of Formal Concept Analysis. In Bernhard Ganter, Gerd Stumme, and Rudolf Wille, editors, *Formal Concept Analysis*, volume 3626 of *Lecture Notes in Computer Science*, pages 149–160–160. Springer Berlin / Heidelberg.

[Schuler2006] Karin Kipper Schuler. 2006. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. Ph.D. thesis, University of Pennsylvania.

[Valverde-Albacete2008] Francisco J. Valverde-Albacete. 2008. Extracting Frame-Semantics Knowledge using Lattice Theory. *Journal of Logic and Computation*, 18(3):361–384, June.

[van den Eynde and Mertens2003] Karel van den Eynde and Piet Mertens. 2003. La valence: l'approche pronominale et son application au lexique verbal. *Journal of French Language Studies*, 13:63–104.

Syntactic Characteristics of Particle Verbs: Empirical Evidence for Complex Predicate Processing in German

Simone Falk
Ludwig-Maximilians-Universität München, Germany.
Simone.falk@germanistik.uni-muenchen.de

Peter Öhl
Ludwig-Maximilians-Universität München, Germany
lmu@oehl.gesus-info.de

Abstract

In this study, we explore empirical aspects of the processing of particle verbs in German from a syntactic-topological point of view. Particle verbs in German form a very heterogeneous group and have given rise to a long-term and still ongoing linguistic debate. The particle and the verb split apart in main clauses, but movement of the particle seems to be very restricted – especially in the middle field – compared to normal phrasal constituents in a sentence. In a series of experiments, we tested empirically the conditions allowing particles to appear in different positions in a sentence, contrasting them to phrasal constituents. In one of these experiments, we tested this difference by applying a self-paced reading paradigm combined with acceptability judgements. Results show that semantically transparent adjectival particles behave differently from phrases both in terms of acceptability ratings and particle/phrasal constituent reading times. Overall, particles occurring in the middlefield are judged as rather non-acceptable compared to phrases. In the prefield, however, readers seem to be uncertain about the status of the particles, judging them as more acceptable than in the middlefield but less acceptable than in the default position and showing longer reading times for particles than in the other topological positions. This underlines the special status of particle verbs and the need of syntactic description of this verbal class.

1 Introduction

This paper has as its subject the group of verbs in Germanic languages that are known as particle verbs (for German cf. Lüdeling 2001; Zeller 2001; Heine & al. 2010, Dutch: Booij 2002; English: Olsen 1998; McIntyre 2001). They consist of at least two parts, a particle and a full verb, where the particle can be of any word class, i.e. prepositional (***unter****gehen*), nominal (***kopf****rechnen*), adjectival (***schön****reden*), adverbal (***zusammen****kommen*) or verbal (***kennen****lernen*). These verbs are known to be problematic for a systematic morphosyntactic analysis. They have lexical entries as complex constructions, their special characteristics, however, cannot be described but in terms of syntactic distribution. The relevant features of particle verbs are (Oehl 2009):

(1) In contrast to morphologically complex verbs like prefixed verbs, the particle separates from the verb in V2-sentences:

Sie ***boten*** *dort Waren* ***an****.*
they offered there goods PTC

(2) In contrast to phrasal constituents like the directional adverbial in (b.) below, particles seem to be blocked from movement within the middle field of a sentence. This is the topological position that spans from the finite verb to the end of the verbal bracket in main clauses or from the conjunction to the end of the verb complex.

...weil sie* ***ein* *nie zuvor gereist sind*
*...because they in(PTC) never before travel-PART-PERF AUX

(3) However, in all kinds of texts, we find particles that appear in the prefield (the position before the finite verb in V2-sentences; Heine & al. 2010; Zeller 2001) as in

Auf *ging die Tür,* ***zu*** *ging das Fenster*
open(PTC) went the door, closed(PTC) went the window)

Thus, as a rule, these particles are heads forming a syntactically complex predicate together with the main verb. Only in specific cases like (3) above, however, they can be positioned in the prefield. To our knowledge, no empirical evidence has been given showing how permutation of particles is constrained, that is to say, evaluated by listeners in online processing. What we suppose on the grounds of the state of the art and what has to be evaluated by empirical testing is that this is determined by three criteria (Oehl, 2009). These are:

- Phrasal status: A particle can be analysed as heading a phrase in that position.
- Semantic transparency: the meaning of both components is transparent, such that the particle verb can be decomposed without losing its referential properties.
- Discourse semantic markedness: movement of the particle implies focus or contrastiveness.

Since common phrasal constituents like adverbials or depictive secondary predicates are necessarily semantically transparent and can be permuted with less restrictions, we started out exploring the empirical correlates of the distinction between such autonomous phrases and particles belonging to the complex predicate. In order to find out whether reading times reflect the distinction or/and are correlated with the (grade of) acceptance, we used a self-paced reading paradigm in combination with an acceptability judgement task.

2 Experiment

In the following experiment, we are testing two hypotheses:
(1) Phrasal constituents should be more acceptable in fronted positions within the middle field than particles which – in linguistic descriptions – are blocked from that position. Phrases are less restricted in their ability to move throughout the middlefield, Overall, higher processing costs are predicted for the middlefield vs. default position in either phrasal constituent or particle condition (Bader & Meng, 2000; Bader, Meng & Bayer, 1999; Bornkessel, Schlesewsky, 2006).
(2) Concerning movement to the prefield position, we do not expect differences between particles and phrasal constituents as long as the criteria above , i.e. semantic transparency and phrasal status, apply (cf. Oehl: 2009).

2.1 Method

2.1.1 Material

The particles used in the first experiment are adjectival and semantically transparent. They combine with the positional verb *halten* ('hold'). This verb was chosen because it can be used with a modal adverbial (phrasal constituent) instead of the particle. 16 particles were used in this experiment: *warm, hoch, still, feucht, dicht, frisch, wach, rein, frei, bereit, gesund, sauber, ruhig, heilig, trocken, geheim*[1]. 16 adjectival modal adverbials were: *vorsichtig, lässig, zitternd, mühelos, mühsam, achtsam, behutsam, unbeholfen, zärtlich, sanft, stolz, liebevoll, ungeschickt, geduldig, widerwillig, lustlos*[2]. Test sentences were created by permutating phrasal constituents and particles between the following positions in the sentence:

- Base/Default Position (Df): Ich habe die Fahne **hoch** (PTC)/**stolz** (adverbial) gehalten.
  I – have – the – flag – ***high/proudly*** *– held*
- Middlefield Position (Mf): *Ich habe* ***hoch/stolz*** *die Fahne gehalten.*
- Prefield Position (Pf): ***Hoch/Stolz*** *habe ich die Fahne gehalten.*

We created 16 sentences that we tested on the basis of two factors (Syntactic Status, Topology) with 2 (particle, phrase) and 3 conditions (Def, Mf, Pf) respectively. Overall, 96 test sentences resulted from that. Furthermore, we added 102 filler sentences. The material was organised in 4 lists, each containing 150 sentences in three blocks of 50 sentences which were randomised per participant. Each participant was tested on 48 test sentences (only one sentence per topological condition) to avoid repetition priming. In a pretest, two people judged the default sentences for their semantic comprehensability.

2.1.2 Procedure

We used a word-by-word self-paced reading paradigm (moving window, Just et al., 1982) using the DMDX software. Participants read sentences word-by-word in their own reading speed on a monitor. They pressed a key to get each word of the sentence. The word was uncovered by the key press and by the next press, it was again replaced by dashes. The key press after the last word of the sentence either revealed a ques-

[1] Translation: warm, high, calm, humid, leak-proof, fresh, awake, clean, free, ready, sane, neat, quiet, holy, dry, secret.

[2] Translation: carefully, casually, trembling, effortlessly, drudgingly, attentively, cautiously, awkwardly, tenderly, gently, proudly, lovingly, clumsily, patiently, grudgingly, half-heartedly.

tion that asked the participant to judge the acceptability of the previously read sentence (4-point scale: 1 = very acceptable, 2 = acceptable, 3 = less acceptable, 4 = not acceptable) or the next test trial. Before the test started, participants went through 12 training sentences.

2.1.3 Participants

32 students of the university of Munich (LMU) participated in the experiment.

2.2 Results

2.2.1 Acceptability judgements

A repeated measures ANOVA (subject-based) with the Factor Topology and dependent variable Acceptability rating was conducted overall and for each Syntactic Status condition separately. We found significant differences in the particle condition, but not in the phrasal constituent condition (s. figure 1). In the particle condition ($F=128.51$, $df = 2$, $p < 0.001$) all three topological positions differed significantly from one another with the lowest acceptability rates for the middlefield.

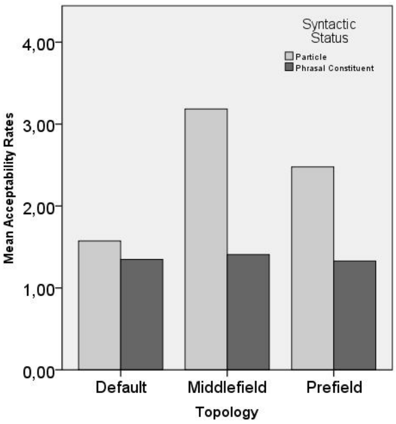


Figure 1: Mean acceptability rates according to Syntactic Status.

Furthermore, in the Particle condition, some of the 16 items showed deviant acceptability rates (see Fig. 2) which resembled the results for phrasal constituents for the particle *still* ('calm') with very high acceptability but non-significant throughout the topological conditions, furthermore *gesund* ('healthy') and *heilig* ('holy') had non-significant and very low acceptability rates.

2.2.2 Reading Times

The reading times of each word of the sentence were cumulated to a whole sentence reading time. Particle/phrasal constituent reading times and verb reading times were also taken into account. First, a repeated measures ANOVA (Factors: Topology, Syntactic Status) was performed for the variables overall reading time and particle/phrasal constituent and verb reading time. No significant effects were found for overall reading times and verb reading times. However, particle/phrasal constituent reading times differed according to topological position ($F=3.225$, $df=2$, $p < 0.05$).

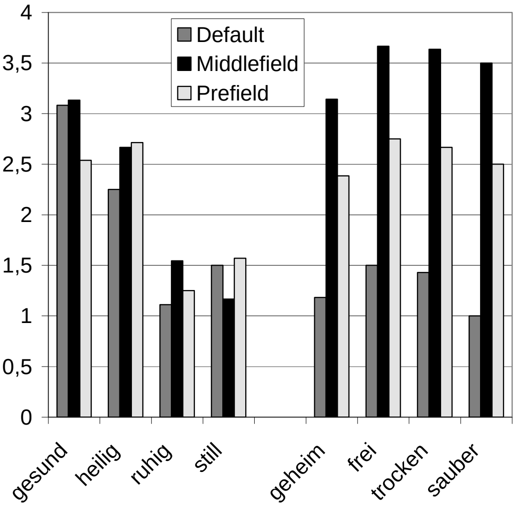


Figure 2: Deviant acceptability ratings of the four particles *gesund*, *heilig*, *ruhig*, *still* contrasted to four particles (*geheim*, *frei*, *trocken*, *sauber*), following the general response pattern.

A repeated measures ANOVA (Factor: Topology) was conducted for each Syntactic Status condition separately and revealed the following highly significant effects: in the particle condition particle reading times differed significantly ($F=6.35$, $df=2$, $p < 0.005$), pairwise comparison of the topological positions shows that this is due to significantly longer reading times for the particle in the prefield (mean = 448 ms) compared to middlefield (mean = 390 ms) and default (mean = 393 ms), but no difference was found between middlefield and default position. In the phrasal constituent condition, the constituent reading times show a different picture: the overall significant difference between the topological positions ($F = 3.91$, $df = 2$, $p = 0.05$) is due to significantly longer reading times of the phrasal constituent in the middlefield (mean = 475 ms) compared to the default (mean = 398 ms) and prefield (mean = 421 ms).

3 Discussion and Conclusions

Results confirm parts of our first hypothesis, namely that particles are less to non-acceptable when fronted in the middlefield compared to phrasal constituents that are highly acceptable in this position. However, only phrasal constituents – as already shown in the literature – and not particles showed significantly longer reading times in this position. We hypothesize that this is the case because readers, in the specific task of this experiment, instantly evaluate the non-acceptability of a particle when encountering it in the middlefield. Our second hypothesis that the prefield condition should be comparable for both particles and phrasal constituents was not confirmed. Whereas phrasal constituents were highly acceptable in this position, particles were yet judged better than in the middlefield but less acceptable than in the default position. Furthermore, reading times for particles in the prefield increase significantly compared to the default or middlefield position. One reason for this pattern could be that readers are somehow uncertain about the status of the particles in the prefield. This might be due to the fact that in this experiment, one of the criteria for fronting of particles to the prefield (Oehl, 2009) – discourse semantic markedness – might not have been as obvious to the readers as it should be for full acceptability of particles in the prefield. We think that emphasis of potential discourse semantic features of those particles (focalizing, contrasting with other particles, addition of more contextual information) might further improve their acceptability, as their phrasal status would be more obvious. This has to be left for further testing.

Some particles showed a pattern in acceptability rates divergent from that of other particles. The particle *still* ('calm') for example which had overall high acceptability rates in all topological positions seems to be ambiguous: it follows the phrasal constituent pattern and seems to be interpreted as a modal adverbial in those positions that were disprefered for a particle reading. The different pattern found with the particle verbs *gesund halten* and *heilig halten* may be due to the fact that they are relatively infrequent and therefore might have been judged as rather unacceptable due to non-familiarity with the construction. Thus, frequency and semantic content of individual particle verbs should be considered in further testing.

One restriction of the experiment was the exclusive use of the verb *halten* 'hold', which may be semantically bleached if combined with particles. We therefore conducted another experiment (which we have not the space to present here, but see Oehl & Falk (forthcoming), where we included additional verbs with semantically more concrete meanings as *trinken, bügeln, binden, schlagen, klopfen, kochen*[3]. This experiment replicated the results described above in terms of acceptability ratings in the different topological positions found with the verbal base *halten*. We therefore think that our results are valid and generalizable to the whole class of particle verbs. To conclude, this study shows that particle reading times combined with acceptability judgements constitute empirical correlates that can be employed for further testing of syntactic characteristics of particle verbs. The topological options distinguish the verbal particles from common phrasal constituents. This underlines their special status as parts of syntactically complex predicates and the need for a differentiated syntactic description of the class of particle verbs.

References

Booij, Geert (2002): Separable Complex Verbs in Dutch: A Case of Periphrastic Word Formation. In: Nicole Dehé/Ray Jackendoff/Andrew McIntyre/Silke Urban (Hg.): *Verb-Particle Explorations.* Berlin/New York: De Gryuter. (= Interface Explorations 1), 21–41.

Bornkessel, Ina, Matthias Schlesewsky. 2006. The Role of Contrast in the Local Licensing of Scrambling in German: Evidence from Online Comprehension. *Journal of Germanic Linguistics* 18 (1), 1–43.

Bader, Markus & Michael Meng, Michael. 1999: Subject-Object Ambiguities in German Embedded Clauses: An Across-the-Board Comparison. *Journal of Psycholinguistic Research* 28 (2), 121-143.

Just, M. A., P.A. Carpenter & J.D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111: 228-238.

Luedeling, Anke. 2001. *On Particle Verbs and Similar Constructions in German.* Stanford: CSLI.

McIntyre, Andrew. 2001. Argument Blockages Induced by Verb Particles in English and German: Event Modification and Secondary Predication. In Nicole Dehé & al. (eds.), *Structural aspects of semantically complex verbs.* Frankfurt am Main (etc.): Lang. 131-164.

Meng, Michael, Markus Bader & Josef Bayer. 1999. Die Verarbeitung von Subjekt-Objekt Ambiguitäten im Kontext. *Proceedings der 4. Fachtagung der Gesellschaft für Kognitionswissenschaften*, Hg v. Ipke Wachsmuth & B. Jung, St. Augustin: Infix Verlag. 244–249.

[3] Translation: drink, iron, bind, beat, knock, cook.

Oehl, Peter. 2009. Zur Distribution der so genannten 'Verbpartikeln' im Deutschen. In: Cseresznyák, Mónika & Szatmári, Petra (eds.): *Zwischen-Bilanz. 20 Jahre Germanistik in Szombathely*. Szombathely/Wien: Savaria University Press/Praesens Verlag (= Acta Germanistica Savariensia 11). 61-78.

Oehl, Peter, Simone Falk (forthcoming). Syntactic competence and performance-based variation: the case of German particle verbs. Michael Putnam (ed.), *Leuvense Bijdragen, Special issue: Syntax and Semantics of Particle Verbs.*

Olsen, Susan. 1998. Prädikative Argumente syntaktischer und lexikalischer Köpfe: Partikelverben im Deutschen und Englischen. *Folia Linguistica* 31, 301-329.

Zeller, Jochen. 2001. *Particle verbs and local domains.* Amsterdam: John Benjamins.

Aggregating Entries of Semantic Valence Dictionary of Polish Verbs

Elżbieta Hajnicz
Institute of Computer Science, Polish Academy of Sciences
ul. Ordona 21, 01-237 Warsaw, Poland
hajnicz@ipipan.waw.pl

Abstract

In this paper the phase of semantic valence dictionary of Polish verbs consisting in aggregating entries to semantically coherent sets is presented. Two methods: a simple agglomerative one and minimal spanning trees method are discussed and compared. Both methods use a predefined similarity measure of semantic frames.

1 Introduction

The primary task of our research is to create a semantic valence dictionary in an automatic way. To accomplish this goal, the valence dictionary of Polish verbs is supplemented with semantic information, provided by wordnet's semantic categories (Hajnicz, 2009d; Hajnicz, 2009c) or synsets (Hajnicz, 2009a) of nouns. In our present work we focus on arguments taking form of nominal phrases NPs and prepositional-nominative phrases PrepNPs, whose semantic heads are nouns. We discuss the case of 26 predefined semantic categories of nouns, which is simpler than the case of actual wordnet synsets. In the current phase of work we want to discuss in this paper, we have in our disposal two resources:

- purely syntactic valence dictionary,
- a syntactically and semantically annotated corpus.

In theory, it is not important whether these resources were prepared manually or automatically. In practice, the difference is quite significant, because errors obtained during automated data processing are cumulated.

Typical approaches, e.g., VerbNet (Dang et al., 1998) or VerbaLex (Hlaváčková and Horák, 2006), consider one strongly preferred sense per argument. In contrast, we present a solution in which all appropriate senses are aggregated.

2 Data resources

We used an extensive valence dictionary based on Świdziński's (1994) valence dictionary containing 1064 verbs. It was specially modified for our task. Świdziński's dictionary was supplemented with 1000 verb entries from the dictionary automatically obtained by Dębowski and Woliński (2007) to increase the coverage of used dictionary on SEMKIPI (cf. below). The most carefully elaborated part of the valence dictionary concerns the set of 32 verbs manually chosen for the experiments (Hajnicz, 2009c). They were chosen manually in order to maximise the variability of their syntactic frames (in particular, diathesis alternations) on one hand and the polisemy within a single frame on the other. Their frequency was the important citerion for this choice as well.

A syntactic dictionary $\mathcal{D}$ is a set of entries representing schemata for every verb considered. Formally, $\mathcal{D}$ is a set of pairs $\langle v, g\rangle$, where $v \in V$ is a verb and $g \in G$ is its syntactic schema. Below we list syntactic dictionary entries for verb *interesować* (*to interest*). np:case are nominal phrases, sentp:wh are wh-clauses, whereas sie is a reflexive marker.

(1) interesować np:acc np:nom
interesować np:inst np:nom sie
interesować np:nom sentp:wh sie

The main resource used in our experiments was the IPI PAN Corpus of Polish written texts (Przepiórkowski, 2004). A small subcorpus was selected from it, referred to as SEMKIPI containing 195 042 sentences predicated by chosen verbs. SEMKIPI was parsed with the *Świgra* parser (Woliński, 2004) based on the metamorphosis grammar GFJP (Świdziński, 1992) provided with the valence dictionary presented above.[1] The complete frequency list of verbs in the IPI PAN Corpus contains about 15 000 verbs, with 12 000 of them occurring at least 5 times. Grammatical dictionary of Polish (Saloni et al., 2007) lists 29 000 verbs.

In order to reduce data sparseness, in the present experiment we considered only the topmost phrases being the actual arguments of a verb (i.e., a subject and complements included in its valence schemata). This means that each obtained

[1] In particular, the parser links genitive of negation with accusative in the corresponding valence schema.

parse was reduced to its “flat” form identifying only these top-most phrases. Semantic annotation concerning verb argument heads only was based on the Polish WordNet (Derwojedowa et al., 2007; Derwojedowa et al., 2008a; Derwojedowa et al., 2008b; Piasecki et al., 2009). The Polish WordNet is a network of lexical-semantic relations modelled on the Princeton WordNet (Fellbaum, 1998) and wordnets constructed in the EuroWordNet project (Vossen, 1998).

3 Semantic valence protodictionary

The process of collecting a semantic valence protodictionary on the basis of SEMKIPI for semantic categories was described in (Hajnicz, 2009b).

Formally, a semantic protodictionary $\mathfrak{D}$ is a set of tuples $\langle\langle v, g, f\rangle, n_g, m_f\rangle$, where $\langle v, g\rangle \in \mathcal{D}$ is a schema of a verb, $f \in F_g$ is one of its semantic frames, n_g is the frequency of $\langle v, g\rangle$ and m_f is the frequency of $\langle v, g, f\rangle$. A frame is a list of arguments, among which only NPs and PrepNPs are semantically interpreted, i.e., supplied with semantic categories $c \in C$.

An exemplary subset of the set of frames connected with the schema `np:acc np:dat np:nom` of the verb *proponować* (*to propose*) is shown in (2). In the second column the frequencies of frames are given.

(2) proponować

np:acc np:dat np:nom	573
np:acc: act; np:dat: person; np:nom: person	51
np:acc: act; np:dat: group; np:nom: person	50
np:acc: act; np:dat: act; np:nom: person	31
np:acc: act; np:dat: person; np:nom: group	22
np:acc: act; np:dat: group; np:nom: group	16
np:acc: act; np:dat: location; np:nom: person	9
np:acc: act; np:dat: act; np:nom: group	8
np:acc: act; np:dat: feeling np:nom: group	4
np:acc: act; np:dat: group; np:nom: event	1

4 The process of aggregation

A protodictionary has plenty of entries (simple semantic frames), with a single category assigned to each syntactic slot. This does not reflect the actual semantics of a verb, since different categories of arguments do not entail different meanings of the verb. In other words, such classification is too fine-grained. For instance in sentences (3) we have different meanings of the verb *przejechać*. These differences are reflected in different English translations of the verb: *to cross* in the first sentence and *to run over* in the second. Hence, we want to have two different entries for it in the valence dictionary, with location and animal on the object position, correspondingly. On the other hand, in sentences (4) we deal with the same meaning of the verb *kupić* (*to buy*), and we want to have one entry for it. In order to differentiate these situations we defined a similarity measure d between two categories. Its value varies from 1 to 6 for two “neighbouring” categories. The similarity measure between semantic categories is presented in Figure 1 in a form of graph in which nodes represent categories. $d(c_1, c_2)$ is the shortest path linking categories c_1 and c_2, interpreted as a sum of edges labels.[2]

Usage of the measure is based on the assumption that two categories are put together only if all categories located in between by means of a particular similarity measure occur at a considered slot of a schema as well. Observe that one can buy almost everything, in particular things having semantic categories positioned in between animal and location (in particular, food, substances, artifacts, some physical objects and groups of things, cf. Figure 1). Contrary, objects of *crossing* and *running over* are separated.

Synsets for which there is not a path in hiponymy relation and that are not top ones are not similar by definition.

(3) *Piotr*$_{\text{person}}$ *przejechał park*$_{\text{location}}$ *samochodem*$_{\text{artifact}}$.
(*Piotr cross his park in a car.*)
Piotr$_{\text{person}}$ *przejechał psa*$_{\text{animal}}$ *samochodem*$_{\text{artifact}}$.
(*Piotr run over his dog by a car.*)

(4) *Piotr*$_{\text{person}}$ *kupił bratu*$_{\text{person}}$ *park*$_{\text{location}}$.
(*Piotr bought his brother a park.*)
Piotr$_{\text{person}}$ *kupił bratu*$_{\text{person}}$ *psa*$_{\text{animal}}$.
(*Piotr bought his brother a dog.*)

Thus, we want to aggregate simple frames into compound ones, in which every syntactic slot is supplied with a list of semantic categories. A compound frame is supposed to determine a single meaning of a verb. To obtain this, we have applied two clustering methods. Both are based on a similarity measure between frames D_n, where n is a space dimension (number of NPs/PrepNPs). D_n is defined on the basis of similarity measure between categories d applied for all NPs/PrepNPs in Euclidean way. Namely,

$$D_n(f^A, f^B) = \sqrt{\sum_{i=1}^{n}(d(c_i^A, c_i^B))^2}$$

for $g = \langle r_1, \ldots, r_n\rangle$ and $f^A = \langle\langle r_1, c_1^A\rangle, \ldots, \langle r_n, c_n^A\rangle\rangle$, $f^B = \langle\langle r_1, c_1^B\rangle, \ldots, \langle r_n, c_n^B\rangle\rangle$.

The first method is a simple agglomerative method (Aggl) based on choosing the most frequent simple frame and joining it with other elements of a compound frame under creation that

[2]Please note that the graphical composition of a picture is not meaningful; in particular, the length of arcs is not proportional to the actual distance between nodes. Observe that the measures are not 2D, there are only visualised on a plane.

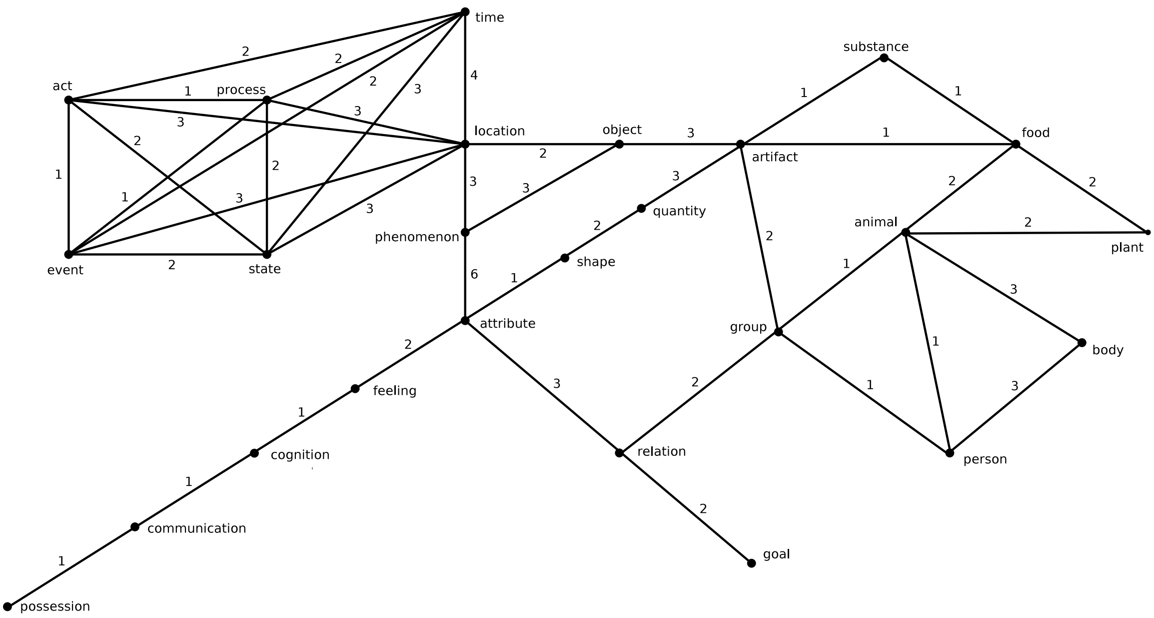


Figure 1: Similarity measure between semantic categories

(5) proponować np:acc np:dat np:nom				573
	acc: act,event,place, state,time;	dat: cognit.,communic.,feel.,group, person,poss.,quality,relation;	nom: group,person, relation	264
	acc: act,place,state;	dat: act,event,place,state,time;	nom: group,person	105
	acc: cognit.,communic.,feel.,group, person,poss.,quality,relation;	dat: group,person;	nom: group,person, relation	49
	acc: act;	dat: artifact;	nom: group,person	22
	acc: act;	dat: group,person;	nom: artifact	8
	acc: act,event;	dat: group;	nom: act,event	7
	acc: act;	dat: quantity;	nom: group	5
	acc: act;	dat: act;	nom: artifact	4
	acc: act;	dat: cognit.;	nom: artifact	2
	acc: act;	dat: quality;	nom: artifact	2
	acc: act;	dat: quantity;	nom: artifact	2
	acc: act;	dat: person;	nom: quantity	1

are "sufficiently" similar, i.e., D_n does not exceed a particular threshold ρ^A.

A fragment of the aggregated dictionary $\tilde{\mathfrak{D}}$ for the schema `np:acc np:dat np:nom` of the verb *proponować* (*to propose*) is shown in (5).

The second method is a popular clustering method based on similarity measure called *minimal spanning trees* (MST) proposed by Zahn (1971). The algorithm was performed for each verb schema independently. Simple frames represented graph nodes, and edges were labelled with distances defined by D_n. The heuristics for determining threshold used for removing outlying edges $\rho_{\langle v,g \rangle}$ was based on local criteria (the median $\mu_{\langle v,g \rangle}$ and q's percentile $\Phi^q_{\langle v,g \rangle}$ of a distribution of lengths of edges between frames of a particular syntactic schema) and global criteria (the median μ_n and q's percentile Φ^q_n of a distribution of lengths of edges between frames of all syntactic schemata with n NPs/PrepNPs). Namely,

$$\rho^q_{\langle v,g \rangle} = \max(\mu_n, \mu_{\langle v,g \rangle}, \min(\Phi^q_n, \Phi^q_{\langle v,g \rangle})).$$

Medians ensure that too short edges will not be cut, percentiles ensure that too long edges will not stay.

5 Experiments

The experiments were performed with $\rho^A = 2$ for agglomerative method and percentiles $q = 80, 90$ for MST. Observe that the greater ρ^A (or the higher q) the larger compound frames are obtained.

5.1 Manually prepared semantic dictionary

$\mathfrak{D}^H$ differs from $\tilde{\mathfrak{D}}$ in that it has no frequencies assigned to frames. Moreover, it is rather exhaustive, i.e., frames contain all corresponding semantic categories of slots. This means that such a dictionary should be interpret in a manner of selectional restrictions rather than selectional preferences (Resnik, 1993). $\mathfrak{D}^H$ was prepared independently from corpus data. Thus, it contains simple frames having no counterparts in $\mathfrak{D}$ (and

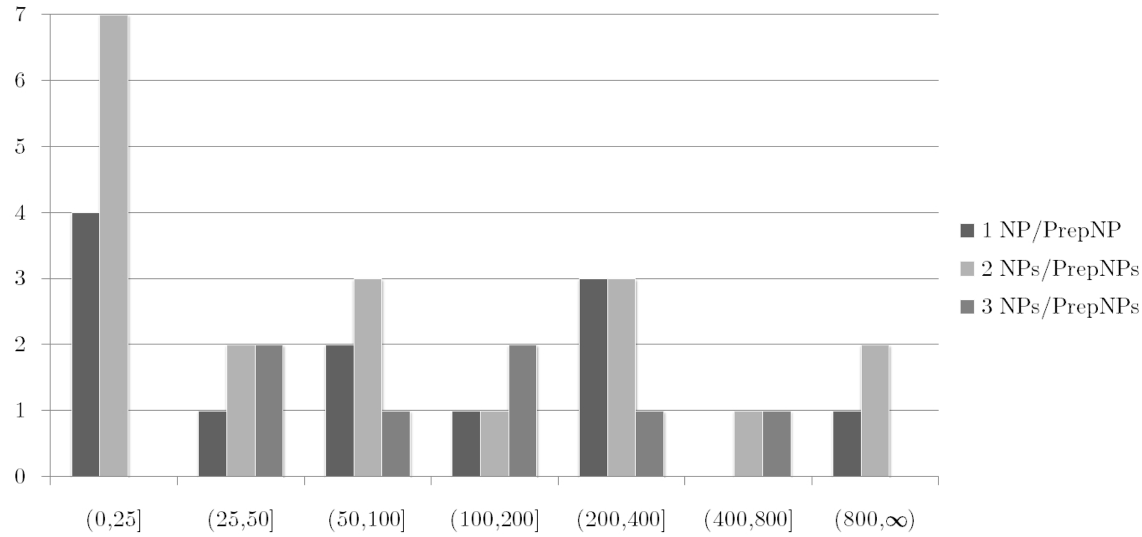


Figure 2: Frequencies of schemata from $\mathfrak{D}^H$ in $\mathfrak{D}$

SEMKIPI), because of sparseness of data. On the other hand, due to data processing errors of SEMKIPI (Hajnicz, 2009d; Hajnicz, 2009c), some frames from $\mathfrak{D}$ are absent in $\mathfrak{D}^H$.

The results were validated w.r.t. a small manually prepared semantic dictionary $\mathfrak{D}^H$ composed of all syntactic schemata and corresponding compound semantic frames for 5 verbs: *interesować* (*to interest*: 3 schemata), *minąć* (*to pass*: 5 schemata), *proponować* (*propose* 10 schemata), *rozpocząć* (*to begin*: 8 schemata) and *widzieć* (*to see*: 13 schemata), which gives total number of 39 schemata. These verbs were selected from the set of 32 ones considered in SEMKIPI in a manner maximising their syntactic diversity. The frequency was not a criterion for this choice. However, since the process of aggregation is performed for each syntactic schema separately, their frequency is more important to validate the process. We should also remember that the task complexity depends on the number of NPs/PrepNPs in the schema. In $\mathfrak{D}^H$ there are 12 schemata with 1 NP/PrepNP, 19 schemata with 2 NPs/PrepNPs and 8 schemata with 3 NP/PrepNP. Their frequencies in $\mathfrak{D}$ are given in Figure 2. The Figure shows that frequencies of schemata are sufficiently differentiated.

5.2 Validation

There exist three popular clustering validation methods based on co-occurrence of two elements (simple frames) in two partitions of a particular data set. Let

- b be the number of pairs co-occurring in both sets,
- c be the number of pairs co-occurring only in the validated set ($\tilde{\mathfrak{D}}$),
- g be the number of pairs co-occurring only in the gold standard ($\mathfrak{D}^H$),
- n be the number of pairs co-occurring in neither of sets.

Then *Rand statistics* (R), *Jaccard coefficient* (J) and *Folkes and Mallows index* (FM) are given by the equations (Halkidi et al., 2001):

$$\begin{aligned} R &= \frac{b+n}{b+c+g+n}, \\ J &= \frac{b}{b+c+g}, \\ FM &= \frac{b}{\sqrt{b+c}\sqrt{b+g}}. \end{aligned}$$

Rand statistics resemble in a way accuracy measure used in typical lexical acquisition tasks. With such point of view, Jaccard Coefficient and Folkes and Mallows index could be interpret as counterparts of combinations of precision and recall.

In order to apply them to our data ($\tilde{\mathfrak{D}}$ and $\mathfrak{D}^H$), we need to bear in mind the specificity of the problem of aggregating semantic dictionary. First, instead of a one large set of data we have plenty of verb syntactic schemata, which frames are aggregated separately. Their validation may be calculated cumulatively or in average. Moreover, there exist some "lonely" frames properly not aggregated with any other frames. In order to take into account such frames (single-element clusters) we consider obvious co-occurrence with itself. Next, the partitioned data sets are different (even though overlapping). Because of that we have counted the above indexes both for all simple frames ($\bigcup$) and for the ones belonging to both dictionaries ($\bigcap$).

		average			cumulative		
		R	J	FM	R	J	FM
$\bigcup$	Aggl	77.6	26.7	40.1	83.6	9.7	17.8
	MST-80	73.4	22.3	35.2	79.3	5.6	10.8
	MST-90	63.6	19.5	30.1	67.7	2.8	8.1
$\bigcap$	Aggl	91.3	86.4	91.8	82.8	69.9	82.3
	MST-80	87.9	82.6	89.1	77.7	59.7	75.0
	MST-90	83.3	78.0	86.5	66.5	55.0	73.9
hand	Aggl	87.5	73.3	82.9	92.6	68.8	81.5
	MST-80	75.2	51.4	66.8	82.9	17.2	39.4
	MST-90	77.8	57.2	71.5	83.3	20.9	42.7

Table 1: Validation of aggregation of frames

The results of the validation are presented in Table 1. They show that the best results are obtained for the agglomerative method. The results are mostly better for frames belonging to both dictionaries than for frames belonging to any of them, which is the obvious consequence of the indexes being used: a frame belonging only to one dictionary cannot co-occur with any frame in the second dictionary.

The improvement of Rand statistics calculated cumulatively w.r.t. the one calculated in average indicates the influence of a proportionally large value of n for large schemata.[3] The deterioration of Jaccard coefficient and Folkes and Mallows index calculated cumulatively w.r.t. the one calculated in average indicates the influence of a proportionally large values of c and g. Observe that the larger indexes are the smaller is the difference between cumulative and average method of calculating them.

In order to validate the actual methods without any influence of the corpus preprocessing, we applied the algorithms to $\mathfrak{D}^H$ distributed back to protodictionary. The results of validation for this case are denoted in Table 1 as *hand*. The superiority of the agglomerative method is in this case even more apparent.

The fact that the results are better for agglomerating $\mathfrak{D}$ calculated for intersection of dictionary than for redistributed and re-agglomerated $\mathfrak{D}^H$ is a bit surprising. This makes an impression that false simple frames help to agglomerate proper ones. The possible reasons for this could be errors in the similarity measure definition or in the preparation of $\mathfrak{D}^H$. However, the most probable explanation of this fact is that simple frames belonging to both dictionaries are most "obvious", "natural" ones and hence they are easier to agglomerate. Simple frames belonging only to $\mathfrak{D}^H$ are rare and "unusual", and hence they harder to agglomerate. The small size of $\mathfrak{D}^H$ could influence the results as well.

[3] Schemata with a large number of simple frames are called "large".

6 Conclusions

In this paper two methods of aggregating simple semantic frames into semantically coherent compound ones were discussed and compared. The fact that a simple agglomerative method was better than MST is indication to apply more sophisticated agglomerative methods.

We also plan to extend $\mathfrak{D}^H$, which will enable us to perform the more reliable validation. In particular, the validation w.r.t. the number of NPs/PrepNPs in a schema and/or the number of simple frames in it will be possible, which is disabled by the present small size of $\mathfrak{D}^H$.

References

Hoa Trang Dang, Karin Kipper, Martha Palmer, and Joseph Rosenzweig. 1998. Investigating regular sense extensions based on intersective Levin classes. In *Proceedings of the COLING-ACL'98 Conference*, pages 293–299, Montreal, Canada.

Magdalena Derwojedowa, Maciej Piasecki, Stanisław Szpakowicz, and Magdalena Zawisławska. 2007. Polish WordNet on a shoestring. In *Data Structures for Linguistic Resources and Applications: Proceedings of the GLDV 2007 Biannual Conference of the Society for Computational Linguistics and Language Technology*, pages 169–178, Universität Tübingen, Tübingen, Germany.

Magdalena Derwojedowa, Maciej Piasecki, Stanisław Szpakowicz, Magdalena Zawisławska, and Bartosz Broda. 2008a. Words, concepts and relations in the construction of Polish WordNet. In Attila Tanacs, Dora Csendes, Veronica Vincze, Christiane Fellbaum, and Piek Vossen, editors, *Proceedings of the Global WordNet Conference*, pages 162–177, Seged, Hungary.

Magdalena Derwojedowa, Stanisław Szpakowicz, Magdalena Zawisławska, and Maciej Piasecki. 2008b. Lexical units as the centrepiece of a wordnet. In Mieczysław A. Kłopotek, Adam Przepiórkowski, and Sławomir T. Wierzchoń, editors, *Proceedings of the Intelligent Information Systems XVI (IIS'08)*, Challenging Problems in Science: Computer Science, Zakopane, Poland. Academic Publishing House Exit.

Łukasz Dębowski and Marcin Woliński. 2007. Argument co-occurrence matrix as a description of verb valence. In Zygmunt Vetulani, editor, *Proceedings of the 3rd Language & Technology Conference*, pages 260–264, Poznań, Poland.

Christiane Fellbaum, editor. 1998. *WordNet — An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Elżbieta Hajnicz. 2009a. Generalizing the EM-based semantic category annotation of NP/PP heads to wordnet synsets. In Zygmunt Vetulani, editor, *Proceedings of the 4th Language & Technology Conference*, pages 432–436, Poznań, Poland.

Elżbieta Hajnicz. 2009b. Problems with pruning in automatic creation of semantic valence dictionary for polish. In Václav Matoušek and Pavel Mautner, editors, *Proceedings of the International Conference on Text, Speech and Dialogue TSD 2009*, volume 5729 of *LNAI*, pages 131–138, Berlin, Heidelberg. Springer-Verlag.

Elżbieta Hajnicz. 2009c. Semantic annotation of verb arguments in shallow parsed Polish sentences by means of EM selection algorithm. In Małgorzata Marciniak and Agnieszka Mykowiecka, editors, *Aspects of Natural Language Processing*, volume 5070 of *LNCS*, pages 211–240. Springer-Verlag.

Elżbieta Hajnicz. 2009d. Towards extending syntactic valence dictionary for Polish with semantic categories. In Gerhild Zybatow, Uwe Junghanns, Denisa Lenertová, and Petr Biskup, editors, *Studies in Formal Slavic Phonology, Morphology, Syntax, Semantics and Information Structure*, pages 279–290, Frankfurt am Main. Peter Lang.

Maria Halkidi, Yannis Batistakis, and Michalis Vazirgiannis. 2001. On clustering validation techniques. *Journal of Intelligent Information Systems*, 17(2/3):107–145.

Dana Hlaváčková and Aleš Horák. 2006. Verbalex — new comprehensive lexicon of verb valences for Czech. In *Proceedings of the Third International Seminar on Computer Treatment of Slavic and East European Languages*, pages 107–115, Bratislava, Slovakia.

Maciej Piasecki, Stanisław Szpakowicz, and Bartosz Broda. 2009. *A Wordnet from the Ground Up*. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.

Adam Przepiórkowski. 2004. *The IPI PAN corpus. Preliminary version*. Institute of Computer Science, Polish Academy of Sciences, Warsaw.

Philip Resnik. 1993. *Selection and Information: A Class-Based Approach to Lexical Relationships*. Ph.D. thesis, University of Pennsylvania, Philadelphia, PA, December.

Zygmunt Saloni, Włodzimierz Gruszczyński, Marcin Woliński, and Robert Wołosz. 2007. *Słownik gramatyczny języka polskiego*. Wiedza Powszechna, Warsaw.

Marek Świdziński. 1992. *Gramatyka formalna języka polskiego*. Rozprawy Uniwersytetu Warszawskiego. Wydawnictwa Uniwersytetu Warszawskiego, Warsaw.

Marek Świdziński. 1994. *Syntactic Dictionary of Polish Verbs*. Uniwersytet Warszawski / Universiteit van Amsterdam.

Piek Vossen, editor. 1998. *EuroWordNet: a multilingual database with lexical semantic network*. Kluwer Academic Publishers, Dordrecht, Holland.

Marcin Woliński. 2004. *Komputerowa weryfikacja gramatyki Świdzińskiego*. PhD thesis, Institute of Computer Science, Polish Academy of Sciences, Warsaw.

Charles T. Zahn. 1971. Graph-theoretical methods for detecting and describing gestalt clusters. *IEEE Transactions on Computers*, C-20(1).

Subject reactivation depends on the syntactic position of the argument: evidence from eye movements

Loes Koring
Utrecht University
Utrecht, The Netherlands
L.Koring@uu.nl

Pim Mak
Utrecht University
Utrecht, The Netherlands
W.M.Mak@uu.nl

Abstract

This paper investigates the on-line processing of different types of intransitive verbs. It shows that the argument of all verb types is reactivated during processing. However the point of reactivation differs per verb type. In agentive verbs, which take a syntactic subject, and assign the agent-role, we find an early reactivation. In unaccusative verbs, which take a syntactic object and assign the theme-role, we find a late reactivation. Mixed verbs (such as *sparkle*), which take a syntactic subject but assign a theme-role, pattern with agentive verbs. This indicates that the point of reactivation of the argument depends on the syntactic position.

1 Introduction

Intransitive verbs (verbs which take only one argument) can be classified on the basis of the thematic role the verbs assign to their argument. The argument can be interpreted as the *agent* (agentive verbs as in (1)) or as the *theme* of the event (unaccusative verbs as in (2)) (Marantz, 1984).

(1) The boy jumped *agentive*
(2) The boy fell *unaccusative*

Another difference is that the argument of unaccusative (and not agentive) verbs has properties that are normally associated with syntactic objects, although the argument appears in subject position (Burzio, 1986; Perlmutter, 1978). Hence, a connection exists between the subject and object position in unaccusative verbs.

Previous experimental research on the different types of intransitive verbs shows that the distinction between the two verb types is reflected in processing (Bever and Sanz, 1997; Friedmann et al., 2008; Poirier, 2009). Several experiments show that the argument of an unaccustive verb is reactivated late after verb offset, whereas this has not been found in agentive verbs (for an extensive overview of the studies see Koring & Mak, submitted).

2 Research Questions

A question that arises is what the source is of the processing difference between agentive and unaccusative verbs. Is the difference caused by the difference in thematic roles: agent vs. theme? Or is the difference the result of the distinction between syntactic subjects and objects? Previous studies cannot disentangle these factors as for all verbs the thematic and syntactic structure matched. That is, an agent would always be a syntactic subject and a theme would always be a syntactic object.

A distinct set of verbs exist for which the classification has so far remained unclear (e.g., *sparkle*). The verbs in this class differ from the agentive and the unaccusative verbs. These verbs show a mixed behavior with respect to the unaccusativity diagnostics (so-called unaccusative mismatches (L. Levin, 1985)). Reinhart (2000) argues that this set of verbs is in thematic structure similar to unaccusative verbs; they assign the theme-role. However, in syntactic structure they are like agentive verbs; the argument is a syntactic subject and lacks object properties (see

Koring and Mak (submitted) for the distinguishing properties). Hence, there is a mismatch in syntactic and thematic structure. Hence, we will call them mixed verbs in this paper. Given their mixed structure, investigating the processing of this class of verbs will enable us to differentiate between thematic and syntactic factors. That is, if they pattern in processing with agentive verbs, the processing difference is the result of a difference in the syntactic position of the argument. On the other hand, if they pattern in processing with unaccusative verbs, the difference is caused by a difference in thematic roles.

Another question is whether the argument of agentive verbs is reactivated during processing. Previous probe-studies were not able to detect this. We hypothesize that also in agentive verbs the argument will be reactivated as a result of integrating the argument and verb into one semantic object. In order to test this we designed a new method to measure activation of the argument throughout the complete sentence instead of at particular probe sites.

3 Method

The experiment used a version of the visual world paradigm (Tanenhaus et al., 1995). Huettig and Altmann (2005) showed that people spontaneously fixate on an object (target) that is semantically related to a spoken word. We hypothesize that people will not only fixate on a target upon *hearing* a related word, but also when this word is *reactivated* upon hearing the verb.

3.1 Participants

Forty Dutch native speakers participated in the experiment. Participants were seated in front of a screen showing visual displays with four objects, one in each quadrant. One of the objects was related to the argument. Participants were told to listen carefully to orally presented Dutch sentences, but had no further explicit task. The sentences all contained an intransitive verb in either one of three conditions (unaccusative, agentive, mixed). While they were listening, people's eye movements were measured by a Tobii eye-tracker sampling at 50 Hz.

3.2 Selection of verbs

Verbs were selected on the basis of several diagnostics, among which the type of auxiliary the verb selects (Hoekstra, 1984; Zaenen, 1993), the availability of the impersonal passive construction (Perlmutter, 1978) and whether or not adding *by itself* yielded an acceptable result (Levin and Rappaport, 1995) (see Koring and Mak (submitted) for a complete overview of the diagnostics). The Log transformed mean frequencies of the different verb classes did not differ significantly.

3.3 Pictures and sentences

For each verb an argument was selected that was not semantically related to the verb. Each argument was combined with a target object that *was* strongly related to the argument, but not to the verb. Semantic relations were pre-tested in a semantic relatedness judgment task (Perraudin and Mounoud, 2008) (see Koring and Mak (submitted) for the discussion).

In between the argument and the verb, material was added that was not related to the argument, target, or verb. In addition we added material after the verb in order to avoid end-of-sentence effects which resulted in sentences as in (3) translated from Dutch.

(3) Bert said that *the wood* (argument) of the fat gentleman with the bald head *fell* (V) hard after the heavy thunderstorm had begun with a flash.

This sentence was combined with a visual display showing a saw (target), a shell, a buggy, and a peacock (all distracters). The control condition consisted of the same visual display combined with the same sentence in which the argument was replaced by a word that was *not* related to the target (*clock* instead of *wood*). The difference in strength of relation between argument – target and control argument – target does not differ across verb types. The control condition served as a baseline (for a complete discussion of the method see Koring and Mak (submitted)).

4 Results

Figure 1 shows the distribution of fixation proportions over time. Looks to the target in the control condition (without a related argument) are subtracted from looks to the target in the test condition (with a related argument). The first and biggest increase in looks to the target is the result of presenting the argument itself (looks to the *saw* increase upon presenting *wood*). Later rises are the result of reactivating the argument.

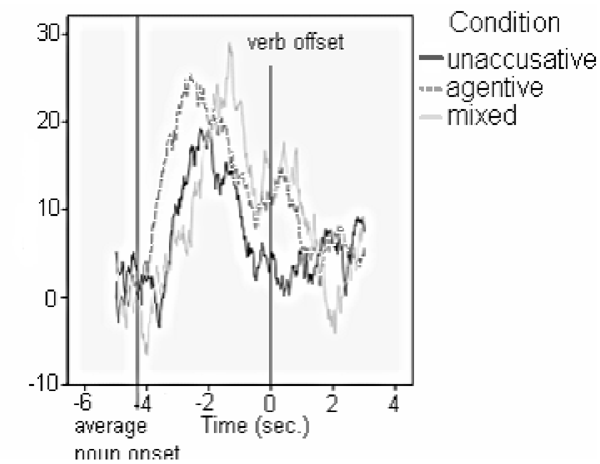


Figure 1. Mean percentage of looks to the target over time in the different conditions. The curves are synchronized to the acoustic offset of the verb.

For analysis we defined two regions on the basis of previous experiments: the verb frame (600 ms. before verb offset until 1000 ms. after verb offset) and the post-verb frame (200 ms. until 1700 ms. after verb offset). The verb frame takes verb offset plus 200 ms. as the mid-point. It does not take exactly verb offset as the mid-point as it takes 200 ms. to initiate and program an eye movement (Huettig and Altmann, 2005). The starting point is 600 ms. before verb offset as this is the average length of a verb. The post-verb frame takes 950 ms. after verb offset as the mid-point (previous literature found reactivation of the argument in unaccusative verbs 750 ms. after verb offset). The starting point is verb offset plus 200 ms.

Results of regression analyses in the verb frame show that the curves in all different conditions have a significant quadratic component. The quadratic component is positive for agentive and mixed verbs, but negative for unaccusative verbs. Growth curve analyses reveal that the quadratic component differs for the unaccusative verbs compared to the agentive and mixed verbs. (see Koring and Mak (submitted) for the details of the analysis).

Regression analyses on the individual curves in the post-verb frame indicate that the curves in the different conditions have a significant linear component which is positive for the unaccusative verbs, but negative for both the agentive and mixed verbs. Growth curve analyses show that unaccusative verbs differ in slope from the mixed and agentive verbs, whereas mixed and agentive verbs do not differ significantly in slope.

5 Discussion

The results of our eye-tracking experiment show that the argument of agentive verbs *is* reactivated during processing: looks to the target start rising from verb onset. That means that the argument of agentive verbs is reactivated much earlier than the argument of unaccusative verbs. Previous probe-tasks were not able to detect this reactivation as they did not probe at verb onset.

Mixed verbs show unaccusativity mismatches. They behave like agentive verbs in some diagnostics, but like unaccusative verbs in others. According to Reinhart (2000) they have the thematic structure of unaccusative verbs (they assign a theme-role), but they have the syntactic structure of agentive verbs. Our processing results show that they pattern in processing with agentive verbs. The argument of mixed verbs is reactivated as early as the argument of agentive verbs. The finding suggests that the point of reactivation of the argument depends on the syntactic position of the argument, and not on the thematic role that is assigned to the argument.

References

Bever and Sanz. 1997. Empty Categories Access Their Antecedents during Comprehension: Unaccusatives in Sapnish. *Linguistic Inquiry*, 28(1):69-91.

Burzio. 1986. *Italian syntax: a government-binding approach*. Reidel, Dordrecht.

Friedmann, Taranto, Shapiro and Swinney. 2008. The leaf fell (the leaf): the on-line processing of unaccusatives. *Linguistic Inquiry* 39:355-377.

Hoekstra. 1984. *Transitivity: Grammatical Relations in Government-Binding Theory*. PhD Dissertation

Huettig and Altmann. 2005. Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition* 96:23-32.

Levin, L. 1985. Operations on lexical forms: Unaccusative rules in Germanic languages. MIT dissertation.

Levin and Rappaport. 1995. *Unaccusativity*. MIT Press, Cambridge, Massachusetts.

Marantz. 1984. *On the Nature of Grammatical Relations*. The MIT Press, Cambridge, MA.

Mirman, Dixon, Magnuson. 2008. Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language* 59(4):475-494.

Perlmutter. 1978. Impersonal Passives and the Unaccusative Hypothesis. *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*: 157-189.

Perraudin & Mounoud. 2009. Contribution of the priming paradigm to the understanding of the conceptual developmental shift from 5 to 9 years of age. *Developmental Science, 12*, 956-977.

Poirier. 2009. Finding Meaning in Silence: The Comprehension of Ellipsis. PhD Dissertation, University of California.

Reinhart. 2000. The theta system: syntactic realization of verbal concepts. *OTS Working Papers in Linguistics*.

Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. Science, 268(5217), 1632–1634.

Zaenen. 1993. Unaccusativity in Dutch: Integrating Syntax and Lexical Semantics. In J. Pustejovsky (Ed.), *Semantics and the Lexicon*. (pp. 129-161). London: Kluwer.

The processing of flexible syntax–semantics mappings: A neurophysiological investigation of split-intransitivity in German

Dietmar Roehm
Department of Linguistics,
University of Salzburg
dietmar.roehm@sbg.ac.at

Antonella Sorace
Department of Linguistics,
University of Edinburgh
antonell@ling.ed.ac.uk

Ina Bornkessel- Schlesewsky
Department of Germanic
Linguistics, University of
Marburg
iboke@staff.uni- marburg.de

Abstract

Language comprehension requires a real time mapping between form and meaning. According to the Split Intransitivity Hierarchy, the interaction of telicity and agentivity creates gradient auxiliary preferences for intransitive verbs, ranging from categorical (rigid) to highly variable (flexible). This raises the question of how flexibility at the syntax–semantics interface affects language comprehension. In the present ERP study, we examined whether auxiliary selection with flexible verbs engenders quantitatively or qualitatively different neurophysiological responses to that for rigid verbs, and whether the compositional specification of telicity (via prefixation) leads to different processing signatures compared to the inherent lexical specification. Dispreferred auxiliary choices engendered a biphasic N400/late positivity pattern for rigid and prefixed verbs. The N400 for prefixed verbs had a later onset, indicating a quantitative difference between compositional and lexically specified telicity. Flexible verbs showed no effect of auxiliary choice in the average ERP. However, an additional linear mixed models analysis revealed an interaction between auxiliary selection and individual by-item / by-subject acceptability ratings. While item-based ratings proved a better predictor for N400 amplitude, subject-based ratings proved a better predictor for the late positivity amplitude. Thus, N400 amplitude is closely tied to inherent lexical properties, while the late positivity reflects individual participants' propensity to accomplish pragmatic enrichment.

1 Introduction

Successful language comprehension requires a real time mapping between form and meaning. To this end, the language processing system can exploit correspondences between syntax and semantics. However, in some cases, the syntax–semantics interface displays considerable flexibility, thus raising the question of how multiple possible mappings affect language comprehension. For example, in the domain of split intransitivity, there are verbs that can select either BE or HAVE depending on the characteristics of the predicate (Levin & Rappaport Hovav, 1995). Under these circumstances, an auxiliary thus does not provide the language processing system with an unambiguous indication of verb meaning / semantic role of the subject. The present ERP study examined split intransitivity in German as a testing ground for flexibility at the syntax–semantics interface. Sorace (2000) has proposed that intransitive verbs are organized in a Split Intransitivity Hierarchy defined by telicity ("telic change") at the core of unaccusativity and agentivity ("atelic non motional activity") at the core of unergativity. (see Fig. 1).

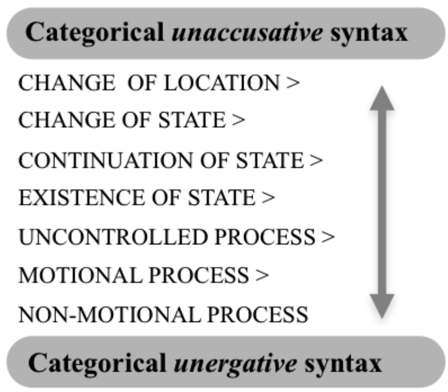


Figure 1. The Split Intransitivity Hierarchy

The interaction of these factors creates gradient auxiliary preferences, ranging from categorical (rigid) to highly variable (flexible). The closer to the core a verb is, the more determinate its syntactic status as either unaccusative or unergative, whereas the distance of a verb from the core correlates with sensitivity to contextual or compositional factors.

2 Materials and methods

Stimulus materials (see Table 1) included core unaccusative (change of location=CH-LOC), and core unergative verbs (controlled non-motional process=CON-PROC) as well as an intermediate verb class that is not inherently specified for telicity (change of state=CH-STATE-(UN)). By presenting each of these verb classes with both BE and HAVE, we examined whether auxiliary selection with non-core (flexible) verbs engenders quantitatively or qualitatively different neurophysiological responses to that for core (rigid) verbs and to what degree these neural processing signatures correlate with acceptability judgements. We additionally included prefixed change of state verbs (CH-STATE-(PRE)) to investigate whether the compositional specification of telicity (via prefixation) leads to different processing signatures compared to the inherent lexical specification. Hence, the experiment employed 4 verb classes (à 8 different verbs). Eighty sentences were constructed per verb class (with 10 different sentence contexts per verb). All sentences had the form NP / AUXILIARY / ADVERB / PAST PARTICIPLE, half included the auxiliary SEIN and half the auxiliary HABEN, thus resulting in 320 sentences plus 80 fillers of the same form (see Table 1 for examples). All sentences were presented visually in a segmented manner (NPs/verbs for 450 ms, all other segments for 400 ms with an ISI of 100 ms), followed by an acceptability judgment and a subsequent probe detection task. Thirty-two right-handed, monolingually raised native speakers of German (17 female) between 20 - 30 years of age (mean age: 23.97) participated in the experiment. The EEG was recorded from 64 Ag/AgCl-electrodes (500 Hz sampling rate, referenced to the left mastoid, and re-referenced to linked mastoids offline). The raw EEG data were filtered offline with a 0.3-20 Hz band pass. Automatic and manual rejections were carried out to exclude periods containing movement or technical artifacts. Average ERPs were calculated per condition per participant from -200 ms to 1000 ms relative to the onset of the critical sentence-final verb, before grand-averages were computed over all participants. For the statistical analysis,

Condition		Acceptability (%)	
Verb class	Example	HABEN	SEIN
CH-LOC	Die Bergsteigerin ist/*hat vorsichtig <u>aufgestiegen</u>. *The mountaineer is/has carefully climbed* 'The mountaineer climbed carefully'	1.9 (3.0)	93.5 (6.0)
CH-STATE-(UN)	Die Dose ist/hat sofort <u>gerostet</u>. *The tin is/has immediately rusted* 'The tin rusted immediately.'	61.0 (17.0)	65.0 (19.6)
CH-STATE-(PRE)	Das Auto ist/*hat langsam <u>verrostet</u>. *The car is/has slowly corroded* 'The car corroded slowly.'	6.0 (4.4)	88.6 (8.6)
CON-PROC	Die Lehrerin *ist/hat dauernd <u>geredet</u>. *The teacher is/has constantly talked* 'The teacher talked constantly.'	94.3 (5.7)	1.2 (2.3)

Table 1. Example sentences for the critical conditions in the present study as well as mean acceptability ratings (critical word underlined). Standard deviations (by participants) are given in parentheses. Abbreviations: CH-LOC = change of location; CH-STATE-(UN) = unprefixed change of state; CH-STATE-(PRE) = prefixed change of state; CON-PROC = controlled non-motional process.

repeated measures analyses of variance (ANOVAs) were computed for mean acceptability ratings using the condition factors VERB and AUX(iliary) and the random factors participants (F_1) and items (F_2), and for average ERPs using the condition factors VERB and AUX(iliary) and the topographical factor region of interest (ROI).

3 Results

Mean acceptability ratings for the judgment task are given in Table 1. The statistical analysis revealed a significant interaction of VERB x AUX ($F_1(3,93) = 246.2$, $p < .001$; $F_2(3,316) = 59.23$, $p < .001$), which was due to a significant preference for HABEN for CON-PROC verbs, a significant preference for SEIN for CH-LOC and CH-STATE-(PRE) verbs (all F_1s and F_2s < 0.001) and no significant preference for one auxiliary over the other for CH-STATE-(UN) verbs.

As is apparent from Figure 2, sentences with a dispreferred auxiliary led to a biphasic N400 - late positivity response for core verbs (CH-LOC; CON-PROC), and verbs for which auxiliary choice was determined via prefixation (CH-STATE-(PRE)). Statistical analyses were conducted in two time windows: 380-530 ms for the N400 and 750-900 for the late positivity.

In the N400 time window, the statistical analysis revealed an interaction ROI x VERB x AUX ($F(15,465) = 3.81$, $p < .002$). Resolving the interaction by ROI showed significant interactions of VERB x AUX in all ROIs (minimal $F(3,93) = 5.09$ in the left-anterior ROI; maximal $F(3,93) = 19.60$ in the left-central and left-posterior ROI)s. The simple effect of AUX was examined for each verb class in each of the ROIs showing an interaction VERB x AUX (for all effects reported, $p < 0.05$): CH-LOC verbs showed an effect of AUX in central and posterior regions; for CH-STATE-(PRE) verbs, the effect of AUX reached significance in all central and posterior regions as well as in the left- anterior ROI; CON-PROC verbs showed an effect of AUX in all regions. By contrast, CH-STATE-(UN) verbs did not show a significant effect of AUX in any region.

For the late positivity time window, the statistical analysis revealed an interaction ROI x VERB x AUX ($F(15,465) = 7.13$, $p < .001$). Analyses within each ROI showed significant interactions of VERB x AUX in all central and posterior regions (minimal $F(3,93) = 3.47$ at the left-central ROI, maximal $F(3,93) = 15.78$ at the right-posterior ROI). The simple effect of AUX was examined for each verb class in each of the ROIs showing an interaction VERB x AUX (for all effects reported, $p < 0.05$): CH-LOC and CH-STATE-(PRE) verbs showed an effect of AUX in all central and posterior regions; CON-PROC verbs engendered effects of AUX in all central and posterior regions except for the left-central ROI. For CH-STATE-(UN) verbs, no region showed a significant effect of AUX.

In sum, there was no evidence for a *qualitative* distinction in the processing of rigid (core verbs) and more flexible (prefixed verbs) syntax-semantics mappings. However, an additional analysis (examination of the effect of AUX in successive 30 ms time windows) showed that the N400 for CH-STATE-(PRE) verbs had a later onset than that for the other verb types, thereby indicating a *quantitative* difference between compositional and lexically specified telicity.

Importantly, auxiliary choice did not show any modulation of the grand average ERP response for lexically indeterminate verbs (CH-STATE-(UN)). Two scenarios appear to be possible: (a) the syntax–semantics mapping is underspecified such that both HAVE and BE fulfill the processing system's expectations for this particular verb class; or (b) the absence of an effect in the grand averages is a result of the averaging procedure.[1] For a more fine-grained analysis of the ERP data for this verb class we, thus used linear mixed models (Baayen et al. 2008) including the fixed factors AUX and ROI and the crossed random factors participants and items. By-participant acceptability (i.e. individual participants' acceptability ratings for each auxiliary type) and by-item acceptability ratings (i.e. acceptability ratings per auxiliary type for individual items) were included as covariates (in separate models). Mean ERP amplitudes for the N400 and late positivity time windows constituted the dependent variable. For both time windows, the inclusion of acceptability ratings led to an improvement of model fits over a base model without acceptability, as revealed by a likelihood ratio test (all *p's* < 0.001). In the earlier time window, the item-based acceptabilities led to a slightly better fitting

[1] I.e. a product of averaging over gradient responses which differ on a trial-to-trial basis.

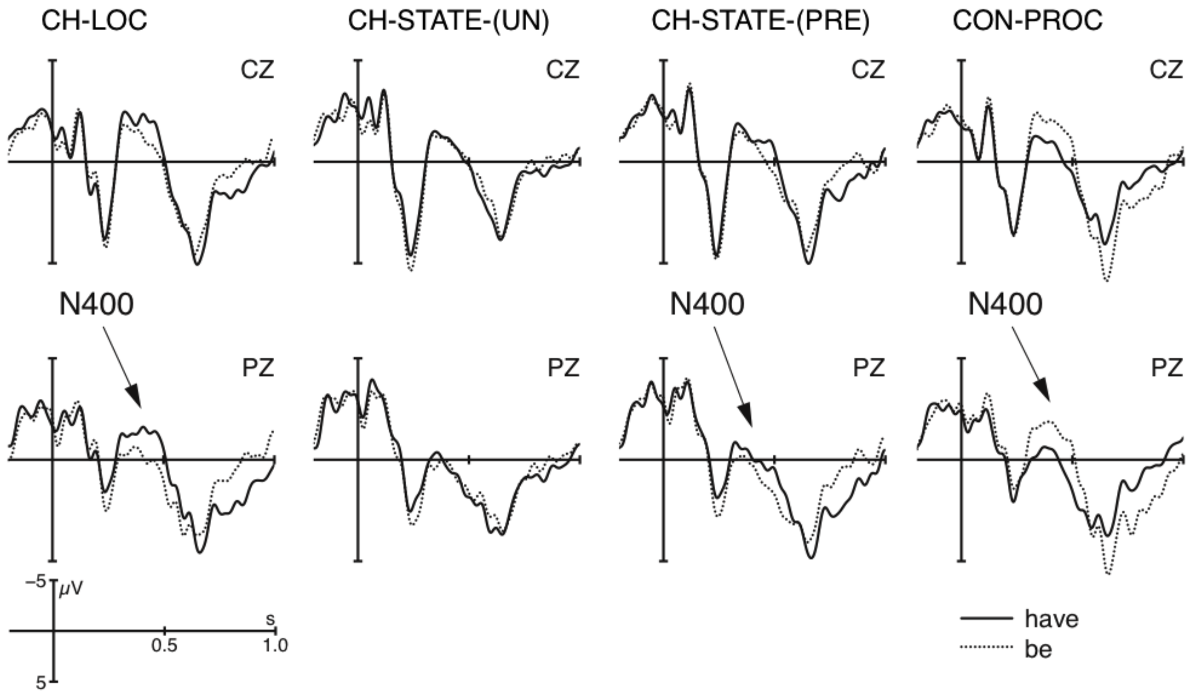


Figure 2. Grand average ERPs at the position of the past participle (onset at the vertical bar) for: (CH-LOC) change of locations verbs, (CH-STATE-(UN)) change of state verbs, (CH-STATE-(PRE)) prefixed change of state verbs, and (CON-PROC) verbs of controlled non-motional process. Negativity is plotted upwards.

model than the subject-based acceptabilities (χ^2 = 2.27, p < 0.001). In the later time window, by contrast, subject-based acceptabilities led to a substantial improvement of the model fit in comparison to item-based acceptabilities (χ^2 = 76.49, p < 0.001).

Discussion

For sentences including a dispreferred auxiliary, we observed a biphasic N400 - late positivity pattern. This pattern was engendered by core unergative verbs (verbs of controlled non-motional process) and core unaccusative verbs (verbs of change of location), i.e. verbs with an inherent lexical specification of the key semantic feature of telicity, and for lexically indeterminate verbs for which the choice of auxiliary was determined by the addition of a telicity-inducing prefix (prefixed change of state verbs). Importantly, there was no evidence for a *qualitative* distinction in the processing of flexible vs. rigid syntax-semantic mappings. The data did, however, provide evidence for a *quantitative* distinction between the processing of these different verb types in the form of an N400 latency shift. Compositional as opposed to lexically specified telicity led to a longer N400 onset latency, which could be taken as an indication of the additional computational effort required by the composition process. These results are consistent with the SIH and allow a more fine-grained understanding of the neurophysiological bases for the gradient behavioural differences among verbs belonging to the same continuum.

All in all, we interpret the N400 effects observed here as correlates of the mismatch between the processing system's expectation for a particular lexical aspect (as induced by the auxiliary) and the properties of the verb that is actually encountered. When the mismatch is compositional (i.e. induced via the combination of a verb and a telicity-inducing prefix), its detection is computationally more complex than in the case of a direct lexical encoding of telicity, thereby leading to a longer N400 latency. By contrast, we assume that the late positivity reflects a categorisation process by means of which the sentences with a dispreferred auxiliary are classified as ill-formed (Bornkessel & Schlesewsky, 2006; Kretzschmar, 2010).

Perhaps the most interesting finding of the present study was that sentences with lexically indeterminate verbs (unprefixed change of state verbs, CH-STATE-(UN)) did not show any differences between BE and HAVE in grand average ERPs, and also led to gradient acceptability ratings. An additional analysis of the ERP data for this verb class using linear mixed models revealed an interaction between auxiliary selection and acceptability ratings. Whereas, in the N400 time window, individual item-based acceptability per condition provided a slightly better-fitting model, individual subject-based acceptability per condition proved to be a better predictor for the late positivity time window. This finding suggests that the absence of an effect was likely due to the averaging procedure rather than resulting from the processing system's general indifference to auxiliary selection with this particular class of verbs.

Crucially, these results support the interpretation advanced above that the N400 effects in the present study index a mismatch between an auxiliary-induced expectation and the actual aspectual properties of the verb. In the N400 time window, the interaction between auxiliary type and item-based acceptability was, in part, due to more negative-going ERPs for the HAVE sentences with decreasing item-based acceptabilities. Considering that a change of state implies telicity at some level, these verbs are clearly close to the "BE end" of the split intransitivity hierarchy (see Figure 1). Thus, it appears plausible that these verbs should show an (albeit weak) mismatch when encountered with HAVE.[2] Strikingly, the late positivity time window showed an inverse pattern with regard to the relationship between ERP effects and acceptability ratings. ERPs tended to be more positive for those individual subjects who showed a higher acceptability for the CH-STATE-(UN) verbs with HAVE.[3] We suggest that this positivity response reflects processes of pragmatic enrichment (see Burkhardt, 2006; Burkhardt & Roehm, 2007; Schumacher, to appear). When used with HAVE, verbs of this type call for an enrichment process in order for the change of state to be interpreted as a process rather than as a telic change. If this process of enrichment is successful, the acceptability of CH-STATE-(UN) verbs with HAVE is higher. The correlation with individual participants' acceptability ratings indicates that some participants may have a propensity for aspectual enrichment processes, thus leading them to consider these verbs more acceptable with HAVE in comparison to participants with a lower tendency to enrich.

Conclusion

The present study demonstrated that indeterminacy at the syntax-semantics interface is, in part, processed in a qualitatively similar manner to consistent (rigid) form-to-meaning mappings, while also providing evidence for some degrees of quantitative and qualitative variation. N400 amplitudes were shown to correlate consistently with item-based aspectual preferences. On a trial-by-trial basis, this even appears to be the case for verbs that are lexically not specified for telicity. In addition, N400 latency was modulated by the compositional complexity of the aspectual information: the N400 onset was delayed for verbs in which the preference for BE was brought about compositionally rather than via lexical specification for telicity. Finally, qualitatively different effects for indeterminacy vs. consistency were observed in the late positivity time window. Whereas violations of consistent (rigid) form-to-meaning mappings engendered a late positivity for unacceptable auxiliary selections, the amplitude of the late positivity for indeterminate verbs correlated positively with individual participants' acceptability ratings. We suggest that this distinction reflects a well-formedness categorisation for consistent mappings on the one hand and processes of pragmatic enrichment for flexible (indeterminate) mappings on the other. The latter appear to vary across individual speakers. These findings, on the whole, allow us to gain a better understanding of the gradient effects obtained in previous studies on the SIH and, more generally, of the cognitive bases of gradient variation in language.

References

Baayen, R. Harald, Douglas J. Davidson, and D.M. Bates. 2008. Mixed-effects modeling with crossed

2 Moreover, the current findings suggest that some of the change of state verbs employed in the present study are more strongly specified for telicity than others.

3 This indicates that, in this case, the positivity response is not simply the reflection of a mismatch between an expected and actually encountered verb type (in which case amplitudes should have been more positive for *lower* acceptabilities).

random effects for subjects and items. *Journal of Memory and Language*, 59:390-412.

Bornkessel, Ina and Matthias Schlesewsky. 2006. The Extended Argument Dependency Model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, 113:787-821.

Burkhardt, Petra. 2006. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, 98:159-168.

Burkhardt, Petra and Dietmar Roehm. 2007. Differential effects of saliency: An event-related brain potential study. *Neuroscience Letters*, 413:115-120.

Kretzschmar, Franziska. 2010. *The electrophysiological reality of parafoveal processing: On the validity of language-related ERPs in natural reading*. Unpublished doctoral dissertation, University of Marburg.

Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax-semantics interface*. MIT Press. Cambridge, MA.

Schumacher, Petra B. to appear. The hepatitis called.: Electrophysiological evidence for enriched composition. In J. Meibauer & M. Steinbach (Eds.), *Experimental pragmatics/semantics*. John Benjamins. Amsterdam.

Sorace, Antonella. 2000. Gradients in auxiliary selection with intransitive verbs. *Language*, 76: 859-890.

Argument Structure Representation: Evidence from fMRI

Einat Shetreet
Tel Aviv University
einat.shetreet@childrens.harvard.edu

Dafna Palti
Tel Aviv University
paltida@post.tau.ac.il

Naama Friedmann
Tel Aviv University
naamafr@post.tau.ac.il

Uri Hadar
Tel Aviv University
Urih1@post.tau.ac.il

Abstract

This paper reports on a series of fMRI experiments testing the processing and representation of various aspects of argument structure of verbs, including the number of complements, the syntactic type of complements, the number of complementation frames, and the optionality of complements. We found that different brain areas are involved in the processing of these different properties.

1 Introduction

Linguistic theory regarding the representation of verbs holds that the lexical entry of verbs includes information about their argument structure (AS). AS specifies the syntactic and semantic environments in which the verb can occur (e.g., Chomsky 1965; Grimshaw, 1979; van Valin, 2001). Several psycholinguistic and neurolinguistic studies have demonstrated that such lexical information affects on-line processing of sentences across different tasks and methods (e.g., Shapiro et al., 1987, 1993; Tanenhaus et al., 1989; Trueswell et al., 1993). The current study used fMRI to examine various lexical-syntactic properties of verbs (number of complements, number of complementation frames, syntactic complexity of complementation frames, and the optionality of complements), to describe the way they influence sentence comprehension and identify their cortical localization, as well as provide evidence appertaining to linguistic controversies regarding verb representation.

2 General Method

The experimental procedure was identical in all the experiments. Hebrew verbs were selected based on the tested property, following judgment procedure by linguists and psycholinguists. In the experiments, each verb was embedded in a few (2-4) sentences. The sentences in each experiment were controlled for the number of phrases, phrase structure, definiteness, duration and frequency. A block design paradigm was used. Each block included 4 sentences and each condition repeated 7 or 8 times. Twelve to nineteen participants were asked, while in the MRI, to listen to the sentences and to decide whether the event described in the sentence is more likely to happen at home or not (for example, for a sentence like "Dan slept in the yellow tent", participants will press a "no" button). This semantic task ensured that participants attended to the sentences and processed them fully. Data was analyzing using SPM2.

3 Experiments

3.1 Experiment 1: Number of complementation frames

It is assumed, based on behavioral studies with reaction times, that all complementation frames of a verb are activated at some stage of processing, regardless of the complement that is used in the sentences. Shapiro et al. (1987, 1989, 1991, 1993) found that this exhaustive access to all complementation frames affects sentence and verb processing such that verbs that have more complementation frames were accessed more slowly than verbs with a single frame. In this experiment (Shetreet et al., 2007), we tested the effect of the processing load of the number of frames on brain activations. This was done by creating a three-point scale, using verbs with one (e.g., *punish*), two (e.g., *discover*) or three complementation frames (e.g., *demand*). The verbs appeared in sentences with either a Noun Phrase (NP) or Prepositional Phrase (PP) complements (e.g., *Dan discovered the mouse in the morning*). Using a parametric design, we looked for brain region in which increasing the number of frames increased the level of activation. We found such graded activation in the left superior temporal gyrus (Wernicke's area), which was assumed to be involved in the processing of the number of complementation options. This assumption was based on the performance of patients with Wernicke's aphasia who did not show sensitivity to the number of complementation options (Shapiro et al., 1993). We also found two clusters of activation in the left inferior frontal gyrus: in BA 47 and in BA 9, which might be involved in selection of competing alternatives (e.g., Thompson-Schill et al., 1997).

In this experiment, we also tried to differentiate between subcategorization frames and thematic frames using verbs that have two subcategorization frames, but one thematic frame (e.g., *nibble* that can appear with either NP or PP complements, both having the thematic role of *theme*). These verbs showed pattern of activation similar to that of verbs with two complementation frames, suggesting that subcategorization cannot be discarded in favor of explanation in purely thematic roles terms, and that it is important for verb processing.

3.2 Experiment 2: Number of complements

Findings from behavioral studies showed that, unlike the number of complementation frames, the number of complements does not affect online processing (Shapiro et al., 1987). Experiment 2 (Shetreet et al., 2007) tested whether the neuronal picture is similar. This was done by comparing verbs that take zero (e.g., *sneeze*), one (e.g., *punish*), or two complements (e.g., *give*), creating a three-point scale of the number of complements. In each sentence, the verb was followed by two constituents, either complements or adjuncts (e.g., two adjuncts: *John sneezed yesterday in bed*; one complement and one adjunct: *Laura broke the glass at midnight*; two complements: *Helen gave the present to Billy*). Here too we used a parametric design to detect graded activations. The pattern of activation in this study crucially differed from that of Experiment 1. The two clusters that showed graded activation according to this gradient have not been traditionally considered to be involved in language processing: one activation cluster was found in the anterior cingulate and one in the medial-precuneus. The activation in the cingulate may stem from its involvement in working memory. Working memory load was expected due to the need to retain a greater amount of information as the number of complements increased. The precuneus has recently been found active in several language studies, including our own (Bachrach, 2008; den Ouden et al., 2009; Shetreet et al., 2009). Bachrach suggested that the precuneus plays a central role in the representation of linguistic syntactic structure. Thus, it seems that the number of complements affected sentence processing, however not in the expected areas (i.e., classic language areas). This may explain the inconsistency with the behavioral results that used interference method. It could be that the resolving the secondary task in the behavioral experiment loaded on linguistic areas, but the processing of the number of complements was done in a different area, using different resources.

3.3 Syntactic complexity of complementation frames

We thus found in Experiment 2 that the number of complements did not load on classic language areas. In this experiment (Shetreet et al., 2009), we tested whether the syntactic complexity of the complement, rather than the number of complements, does show an effect in language regions. For that aim, we used verbs that select sentential complements (CP) or prepositional phrases (PP, e.g., *complain*) and compared them with verbs that select noun phrases (NP) or PPs (e.g., *nibble*). That is, we compared verbs that can take sentential complement and those that cannot. CP complements are syntactically more complex than NP complements, because they include more syntactic layers (lexical Verb Phrase layer, inflectional (IP) layer, and complementizer (CP) layer; Rizzi, 1986). To examine whether the mere inclusion of a CP complement in the lexical entry of the verb, even when the CP was not realized in the sentence, affected the access to the verb, we compared the two verb types when they appeared in the sentence with a PP complement (e.g., *John complained about the cold ice-cream* and *John nibbled at the tasty cake*), thus comparing sentences with identical syntactic structure, but verbs that differ in their lexical information with respect to the syntactic types of their complements. The comparison between verbs that can take a sentential complement and those that cannot yielded activations in bilateral anterior middle temporal gyrus and the precuneus. This indicated that syntactic information regarding the syntactic types of complements in the lexical entry of the verb is reflected in brain activity even when not realized in the sentence.

3.4 Experiment 4: The representation of optional-complement verbs

Finally, we examined the cortical representation of verbs with optional complements, which can appear with and without their complements (e.g., *eat*) in an attempt to provide neurally-based constrains for the linguistic theory (Shetreet et al., 2010). We examined three linguistic approaches for the representation of optional verbs- one that argues that these verbs have two subcategorization frames (one with the complement and one without it; e.g., Engelberg, 2002; van Valin & LaPolla, 1997); an approach that argues that there is one subcategorization frame, with the complement, and its omission is made possible through a syntactic operation (null element; e.g., Cummins & Roberge, 2004); and an approach that argues that there is one frame and that the omission of the complement is made possible through a lexical saturation of the complement (e.g., Bresnan, 1982; Dowty, 1978, 1989). Each of these theories bears different hypotheses with regard to the number of frames and number of complements that sentences with optional complements have. We relied on these distinctions between the theories in our attempt to discriminate between them. First, we assessed the number of frames of verbs with optional complements- to distinguish between the two frames theory and the other two theories, which assume that these verbs have a single complementation frame. To do so, we contrasted verbs with a known number of frames (1 or 2) and compared the identified regions to regions identified in the comparison of verbs with optional complements to verbs with one frame and to verbs with two frames. We found that the comparison between verbs with optional complements and two-frame verbs revealed activations similar to the activation found in the comparison between one- and two-frame verbs. Among the identified regions was the left STG, also identified in Experiment 1 that tested the number of frames. These results suggest that verbs with optional complements have only one frame. In the next stage, we assessed the number of complements in sentences that include verbs with omitted complements. According to the syntactic operation theory, a null element is placed in the position of the omitted complement. Thus, this theory predicts that sentences with omitted complement will be syntactically similar to sentences with a complement. By contrast, according to the lexical saturation account, sentences with omitted complements are similar to sentences containing verbs with no complements, because both are inserted from the lexicon into the sentence without any complement. It is important to note that phonetically null elements like the one assumed by the syntactic theory can be detected by neuroimaging techniques, such as ERP (Featherson et al., 2000; Fiebach et al., 2001; Kluender and Kutas, 1993) or fMRI (Shetreet et al., 2009a). The baseline for this assessment was the comparison between verbs with one complement and verbs with no complements. We contrasted sentences with omitted complements to both no- and one-complement verbs and compared the results of each comparison to the baseline comparison. We found that the sentences with omitted complements were more similar to sentences containing verbs with no complements. One of the areas identified in both of these comparisons was the precuneus that was identified in Experiment 2 in the assessment of activations related to the number of complements. This supports the lexical saturation account for omission of complements. Thus, we concluded that verbs with optional complements have only one subcategorization frame, with the complement, and that a lexical operation enables the complement omission. In addition, by comparing sentences containing verbs with omitted complement to the other conditions, we identified the fusiform gyrus and possibly the temporal-parietal-occipital junction as having a role in lexical saturation and the execution of the omission of optional complements.

4 Conclusion

These experiments revealed on-line effects of some of the critical aspects of verb processing during sentence comprehension, including the number of subcategorization frames and the syntactic properties of the complements. Furthermore, we showed that the processing of lexical-syntactic information regarding the verb's arguments is distributed in a network of regions, which extends the classic language sites. Additionally, the results of this study clearly indicate that the linguistic ideas are reflected in brain activations and provide arguments to decide between linguistics theories.

To conclude, one of the most important and unique aspects of this study is in the interface it suggests between linguistics and neuroscience. The theoretical linguistic framework played a critical role in the interpretation of the brain activations and the brain activations provided neurally-based arguments to linguistic debates. Thus, linguistics and neuroscience can inform and enrich each other, as well as constrain one another and, on the whole, derive scientific gains from the two-way consideration of possible mechanisms.

References

Bachrach, A. (2008). *Imaging neural correlates of syntactic complexity in a naturalistic context*. Doctoral dissertation, MIT.

Bresnan, J. (1982). Polyadicity. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 149-172). Cambridge, MA: MIT Press.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Cummins, S., & Roberge, Y. (2004). Null objects in French and English. In J. Auger, C. Clements, & B. Vance (Eds.), *Contemporary approaches to Romance linguistics* (pp. 121-138). Amsterdam: John Benjamins.

Den Ouden, D. B., Fix, S. C., Parrish, T., & Thompson, C. K. (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, *22*, 196-215.

Dowty, D. (1978). Governed transformations as lexical rules in a montague grammar. *Linguistic Inquiry*, *9*, 393-426.

Dowty, D. (1989). On the semantic content of the notion of ''thematic role''. In G. Chierchia, B. Partee, & R. Turner (Eds.), *Properties, types, and meanings*, vol. 2 (pp. 69-129). Dordrecht: Kluwer.

Engelberg, S. (2002). Intransitive accomplishments and the lexicon: The role of implicit arguments, definiteness, and reflexivity in aspectual composition. J*ournal of Semantics*, *19*, 369-416.

Featherston, S., Gross, M., Münte, T. F., & Clahsen, H. (2000). Brain potentials in the processing of complex sentences: An ERP study of control and raising constructions. *Journal of Psycholinguistic Research*, *29*, 141-154.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research*, *30*, 321-338.

Grimshaw, J. (1979). Complement selection and the lexicon. *Linguistic Inquiry*, *10*, 279-326.

Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*, 196-214.

Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), (pp. 281-337) *Elements of grammar: A Handbook of generative syntax*. Dordrecht: Kluwer.

Shapiro, L. P., Brookins, B., Gordon, B., & Nagel, N. (1991). Verb effects during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 983-996.

Shapiro, L. P., Gordon, B., Hack, N., & Killackey, J. (1993). Verb argument processing in complex sentences in Broca and Wernicke's aphasia. *Brain and Language*, *45*, 423-447.

Shapiro, L. P., Zurif, E. B., & Grimshaw, J. (1987). Sentence processing and the mental representation of verbs. *Cognition*, *27*, 219-246.

Shapiro, L. P., Zurif, E., & Grimshaw, J. (1989). Verb processing during sentence comprehension: Contextual impenetrability. *Journal of Psycholinguistic Research*, *18*, 223-243.

Shetreet, E., Friedmann N., & Hadar U. (2009). An fMRI study of syntactic layers: Sentential and lexical aspects of embedding. *NeuroImage*, *48*, 707-716.

Shetreet, E., Friedmann, N., & Hadar, U. (2009a). The cortical representation of unaccusative verbs. *Language and Brain*, *8*, 21-29. (in Hebrew)

Shetreet, E., Friedmann N., & Hadar U. (2010). The cortical representation of verbs with optional complements: The theoretical contribution of fMRI. *Human Brain Mapping*, *31*, 770-785.

Shetreet, E., Palti, D., Friedmann, N., & Hadar, U. (2007). Cortical representation of verb processing in sentence comprehension: Number of complements, subcategorization and thematic frames. *Cerebral Cortex*, *17*, 1958-1969.

Tanenhaus, M. K., Boland, J., Garnsey, S. M., & Carlson, G. N. (1989). Lexical structure in parsing long-distance dependencies. *Journal of Psycholinguistic Research*, *18*, 37-50.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, J. K., & Farah, M. J. (1997). Role of the left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the USA*, 14792-14797.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. Journal of Experimental Psychology: *Learning, Memory and Cognition*, *19*, 528-553.

van Valin, R. D. (2001). *An introduction to syntax*. Cambridge, UK: Cambridge University Press.

Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning and function*. Cambridge, UK: Cambridge University Press.

Comparing linguistic classifications with sensorimotor data of English and Greek verbs of motion

Marietta Sionti[1], Leonardo Claudino[2], Carolyn P. Rose[3], Stella Markantonatou[4], Yiannis Aloimonos[2]

[1]University of Athens, [2]University of Maryland, [3]Carnegie Mellon University,
[4]Institute for Speech and Language/ 'R.C.Athena'

Abstract: We combine linguistic knowledge from corpus data with sensorimotor data obtained experimentally in an effort to better specify the minimum conceptual representation of a motion event that distinguishes it from all other events. Sensorimotor data are collected by measuring the performance of speakers of Modern Greek and American English. We focus on the clustering of motor actions and its correspondence to previous linguistic classifications of both languages.

1 Introduction

We combine corpus driven linguistic knowledge with experimentally obtained sensorimotor data in an effort to better specify the minimum conceptual representation of a motion event that distinguishes it from all other events. We use American English and Modern Greek data as a case study, in order to enable crosslinguistic analysis. We hope that our work will contribute to a better understanding of both language and perception. Sensorimotor data, here collected by measuring the motor behavior of speakers of American English and Modern Greek, allow for linking embodied experience and language (image schemas are learnt as a sequence of interrelated sensorimotor patterns (Lakoff & Johnson, 1999)). We expect that by delineating conceptual representations of motion events in this way, we will be able to (i) better understand semantic classifications of verbal predicates (ii) perceptually ground abstract notions, eg *transitivity*, that have traditionally been used in linguistics to study and classify verbal semantic and syntactic properties, and by combining them with corpus driven data, (iii) offer a quantitative answer to long standing syntactic questions such as the distinction between "argument" vs. "adjunct", whose binary nature has been questioned (Galen, Grenager&Manning, 2004).

This paper focuses on the collection of sensorimotor data, the clustering of

motor actions and its correspondence to previous linguistic classifications (Levin, 1993; FrameNet; Antonopoulou, 1987). The detailed sensor data are analyzed to identify latent factors that represent stable patterns across the many dimensions of low level data. These factors appear as discrete sets (*synergies*) of joint angles and orientations associated with each action and are correlated with linguistic descriptions.

We draw on the extensive prior work related separately to Cognitive Science (Jackendoff, 1990; Feldman, 2006), Mirror Neurons and their impact on language (Fadiga *et al*, 2006; Arbib, 2008; Kemmerer, 2006) and Computer Science (Santello, 1998; Troje, 2002).

2 Linguistic Classifications

2.1 Modern Greek:

There exist two studies on the classification of Modern Greek Motion Verbs (MGMV) (Antonopoulou, 1987; Mpasea, 1996). MGMV exhibit an overall semantic structure found with motion verbs of several Indo-European languages and, at the same time, present some certain aspectual idiosyncrasies. Antonopoulou (1987) adopted prototype theory as the most suitable method for the investigation for MGMV; prototype theory is by default closer to the cognitive approach adopted here.

Antonopoulou's taxonomic sets were defined with the use of two groups of criteria:

- Criteria of the first group: transitivity, causativity, agentivity, intentionality and aspect; none of them can be measured with sensorimotor methods at the moment.
- Criteria of the second group: change-of-location, directionality, path, dependent motion, change-of-orientation, manner, medium and instrumentality.

2.2 American English

Though the difference in perspective of Levin's (1993) English Verb Classes and Framenet's categorization is well attested (Baker & Ruppenhofer), both these classifications are important for this work. Levin's classes are based on semantic grouping and valence alternations. Very much like Antonopoulou, Levin offers a rich anthology of verbs enriched with syntactic information that is crucial for our long standing goal, namely the distinction *argument* vs *adjunct*. On the other hand, Framenet's grouping of words according to conceptual structures can be easily matched to Antonopoulou's second group of criteria and, finally, to sensorimotor data.

3 Sensorimotor experiment

3.1 Verb collection criteria

The verbs used for the experiments fulfilled the following criteria: (i) one human participant per action, including both intransitive and transitive verbs, and (ii) keeping to more literal meanings mainly due to lab limitations (Table1).

3.2 Experimental procedure:

Method: The detailed sensor data were analyzed to identify latent factors that represent stable patterns across the many dimensions of low level data. These factors appear as discrete sets (*synergies*) of joint angles and orientations associated with each action.

Equipment: A full body Moven system contains 16 inertial motion trackers. Each sensor module comprises 3D gyroscopes, 3D accelerometers and 3D magnetometers. Using advanced sensor fusion algorithms (Moven Fusion Engine) the inertial motion trackers give absolute orientation values which are used to transform the 3D linear accelerations to global coordinates which in turn give the translation of the body segments. The advanced articulated body model (23 segments and 22 joints biomechanical model) implements joint constraints to eliminate any integration drift or foot-sliding.

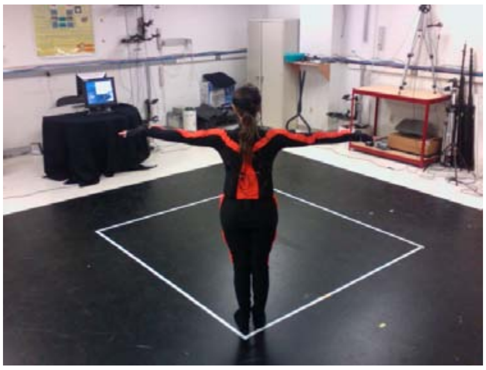


fig.1 A subject while performing an action

Subjects: The age range of the subjects was 25-30 years old and their sexual distribution 5 men and 3 women. Being native speakers, they were encouraged to implement each meaning according to their intuition.

Phases: The sensorimotor experiment was divided in two phases; (i) capturing of the action performance- this part yielded the main dataset for Greek and English (8 subjects each) and (ii) correspondence between the languages.

Action performance: This was a step-wise procedure. The verb or the sentence was uttered by the experimenter. When the verb was performed with the body of the subject only, action was limited to a floor area restricted by a quadrangle. In order to normalize the distance, subjects were encouraged to start acting at a specific corner of the quadrangle (fig.1).

Although several actions could be implemented by using only the body and the prerequisite was to involve as few objects as possible, the subjects asked for items that could be found in the lab:

- a step (verbs 8, 10, 15)
- a ramp (9, 11)
- one or several balls (5, 14)
- table, book, cylinder, chair (22-25)
- chair (20, 21)

In order to standardize the procedure, the same objects were used throughout the experiments (whenever an object was required).

Correspondence of the verbs in the two languages: The Greek participants performed the Greek verbs of motion and one of them was videotaped. Ten English speakers were shown the video segments and were asked for the corresponding English verb that would best describe each action. In controversial answers, we substituted the open question with multiple choices, complementing them with similar entries from WordNet. In the cases that the problem persisted, we asked the English participants of the sensorimotor experiment to perform both choices. Therefore, two tendencies were observed: (a) 1:1 correspondence between the verbs of the two languages and overlapping in the meanings and (b) participants not feeling confident both about the meaning of the verb and how to perform the corresponding action.

4 Analysis of the data and results

4.1 Analysis

In order to describe the motor representation of each verb, we extract its average action. These average actions are normalized in length and further "stacked", forming the base motor data matrix of our work. That matrix is called the Principal Component Analysis (PCA) and is processed as in (Santello, 1998), allowing a two-dimensional visualization of the action scatter.

The two main visual groups of actions are (i) the rectangular that includes walking like actions (leg-related), and (ii) the blue ellipsis that includes manipulation of an object (arm-related) (fig.4, 5). They are projected on both PC1 (distinct use of hips) and PC2 (emphases on knees and shoulders) (fig.6, 7). Therefore, all these actions have approximately the same profile in terms of joints-angles.

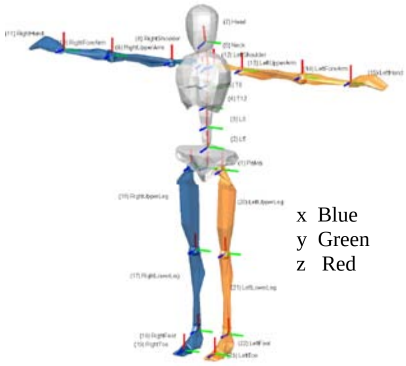


fig.2 X Moven Suit's axes

Principal Component 1 (PC1, fig. 6, 7) emphasizes the hips x (displacement on the sagittal plane, fig.2, 3). Principal Component 2 (PC2) highlights the combination 'right and left knee and shoulders' towards all directions (x, y, z). In fig. 6, 7 each box depicts the amount of energy over time, e.g. a joint is white only in the first half of the box, if it is highlighted only for the first half of the action.

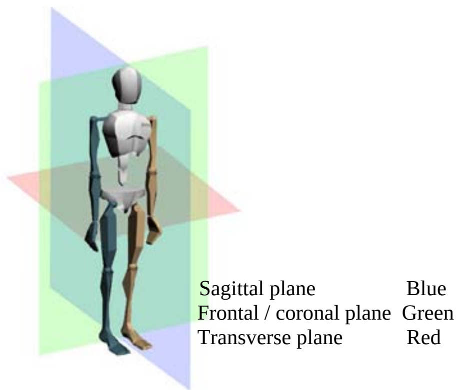


fig.3 X Moven Suit's body planes

The visualization presented here (fig.4, 5) supports the results of brain imaging studies; the schematic of the distributed semantic representation in the

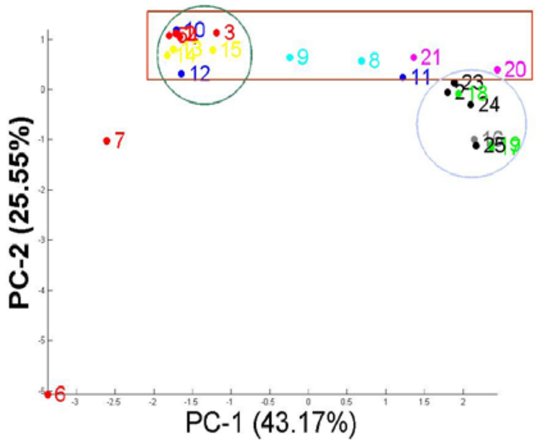


fig.4 Average Actions -Greek Verbs
(for numbers please seeTable 1)

brain of action verb processing is based on the body parts performing them (Wermter *et al.*, 2005), e.g. arm-related and leg-related.

Comparing this visualization with the linguistic classification (Antonopoulou, 1987; Levin; 1993) we see that Antonopoulou's classes are more fine-grained than the two big categories in fig.3, where all actions consisting each of the two linguistic groups appear in the same scheme –rectangular or ellipsis. Especially, the members of the walk-group, *pido*-group and *kateveno*-group are coiled together (green ellipsis). The same occurs for *girizo2*-group, *katevazo* and *anevazo* group (blue ellipsis). Levin's classes are similar to the groups at fig.5.

The actions 20, 21 (*girizo1*) and 22, 23, 24, 25 (*girizo2*) share the same morphological representation *girizo*. These actions present themselves close to each other on the PC1 projection but are separated on the PC2 projection fig.4). The first group -the green ellipsis- is mainly about leg motion. The second group -blue ellipsis- is about object manipulation that necessarily involves arm displacement as is clearly indicated on PC2. In English all *roll* verbs belong to the same class.

Though the plots were based on similar or even identical actions, certain divergences occurred. For instance, the verb *draskelizo* has traditionally been translated as *stride*, since both share longer steps. The Greek subjects always needed a small obstacle, such as a ball or a hole on the ground, to perform longer steps. On the contrary, the English subjects clearly distinguished between *stride* and *step over* (although WordNet assigns this meaning to *stride* as well). Similarly, we would expect *vimatizo* to be closer to *pace* rather than *march*, but it should be noted that the majority of the English participants were unsure for the exact representation of *pace*. When we compare the signals of *vimatizo* and *march*, we see significant similarity in the manner of stepping, while English subjects also emphasize the movement of the arms.

Of particular interest in the Greek plot is the distance of *treho-run* and *mpousoulo-crawl* from the rest leg-related actions (especially for English, *march* and *crawl* show the same behavior). Although we would expect *run* to resemble *treho*, differences occurred due to two reasons: (i) the English subjects tented to use their arms less than the Greek ones, and (ii) each group performed *march* in a different way; this time, the English subjects used their arms more than the Greek ones (as opposed to the performance of *run*). The blue and green ellipsis are projected both on the same PC1 (highlighted hips) and 2 (highlighted knees and shoulders).

Furthermore, the reason why *march*, *treho* and *crawl-mpousoulo* are projected on different PC2 narrows down to (i) the height of the knees (upward-downward and forward-backward respectively) and, (ii) the frequent movement of shoulders in all directions. But still, these verbs are projected on the same PC1. PC1 is about the forward and backward displacement of the hips and emphasizes on the leg related actions, namely the walking like actions, which, in turn, is considered to be the actions' common linguistic characteristic.

The above findings are still consistent with the aforementioned linguistic analysis. The Greek *treho* can form its own class in terms of velocity according to (Antonopoulou, 1983). At the same time, in fig.4, action 7 differs from all the other actions performed because of velocity. This fact is incorporated in the depiction of time in fig. 6. Probably Levin (1993) gives us a hint that these verbs need special treatment, since she enlists them under both the *meander* verbs and the *run* class (it must be kept in mind that

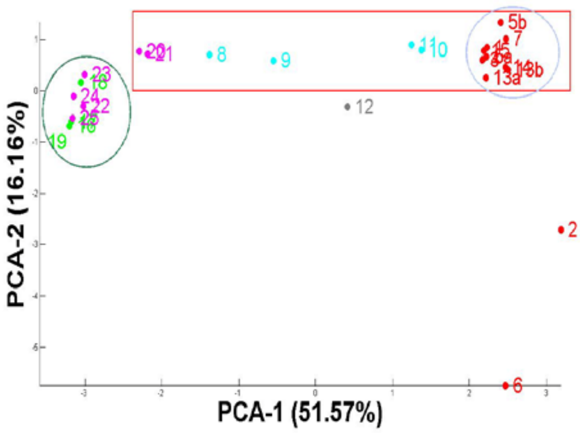


fig.5 Average Actions - English Verbs (colors according to Levin's classes)

probably Levin's classification has taken into account the criterion of intentionality, among others; however, intentionality is still not measurable with sensorimotor techniques).

Based on these first observations, we focus on our ongoing work that aims at proposing a framework that would establish joint-angle-based representations for parameters widely used in linguistic descriptions/classifications, such as path and directionality. The existing literature is mainly constricted in revealing the path of the action based on the gaze (Clark *et al.*, 2000). Intuitively, we could assume that directionality is mainly shown by the gross motion, e.g. head and chest movement or the upward and downward motion derived from the y axis of the knees, but further analysis needs to be done. Last but not least, we will extend our work by projecting the actions of each linguistic class separately and reduce their feature space, in order to focus on the most important synergies.

Table1 Verbs according to classes

Greek (Antonopoulou, 1987)			English (Levin, 1993)			English (FrameNet)	
Perpato verbs (treho can be the head of its own class)	Perpato	1	Run verbs (+meander)	Walk	1	Self motion	Walk
	Vimatizo	2		March	2		March
	Pisopato	3		Step back	3		Step back
	Triklizo	4		Stagger	4		Stagger
	Draskelizo	5		Stride/step over	5a/5b		Stride/step over
	Mpousoulo	6		Crawl	6		Crawl
	Treho	7		Run	7		Run
Aneveno verbs (upward motion)	Aneveno	8	Go verbs	Go up (step)	8	Motion	Go up (step)
	Aniforizo	9		Go up (ramp)	9		Go up (ramp)
Kateveno verbs (downward motion)	Kateveno	10		Go down (step)	10		Go down (step)
	Katiforizo	11		Go down (ramp)	11		Go down (ramp)
	Hamilono (only with the body)	12	Verb of assuming a position	Crouch	12	Posture	Crouch
Pido verbs	Pido (epi topou)	13	Run verbs	Jump/hop	13a/13b	Self motion	Jump/hop
	Pido (pano apo)	14		Jump over	14		Jump over
	Pido (apo kapou)	15		Jump down	15		Jump down
Katevazo verbs (downward motion)	Katevazo	16	Only the combination of the two verbs expresses the same with the Greek	Pick up and put on (lower)/ lower onto	16		Pick up and put on (lower)/ lower onto
Anevazo verbs (upward motion). Though sikono can form its own class	Anevazo	17		Pick up and put on (higher)/ lift onto	17		Pick up and put on (higher)/ lift onto
	Sikono	18	Lift verbs	Lift/raise	18	Body movement (raise not included in the same group)	Lift/raise
	Ipsono	19		Lift high	19		Lift high
$Girizo_1$verbs (rotary motion)	Girizo (antithetic katefthinsi)	20	Roll verbs (around an axis turn, rotate, circle)	Turn around	20	Change direction	Turn (as verb of changing direction)
	Girizo (e.g.giro apo karekla)	21		Circle (e.g. chair)	21	Motion	Circle (e.g. chair)
$Girizo_2$verbs (cause to turn)	Girizo (e.g. selida)	22		Turn (e.g. page)	22	Cause to move in place	Turn (as verb that cause to move in place)
	Peristrefo	23		Rotate	23		Rotate
	Anapodogirizo	24		Turn over	24		Turn over
	Kilo	25		Roll	25	Cause motion	Roll

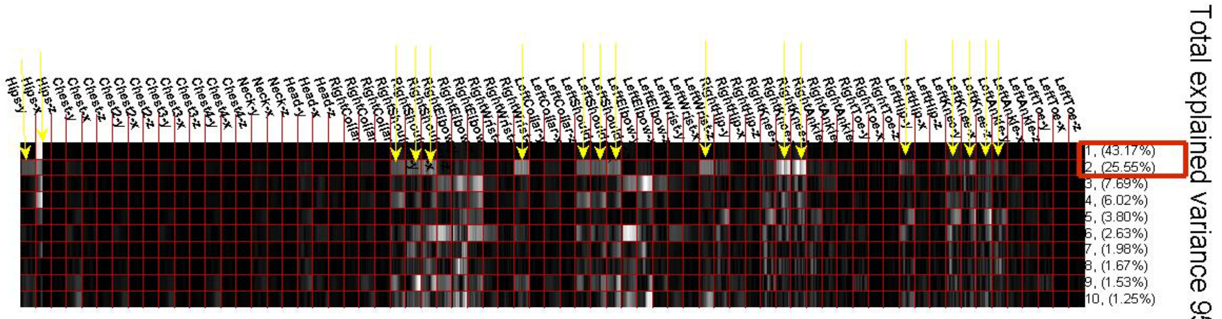


fig.6 Weights of each joint-direction-time feature according to the 10-first principal components (Greek Verbs)

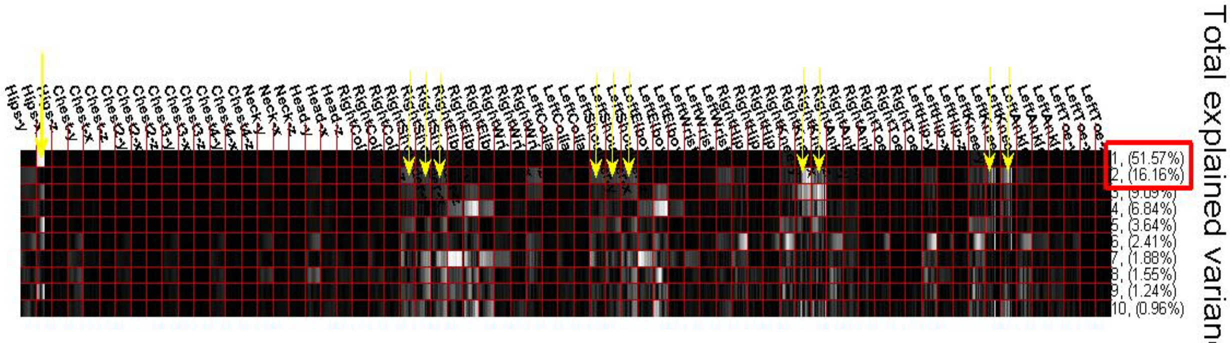


fig.7 Weights of each joint-direction-time feature according to the 10-first principal components (English verbs)

References

Antonopoulou, E. (1987). Prototype theory and the meaning of verbs, with special reference to Modern Greek verbs of motion. Unpublished PhD thesis. University of London

Arbib, M. A. (2008) From Grasp to Language: Embodied Concepts and the Challenge of Abstraction. Journal of Physiology Paris 102: pp. 4-20

Baker, C.F & Ruppenhofer, J. FrameNet's frames, vs, Levin's verbs classes.20-09-2010 <framenet.icsi.berkeley.edu/papers/bakerrup.pdf >

Clark, R.A., Miller, J.M. & Demer, J.L. (2000). Three-dimensional location of human rectus pulleys by path inflections in secondary gaze positions. Invest. Ophthalmol. &Visual Sci. 41

Fadiga, L., et al. (2006). Language in shadow. Social Neuroscience, 1(2): pp.77-89.

Feldman, J.(2006). From Molecule to Metaphor. Cambridge, MA: MIT Press.

Framenet http://framenet.icsi.berkeley.edu/

Galen, A, Grenager. T. & Manning C. (2004). Verb Sense and Subcategorization: Using Joint Inference to Improve Performance on Complementary Tasks. EMNLP 2004: pp. 150-157.

Jackendoff, R. (1990). Semantics Structures. Cambridge, MA: MIT Press.

Kemmerer, D. (2006). Action verbs, argument structure constructions, and the mirror neuron system. In M. Arbib (Ed.), Action to language via the mirror neuron system. Cambridge University Press.

Lakoff, G.& Johnson, M. (1999). Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought. Basic Books, New York

Levin, B. (1993). English Verb Classes and Alternations: A Preliminary Investigation, University of Chicago Press, Chicago, IL.

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. Cognition 98: B75-87

Santello, M., Flanders, M., & Soechting, J.F. (1998). Postural hand synergies for tool use. Journal of Neuroscience, 18.

Troje, N.F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. Journal of Vision, 2: pp. 371–387

Wermter, S., Weber, C., Elshaw, M., Gallese, V. & Pulvermüller, F. (2005). Grounding neural robot language in action. Biomimetic Neural Learning for Intelligent Robots. Springer

AnCora-Net: Mapping the Spanish Ancora-Verb lexicon to VerbNet

Mariona Taulé, M. Antònia Martí, Oriol Borrega
Centre de Llenguatge i Computació (CLiC)
University of Barcelona
Gran Via, 585, 08007 Barcelona.
{mtaule,amarti,oriol.borrega}@ub.edu

Abstract

We present a new multilingual resource, the mapping of AnCora-Verb lexicon to VerbNet, with the aim of enriching both resources with lexical information. A two-way case study is conducted linking first, the VerbNet 13.1 class to the corresponding AnCora-Verb senses, and then a sample of AnCora-Verb senses, belonging to the broader a32 class, to VerbNet. The aim is to analyze the compatibility of both resources. The results show a neat correspondence. Taking into account that VerbNet is related to other semantic resources (PropBank, FrameNet, WordNet, and OntoNotes), AnCora-Verb can be enriched with this information. In the same way, Verb-Net can benefit from the more general classes used in the Spanish resource, and will be related to equivalent verbs in another language that can be useful for comparative studies.

1 Introduction

This paper describes the mapping of AnCora-Verb, a Spanish verbal lexicon, to VerbNet in the framework of the AnCora-Net project. The aim of this project is the integration of the EPEC-ADI Basque lexicon (Aduriz et al., 2006) and the Catalan and Spanish AnCora-Verb lexicons (Aparicio et al. 2008) into the English *Unified Verb Index* (UVI)[1], giving rise to a multilingual ─English, Spanish, Catalan and Basque─ verbal lexicon. This mapping will enrich the contents of the different lexical resources involved by incorporating the semantic information included in each of them.

UVI is constituted by several English resources– VerbNet (Kipper-Schuler, 2006), PropBank (Palmer et al., 2005), FrameNet (Baker et al., 2003), OntoNotes Sense Groupings (Hovy et al., 2006), and WordNet (Fellbaum 1998). The mapping between our lexicons and UVI will be done through PropBank and VerbNet at sense level. This choice is determined by the fact that these resources are the closest to AnCora-Verb. The Spanish lexicon follows the same annotation scheme as PropBank for argument structure, and is organized in semantic classes like VerbNet.

We hypothesize that these resources, though different in the way they semantically classify and represent verbal predicates, can still be compatible and complementary. Moreover, having these resources connected can provide a 'more complete picture of the meaning aspects of a verb' (Čulo et al., 2008).

Given that VerbNet provides the semantic classes for PropBank, we have conducted a case study to analyse the compatibility between AnCora-Verb and VerbNet semantic classifications. We deal with the complex task of comparing and linking two different semantic classifications within two different languages. In particular, we address the following aspects: a) how compatible is the fine grained classification proposed in VerbNet with the coarser grained one adopted in AnCora-Verb; b) whether the VerbNet and AnCora-Verb class mapping is compatible and consistent; and, finally c) how do arguments and thematic roles map.

The remainder of the paper is organized as follows. Section 2 describes the AnCora-Verb lexicons. Section 3 presents the case study carried out, and the analysis of results obtained are discussed in section 4. Finally, in section 5 conclusions, final remarks, and future work are presented.

2 AnCora-Verb lexicons

Ancora-Verb consists of two verbal lexicons, one for Spanish and one for Catalan, used as

[1] http://verbs.colorado.edu/verb-index/index.php

the basis for the semantic annotation of AnCora corpora with arguments and thematic roles (Taulé et al., 2008).[2] In the AnCora-Verb lexicons[3], the mapping between syntactic functions, arguments and thematic roles of each verbal predicate is established taking into account the verbal semantic class and the diathesis alternations in which the predicate can participate. This mapping was manually encoded, and tests of inter-annotator agreement were carried out in order to ensure consistency in the description of predicates. Each verb is divided into different senses ('01, 02... n') and each sense is related to one or more syntactic-semantic frames (transitive, ditransitive, unnaccusative, inergative, etc.). Each frame is characterized according to the four ontological event classes (accomplishment, achievement, states, and activities), and to the diathesis alternations in which a sense can occur.

We only considered very productive diatheses, such as causative/inchoative, active/passive, resultative, oblique subject, transitive/intransitive, object extension, cognate object, and beneficiary alternation (Vàzquez et al., 2000).[4]

The semantic relation that each argument maintains with the event denoted by the verbal predicate is defined by thematic roles. We adopted a set of 20 different thematic roles, each of which can be mapped to several syntactic functions and argument positions.[5] The set of thematic roles is a subset of the 30 different thematic roles used in VerbNet. Following PropBank, the arguments required by the verb sense are incrementally numbered (arg0, arg1, arg2, etc.), expressing their degree of proximity in relation to their predicate. These criteria give rise to a coarse grained verbal classification with a total of 22 different classes compiled.

The Spanish AnCora-Verb lexicon consists of 2,821 lemmas corresponding to 3,934 different senses and 5,481 syntactic-semantic frames.

```
<lexentry lemma="prestar" lng="es" type="verb">
 <sense id="1">
  <frame default="yes" lss="c21" type="state-attributive">
   <argument argument="arg1" function="suj"
    thematicrole="tem"/>
   <argument argument="arg2" function="creg"
    thematicrole="atr"/>
     <constituent type="sp" preposition= "a"/>
    </argument>
   <examples>
    <example>Comparar la piel de los mamíferos y anfibios se presta a una mayor polémica[6]</example>
    </examples>
  </frame>
 </sense <sense id="2">
  <frame default="yes" lss="a32" type="ditransitive">
   <argument argument="arg0" function="suj"
    thematicrole="agt"/>
   <argument argument="arg1" function="cd"
    thematicrole="pat"/>
    <argument argument="arg2" function="ci"
    thematicrole="ben"/>
     <constituent type="sp" preposition= "a"/>
    </argument>
  <examples>
    <example>El Banco Interamericano
 de Desarrollo prestará a Perú 120 millones de
dólares[7] como apoyo a un programa de mejora de la
calidad y cobertura de la educación secundaria en el
país andino</example>
 </examples>
</frame>
<frame default="passive" lss="b12" type="unaccusative-passive-ditransitive">
    <argument argument="arg1" function="suj"
    thematicrole="pat"/>
    <argument argument="arg2" function="ci"
    thematicrole="ben"/>
 …</sense>...
</lexentry>
```

Figure 1: AnCora-Verb lexical entry of *prestar*

Figure 1 shows the information associated with the entry *prestar* 'to lend' in the AnCora-Verb-Es lexicon: the lemma ("prestar"); the category ("verb"); the different senses associated to their corresponding semantic classes ─in this case, the first sense with the state-attribu-

[2] AnCora is a multilingual corpus consisting of 500,000 words in Catalan (AnCora-CA) and Spanish (AnCora-ES) annotated with morphological (POS), syntactic (constituents and functions), semantic (argument structure, thematic roles, named entities, and nominal synsets) and coreference information (Recasens & Martí, 2010). The corpora are freely available at http://clic.ub.edu/corpus/ancora.

[3] http://clic.ub.edu/corpus/ancora-lexics

[4] The specific alternations shared by few verbs were not considered because they do not define general verb classes.

[5] The list of the different thematic labels is the following: "agt" (Agent), "cau" (Cause), "exp" (Experiencer), "scr" (Source), "pat" (Patient), "tem" (Theme), "cot" (Cotheme), "atr" (Attribute), "ben" (Beneficiary), "ext" (Extension), "ins" (Instrument), "loc" (Locative), "tmp" (Time), "mnr" (Manner), "ori" (Origin), "des" (Goal), "goal" (Purpose), "ein" (Initial State), "efi" (Final State) and "adv" (Adverbial).

[6] 'To compare the skin of mammals and amphibians lends itself to a more controversial'.

[7] 'The Inter-American Development Bank will lend 120 million U.S. dollars to Peru'.

tive semantic class (lss="c21"), and the second with the ditransitive class (lss="a32")[8]; the mapping between syntactic function, argument and thematic role (for instance, in the second sense the subject "suj" correspond to the first argument "arg0" with the thematic role of agent "agt"); and, the diatheses alternations in which the verb occurs (in the second sense, *prestar* can appear in passive "unaccusative-passive-ditransitive"). As we can observe, the expression of the passive alternation entails an argument crossing: the affected object, appears as direct object ("cd") in the active structure and as subject ("suj") in the passive structure, being in both cases the argument ("arg1") with the thematic role of patient ("pat"). Furthermore, the expression of this alternation also involves an aspectual change, since the active reading corresponds with an accomplishment (lss="a32") and the passive reading with an achievement (lss="b12"). Finally, examples are also included.

3 A case study

The case study consists in comparing and analyzing the compatibility of AnCora-Verb and VerbNet. It is conducted in two phases: 1) First, a VerbNet class is selected and all the members of this class are mapped to the corresponding verb senses in AnCora-Verb. 2) Secondly, the AnCora-Verb classes involved in the mapping process are obtained and then the ‘inverse’ mapping is conducted. That is, the remaining verb senses belonging to the selected AnCora-Verb classes are mapped to VerbNet. Notice that the mapping is established between PropBank senses, which are linked to VerbNet semantic classes, and AnCora-Verb senses.

We assume the following hypotheses: a) the VerbNet class will map to only one AnCora-Verb class; and b) the broader AnCora-Verb class will map into a restricted set of VerbNet classes. This set will share some characteristics being able to constitute a more general class in VerbNet.

[8] The first sense (*prestarse a*) corresponds to a pronominal verb with the meaning of ‘to lend itself; to accommodate or offer (itself) to: your words lend themselves to confusion’, while the second sense refers to ‘to permit the use (of something) with the expectation of return to the same or an equivalent: she lent me the book’.

3.1 Mapping VerbNet to AnCora-Verb

The linking of VerbNet[9] to AnCora-Verb is carried out automatically mapping the two resources through WordNet 3.0 and PropBank, and then manually validating the mapping obtained, or connecting those verbal senses without automatic correspondences. WordNet 3.0 is used as a bilingual lexicon and PropBank allows for the argument mapping. The mappings between VerbNet and WordNet 3.0 and PropBank are already established in UVI.

For this case study we selected the verbs belonging to VerbNet class ‘give 13.1’ (i.e., lend, loan, pass, peddle, refund, render) and to the subclass ‘give 13.1-1’ (i.e., give, hock, lease, pawn, rent, sell). They are defined as verbs of change of possession most of which display the dative alternation (Levin, 1993). This kind of alternation does not occur in Spanish. The order of constituents can vary (*El Banco prestará 120 millones a Perú* vs. *El Banco prestará a Perú 120 millones*)[10], but Spanish does not allow for the double object construction (**El Banco prestará Perú 120 millones*) as in English. All verbs belonging to these classes map to the *a32* AnCora-Verb class (lss=“a32” in figure 1), that is, the class including ditransitive agentive benefactive verbs. Figure 2, summarizes the syntactic and semantic properties of the *a32* class.

LSS:
[[x do-something] cause [become[y<place[11]>z]]]
Argument Structure:
arg0=agt
arg1=pat
arg2=ben
Diatheses Alternations:
[+causative], [+impersonal],
[+/-resultative], [+passive]
Spanish verbs: *enviar* (to send), *dar* (to give), *prestar* (to lend/loan), *vender* (to sell/peddle), *alquilar* (to rent), *rendar* (to lease), *decir* (to say)...

Figure 2: *a32* AnCora-Verb class

a32 class associates a causer argument (*x*) with the semantic predicate *do*, and a third argument (*z*) with a location in space with semantic traits such as [+animate] or [+human].

[9] We use VerbNet 3.0.

[10] ‘The Bank will lend 120 millions to Perú’ vs. ‘The Bank will lend (to) Perú 120 millions’

[11] ‘place’ has to be interpreted in a very general way, including the goal of physical as well as verbal transfer.

Since all the verbs belonging to this class allow for passive alternation, argument *x* is referred to as arg0-agent and argument *y* as arg1-patient. Argument *z* is referred to as arg2-beneficiary. arg0 is syntactically the subject (*suj*), while arg1 is the direct object (*cd*), and arg2 the indirect object (*ci*) in the default frame. This class accounts approximately for 3% of the verbs represented in AnCora-Verb.

3.2 Mapping AnCora-Verb to VerbNet

In order to carry out the inverse mapping, the 30 most frequent *a32* AnCora-Verb class senses are selected (20% of the total) and mapped to VerbNet. In this case study, this mapping is manually done. In the AnCoraNet project it is carried out automatically. In figure 3 there is an example of the mapping between the involved resources. The tool used throughout all the process is AnCora-Pipe (Bertran et al., forthcoming).

Figure 3: a sample of the mapping

As Figure 3 shows, the mapping is established at two levels: sense and arguments. Sense mapping can be many-to-many (one AnCora-Verb sense may map to more than one VerbNet senses and vice-versa). In the bottom row of the sample, the link is illustrated: from *vender.1* (*a32*) to *sell.01* (*13.1-1*). As for argument mapping, the link is established taking into account argument position (*arg0, arg1,...*) as well as, primarily, thematic role (*agent*, *patient,...*). That is, the anchor of the mapping is argument position, but what determines the final link is thematic role. As highlighted in the sample, a non-argumental complement (*argM*) in AnCora-Verb maps to an argumental complement (*arg3*) in VerbNet because they share the same thematic role (*adv~price paid*).

The members of AnCora-Verb *a32* class map to 20 different VerbNet classes. Half of them map to 'say 37.7' (6 links), 'reflexive_appearance 48.1.2' (6 links) and 'order 60' (4 links) classes. The remaining ones map to 17 different classes (one or two links each) as shown in table 1.

Matches	VerbNet class	AnCora-Verb
6	37.7	Responder, confesar, declarar, afirmar, proponer, manifestar
6	48.1.2	Proponer, presentar, mostrar, manifestar, expresar, afirmar
4	60	Pedir, solicitar, reclamar, exigir
2	11.1	Entregar, enviar
2	13.1	Vender, dar
2	13.3	Ofrecer, conceder
2	78	Indicar, mostrar
1	10.5	Quitar
1	13.2	Devolver
1	22.2	Presentar
1	37.1.1	Explicar
1	37.1.2	Preguntar
1	37.10	Confesar
1	37.11	Comentar
1	37.13	Asegurar
1	37.9	Asegurar
1	58.2	Pedir
1	63	Imponer
1	64	Permitir
1	68	Pagar
Ø	No match	Atribuir

Table 1: AnCora-Verb to VerbNet mapping

The first column in Table 1 displays the number of matches. The second one, VerbNet class codes, which correspond to: steal (10.5), send (11.1), give (13.1), contribute (13.2), future_having (13.3), amalgamate (22.2), transfer_mesg (37.1.1), inquire (37.1.2), say (37.7), advise (37.9), confess (37.10), lecture (37.11), promise (37.13), reflexive_appearance (48.1.2), beg (58.2), order (60), enforce (63), allow (64), pay (68) and indicate (78). Finally, the third column displays AnCora-Verb lemmas involved in the mapping. There are some senses not represented in VerbNet, so the mapping of certain AnCora-Verb senses is not possible. As seen in the last row of Table 1, no satisfactory match was found for *atribuir* ('to attribute' is not represented in VerbNet, though it is in PropBank). In next section we analyze and discuss these results.

4 Analysis of results

The results confirm that, in principle, verbs belonging to one VerbNet class (‘give 13.1’ and ‘give 13.1-1’) correspond to a unique AnCora-Verb class (*a32*). In this case we have a neat correspondence with no dispersion in the class members. The inverse mapping presents a completely different scenario: the 30 verbs of *a32* AnCora-Verb class map into 20 different VerbNet classes.

At first sight, this result could be interpreted as a lack of consistency between both resources, but in a deeper analysis some interesting generalisations arise. First, the VerbNet verbs mapped to *a32*, in spite of belonging to 20 different classes, are all ditransitive. This information is not explicitly declared in VerbNet, but it is captured thanks to the mapping with AnCoraVerb. The different criteria applied to the creation of these resources make them complementary: VerbNet is a semantic classification while AnCora-Verb is more syntax-oriented. The mapping between these resources makes information that is not explicitly declared evident, as for instance the ditransitivity shared by verbs coming from different classes in VerbNet. In a similar way, AnCora-Verb is enriched with more fine grained semantic information.

Second, all the verbs involved in the mapping process have three arguments: an arg0 agent, an arg1 theme and a third argument, that can be recipient, beneficiary, destination or patient (all of them with the feature [+animate]) depending on the meaning component represented in the corresponding VerbNet classes. For instance, in ‘give 13.1’ class the basic meaning component is ‘transfer’ and the third argument is a recipient, while in ‘send 11.1’ class the basic meaning components are motion and location, and the third argument is a destination. Despite the fact that these VerbNet classes grouped in the *a32* AnCora-Verb class have different semantic components, and that the verbs display different alternations, it is also true that they share a general syntactic-semantic behaviour and, probably the common and more general meaning component of ‘transfer’ (Vàzquez et al., 2000).

5 Conclusions and remarks

The case study presented in this paper shows that the mapping between AnCora-Verb and VerbNet, two verbal semantic classifications within two different languages and based on different criteria, is a way to enrich and complement the contents of the involved resources, showing that the information included is compatible, consistent and complementary. VerbNet can benefit from the more general classes used in AnCora-Verb to create a coarser class classification capturing generalizations currently scattered throughout different VerbNet classes, for instance ditransitivity. AnCora-Verb classes can be grouped in a more fine grained subclasses considering VerbNet classification. Moreover, taking into account that VerbNet is also mapped to PropBank, FrameNet, OntoNotes, and WordNet in UVI, AnCora-Verb will be enriched with semantic information provided by these resources (VerbNet class, PropBank rolesets, FrameNet conceptual frames, WordNet synsets, OntoNotes Sense Groupings). Furthermore, all resources commented will also be enriched with the linking of Catalan AnCora-Verb lexicon, and the Basque EPEC-ADI lexicon, which are already being built. Finally, the resources in UVI will be related to equivalent verbs in three more languages, giving rise a multilingual verbal classification.

Acknowledgments

This research has been supported by the projects TEXT-MESS 2 (TIN2009-13391-C04-04) and AnCora-Net (FFI2009-06497-E/FILO) from the Spanish Ministry of Education and Science.

References

Aduriz I., Aranzabe M., Arriola J., Atutxa A., Díaz de Ilarraza A., Ezeiza N., Gojenola K., Oronoz M., Soroa A., Urizar R. , 2006. Methodology and steps towards the construction of EPEC, a corpus of written Basque tagged at morphological and syntactic levels for the automatic processing. *Language and Computers*. Vol 56.

Baker, C., Collin F., Charles J. Fillmore, B. Cronin 2003. The Structure of the Framenet Database. *International Journal of Lexicography*, Volume 16.3: 281-296.

Bertran, M., M.A. Martí, M. Taulé, O. Borrega. 2010. AnCoraPipe: a new tool for corpora annotation. Forthcoming.

Čulo, O., K. Erk, S. Padó, S. Schulte im Walde. 2008. Comparing and combining semantic verb classifications. *Language Resources and Evaluation*, 42: 265-291, Spinger.

Fellbaum, C., ed. 1998. WordNet: An Eletronic Lexical Database. *Language, Speech and Communications*. MIT Press, Cambridge, Massachusetts.

Hovy, E., et al., 2006. OntoNotes: the 90% solution. Proceedings of the Joint Human Language Tecnhology and North American Association of Computational Linguistics Conference. New York.

Kipper, K., M. Palmer, O. Rambow. 2002. Extending PropBank with VerbNet Semantic Predicates. *Workshop on Applied Interlinguas*, held in conjunction with AMTA-2002. Tiburon, CA.

Levin, B. 1993. *English Verb Classes and Alternations: A preliminary investigation*. University of Chicago Press, Chicago.

Palmer, M., P. Kingsbury, D. Gildea. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, volume 21, MIT Press, USA.

Recasens, M., Martí, M.A. (2010) AnCora-CO: Coreferentially annotated corpora for Spanish and Catalan. *Language Resources and Evaluation*. Springer Science.

Vàzquez, G., A. Fernández, M.A. Martí . 2000. Clasificación verbal. Alternancias de diátesis. Ed. Universitat de Lleida , Spain.

(Ir)regularity of verbs revisited: Evidence for a lexical entry complexity based account

Helena Trompelt
University of Potsdam
Potsdam, Germany

trompelt@uni-potsdam.de

Denisa Bordag
University of Leipzig
Leipzig, Germany

denisav@uni-leipzig.de

Thomas Pechmann
University of Leipzig
Leipzig, Germany

pechmann@uni-leipzig.de

Abstract

To explore the representation and encoding of regularity as well as the inflectional processes involved in the production of regular and non-regular verbs, we investigated three groups of German verbs: regular, irregular and hybrid verbs. In a picture naming experiment and a picture word interference experiment, articulation latencies were measured while participants named pictures of actions, producing the 3^{rd} person singular of German verbs in present and past tense. The differences in naming latencies in the three groups of verbs in the two tenses suggest that the complexity of lexical entries of verbs is a decisive factor in the production of verbs. We propose a lexical entry complexity account which can explain the pattern of the presented data while the blocking mechanism (e.g. Pinker, 1991; Clahsen, 1999) cannot.

1 Introduction

One of the prominent accounts of processing irregular vs. regular verbs, the Dual Mechanism Model (DMM: Pinker, 1991, 1998; Jaeger, 1996; Clahsen, 1999), assumes different routes for processing regular and irregular inflection. Whereas regular forms are generated by concatenating verb stems and corresponding suffixes in a rule-governed process, irregular forms are stored as ready-made entries in the lexicon and must be looked up individually. The standard finding that the production of irregular verbs takes longer than that of regular verbs is explained in the following way. The regular route is a default mechanism that starts to process each verb irrespective of its (ir)regularity status. In the case of the irregular forms, this regular default mechanism must be blocked and the irregular form is retrieved from the lexicon. Suppressing the rule is costly and time-consuming (Pinker, 1991; Jaeger, 1996; Clahsen, 1999), which explains why the production of irregular forms takes longer.

An interesting psycholinguistic contribution on the core issue about the structure of lexical entries of verbs is made by Clahsen (1999) and Clahsen et al. (2002) adopting the approach of Minimalist Morphology (Wunderlich, 1996). Minimalist Morphology teases apart regular inflection and lexically driven inflection and assumes two qualitatively distinct linguistic mechanisms for them. While regular inflection is pursued by a combinatorial affixation process, it is claimed (Wunderlich & Fabri, 1995), that irregular past tense forms, e.g. *ran*, are represented as subnodes of lexical entries. Based on these concepts, Clahsen and colleagues argue that irregular participles such as (*ge*)*trunken* [drunk] are mentally represented as structured, underspecified lexical entries. Stem alternants are represented as subnodes of a hierarchically higher mothernode. Subnodes are underspecified feature pairs formed upon the pattern <phonological string, morphological feature value> which get features from the mother node by inheritance. The subnodes are shared by verbs of the same class.

Previous studies did not consider the representation of (ir)regularity itself: whether is represented as a property of individual forms (implicit assumption of DMM) or of whole inflectional paradigms. The following experiments investigated hybrid German verbs to test these two hypotheses.

2 Study

As already suggested, there are some aspects of the DMM that are in need of further inspection. The role of (ir)regularity has been explored so far only in the past tense, so that it is not clear whether the (ir)regularity of a verb (paradigm) per se plays the decisive role, or whether the concept of (ir)regularity is bound to the individual irregular forms. In our research we explored the production of three types of German verbs in past and present tense to differentiate between the two options and to clarify further issues concerning the production of regular and irregular verbs.

The German verb system is organised in a greater diversity than shown so far and greater than for example the English one. It is set up by *three* basic paradigms: the *regular* paradigm and the non-regular paradigm which is comprised of *hybrid* and *irregular* verbs. The second and third type of verbs is traditionally labelled as *irregular* verbs disregarding possible dissociations. Regular verbs (e.g. *spielen* [play]: *er spielt, er spielte, er hat gespielt*) have only one stem and take regular affixes in both past (–te) and present tense (-t). Irregular verbs have several stems and take on irregular forms both in present and past tense (e.g. *brechen* [break]: *er bricht, er brach, er hat gebrochen*). Hybrid verbs also have more than one stem but their present conjugation is completely regular, while their past forms are irregular (e.g. *singen* [sing]: *er singt, er sang, er hat gesungen*).

2.1 Material

Nine intransitive German verbs were chosen for each type of verb. The three groups were equated in terms of word form and lemma frequency, length, initial phoneme, ablaut patterns and transitivity. Verbs containing allomorphy (ə-epenthesis *<u>blut</u>en, ich <u>blut</u>e, ich <u>blute</u>te* versus none in *<u>lach</u>en, ich <u>lach</u>e, ich <u>lach</u>te*) were excluded to avoid ə-epenthesis to affect reaction times. Actions were depicted in black and white line drawings. Some were taken from Masterson & Druks (1998), but several were designed for this purpose in the same style and comparable complexity (see Figure 1).

2.2 Picture naming experiment

In Experiment 1, participants named pictures of actions in the 3rd person singular present or past tense within a sentential context in a picture naming paradigm - a task that involves conceptualisation and avoids possible priming between the presented and elicited forms. Tense was blocked and counterbalanced across subjects. Measurement of articulation latencies started when the picture appeared on the screen. A voice key was triggered by the first phoneme of the participants' utterances. Wrong namings, hesitations or technical errors during measurement were excluded from the analyses.

A two-way repeated measures ANOVA yielded significant main effects of Regularity, [$F1(2,70) = 116.61$, $MSE = 3579.28$, $p < .001$; $F2(2,16) = 21.61$, $MSE = 5312.15$, $p < .001$] and Tense [$F1(1,35) = 41.41$, $MSE = 4058.29$, $p < .001$; $F2(1,8) = 168.14$, $MSE = 254.98$, $p < .001$]. The interaction between Regularity and Tense reached significance by subjects and very scantily by items [$F1(2,70) = 6.92$, $MSE = 1153.38$, $p < .01$; $F2(2,16) = 3.64$, $MSE = 485.99$, $p = .05$]. A post hoc Scheffé-Test ($diff_{crit;\ p<.05} = 20.2$) revealed that reaction times do not differ between irregular and hybrid verbs and that their articulation latencies depend on the factor Tense (production is faster in present tense) whereas that particular Tense effect is not significant for regular verbs.

Regular verbs in the present and past tense were produced significantly faster than all other verbs (see Table 1). Crucially, the naming latencies of hybrid and irregular verbs did not differ from each other in both tenses suggesting that (ir)regularity is not a property of individual verb forms, but generalizes to all forms within a paradigm.

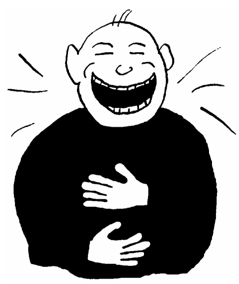


Figure 1. Example for stimuli in Experiment 1. Pictures were named with inflected verbs (e.g. *lacht* [is laughing]) in present and past tense in a sentential context provided by the pronoun *jemand* [somebody].

Table 1. Mean Response Latencies (RT, in Milliseconds, standard deviations in parentheses), (Experiment 1).

	Regularity			
Tense	**irr**	**hyb**	**reg**	**M**
past	638	619	480	577
	(185)	(171)	(106)	(171)
present	577	547	447	520
	(170)	(158)	(93)	(153)
M	606	583	462	549
	(180)	(169)	(101)	(165)

2.3 Picture-word-interference experiment

In Experiment 2, we tested whether (ir)regularity, once not bound to individual verb forms, is represented in form of abstract (ir)regularity nodes, as assumed for gender or conjugational class (Levelt et al., 1999; Bordag & Pechmann, 2009). In a picture-distractor paradigm, participants named pictures of actions with verbs in the 3rd person singular present and past tense (same material as in Experiment 1). Additionally, a written distractor verb appeared over or above the picture which should be ignored by the participants. In the congruent condition, the picture and the distractor were either both regular or non-regular, in the incongruent condition they differed in regularity. An identical condition where the name of the picture served as distractor was applied as control condition. Materials were counterbalanced so that each item appeared in each condition. We expected an (ir)regularity congruency effect (slower RTs in the incongruent condition), reflecting competition between abstract grammatical features for (ir)regularity. The experiment proceeded as Experiment 1.

The statistical analysis revealed effects that corresponded to those of Experiment 1 with differences that were expected due to the changes of the paradigm (see Table 2). The 3x3x2 ANOVA showed main effects for Distractor [$F1(2,34) = 67.62$, $MSE = 4085.17$, $p < .001$; $F2(2,16) = 110.20$, $MSE = 1275.68$, $p < .001$], Regularity [$F1(2,34) = 162.39$, $MSE = 3037.14$, $p < .001$; $F2(2,16) = 13.87$, $MSE = 18869.99$, $p < .001$] and Tense [$F1(1,17) = 5.03$, $MSE = 5603.75$, $p < .05$; $F2(1,8) = 18.37$, $MSE = 991.41$, $p <.01$].

The regular present forms of the hybrid verbs were again produced more slowly than the regular present forms of the regular verbs. Moreover, there was no statistical difference between naming latencies of regular present tense forms of hybrid verbs and the irregular present tense forms of the irregular verbs. All main effects were significant and replicated the results of Experiment 1. However, the critical conditions did not exhibit the expected congruency effect. Consequently, we assume that rather than through abstract node representation, the paradigmatic effects could be explained as a result of complexity of lexical entries and (ir)regularity might be coded by the lexical entries' complexity.

Table 2. Mean Response Latencies (RT, in Milliseconds, standard deviations in parentheses), (Experiment 2).

Past Tense	Regularity		
Distractor	**irr**	**hyb**	**reg**
identical	676 (111)	658 (104)	610 (91)
incongruent	735 (107)	727 (107)	656 (92)
congruent	730 (106)	722 (104)	652 (84)
M	714 (111)	703 (109)	639 (91)
Present tense			
identical	665 (110)	660 (112)	703 (109)
incongruent	709 (119)	703 (115)	643 (94)
congruent	714 (112)	710 (105)	644 (87)
M	699 (113)	691 (113)	630 (93)

3 Conclusion

We argue that postulating two different mechanisms for the processing of regular and irregular inflection (DMM) and a blocking mechanism cannot account for all data, in particular not for the fact that even regular forms of hybrid verbs are produced more slowly than regular forms of the regular verbs. The results are most likely not due to general form effects, for which the material was carefully controlled.

We propose that the crucial explanatory factor for the observed results is the complexity of the lexical entry: If a verb has alternating stems (irregular and hybrid verbs), the retrieval of the appropriate one takes longer than the retrieval of a single stem entry (regular verbs). The generation of the correct word form is more costly for hybrid and irregular verbs because more stems are related to their lemmas, e.g. *brechen* [to break]: *brech-e*, *brich-st*, *brach-∅*, *ge-broch-en*, *bräch-e*. Hence, compared to regular verbs, *selection* is necessary to access non-regular verbs as opposed to the mere lexical *retrieval* of a single stem from a single lemma.

This proposal is consistent with Clahsen (1999) and Clahsen et al. (2002) and can be extended with the new empirical data to the present tense. It assumes internally structured lexical entries in form of feature pairs of phonological and morphological information.

Blocking was a promising attempt by Pinker & Prince (1994) and Pinker (1999) to explain empirical data. However, the blocking mechanism of the DMM cannot account for the fact that even regular forms of hybrid verbs are produced more slowly than regular forms of regular verbs. According to the DMM, blocking is kind of waiting of a non-regular form for spell out: a quite counterintuitive and uneconomic mechanism. Therefore, once we can explain longer reaction times for non-regular verbs with lexical entry complexity we can abandon the idea of a blocking mechanism (cf. Ockham's razor: the simple explanation with fewer assumptions is the better one).

Whether the inflection for person and number for all three types of verbs proceeds similarly or not is in need of further investigation.

Reference

Bordag, D. & Pechmann, T. (2009). Externality, internality, and (in)dispensability of grammatical features in speech production: Evidence from Czech declension and conjugation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35(2), 446–465.

Clahsen, H. (1999). Lexical entries and rules of language: A multi-disciplinary study of German inflection. *Behavioral and Brain Sciences, 22,* 991-1060.

Clahsen, H., Prüfert, P., Eisenbeiss, S. & Cholin, J. (2002). Strong Stems in the German Mental Lexicon: Evidence from Child Language Acquisition and Adult Processing. In: Ingrid Kaufmann & Barbara Stiebels (eds.), More than Words. A Festschrift for Dieter Wunderlich. Berlin: Akadamie Verlag, 91-112.

Jaeger, J. J., Lockwood, A.H., Kemmerer, D.L., Van Valin, R.D., Murphy, B.W., & Khalak, H.G. (1996). A positron emission tomographic study of regular and irregular verb morphology in English. *Language, 72,* 451-497.

Levelt, W. J. M., Roelofs, A. & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38.

Masterson, J. & Druks, J. (1998). Description of a set of 164 nouns and 102 verbs matched for printed word and frequency, familiarity and age-of-acquisition. *Journal of Neurolinguistics, 11*(4), 331–354.

Pinker, S. (1991). Rules of language. *Science*, *253,* 530-535.

Pinker, S. (1998). Words and rules. *Lingua, 106,* 219-242.

Wunderlich, D. (1996). Minimalist morphology. The role of paradigms. *Yearbook of Morphology 1995*, 93–114.

Wunderlich, D. & Fabri, R. (1995) Minimalist morphology: An approach to inflection. *Zeitschrift für Sprachwissenschaft* 14:236–94.

Event Types in the Mind and in the Corpus

Alessandra Zarcone
University of Stuttgart
Stuttgart, Germany
zarconaa@ims.uni-stuttgart.de

Alessandro Lenci
Università di Pisa
Pisa, Italy
alessandro.lenci@ling.unipi.it

Abstract

Event types (ET) have received considerable attention in formal semantics, but their importance in experimental linguistics has developed only recently. The aim of this work is to compare the performance of human annotators and corpus-based models in ET classification of Italian verbs

1 Event Types in experimental linguistics

Event types (ET) play a crucial role in verb semantics, contributing to the temporal constitution of the sentence. We refer here to Vendler's (1967) standard classification of predicates into *states* (STA), *activities* (ACT), *accomplishments* (ACC) and *achievements* (ACH), which can be further cross-classified with respect to the features of dynamicity (DYN), durativity (DUR) and resultativity (RES):

Table 1: Features of Vendler's ETs

ET	[DYN]	[DUR]	[RES]	example
STA	−	+	−	*to know, to be tall*
ACT	+	+	−	*to sing, to walk*
ACC	+	+	+	*to write a book, to walk to the fence*
ACH	+	−	+	*to stumble, to die*

The ET of a sentence is the result of a complex interaction between the verb lexical item and the sentence context (arguments, adjunts, verb morphology) (Verkuyl, 1972); contrast for example *to walk* (ACT) and *to walk to the fence* (ACC). Such an interplay emerges very clearly in ET polysemy and ET coercion, which need to be accounted for by any theory of ETs. ET polysemy (Bertinetto, 1986; Lucchesi, 1971) is a fairly regular phenomenon: some verbs show different ETs in different contexts (e.g. ACH/STA in Italian: *impugnare*, "hold"/"get hold of"; *indossare*, "wear"/"put on"). ET coercion (Pustejovsky, 1995; Rothstein, 2004) occurs when contextual features trigger a reinterpretation of a verb into a new ET class: e.g. The student *ate* two sandwiches (ACT ⇒ ACC, countable direct object with numeral modifier); Guests *have been arriving* for hours (ACH ⇒ ACT, bare plural subject, for x time).

ETs have received considerable attention in formal semantics, but their importance in experimental linguistics has developed only recently. We believe that the study of ETs, like a number of other research areas in linguistics, could benefit from a cross-contamination among different fields and methodologies.

Antinucci and Miller (1976) showed that strong correlations between Aspect and ETs emerge in language acquisition, along the axes of telicity/perfectivity/past-reference and atelicity/imperfectivity/present-reference; such correlations also emerged in the distributional model in Li and Shirai (2000). The correlation, though relaxed, can still be detected in adult language, along with other associations between ETs and context features, by computational models such as those in Zarcone and Lenci (2008) and Im and Pustejovsky (2010). Finocchiaro and Miceli (2002) found an effect of ET on the performance of aphasic subjects, showing a double dissociation between STA and ACT and thus supporting the idea that ETs are one fundamental principle of organization of the mental lexicon in the brain. Behavioral studies have been conducted using the paradigms of self-paced reading (Gennari and Poeppel, 2002), ERP (Bott, 2008) and semantic priming (Bonnotte, 2008; Zarcone and Lenci, 2010).

A close interaction between cognitive methods and corpus-based computational methods seems to be promising to explain how this interplay between contextual elements and verb lexical items influences the speakers in determining the ET of a sentence. In this paper we have a twofold goal. First of all, we are going to test the subjects' competence of ETs in a series of cross-modal annotation experiments for English and Italian. Secondly, we are going to compare the performance of human annotators with results from corpus-based models in a task of ET classification of Italian verbs, in order to investigate potentially interesting differences among ET classes and to evaluate the contribution of cognitive and corpus-based methods to the study of ETs.

2 Competence of Event Types

We carried out four web-based annotation experiments: Experiment 1 and 2 for linguistic stimuli (Italian and English), Experiment 3 and 4 for picture stimuli (with Italian speakers and English speakers). Experiments requiring English speakers (2 and 4) were conducted using the crowdsourcing paradigm[1].

2.1 Design and procedure

Experiment 1: Materials for Experiment 1 were 138 Italian predicates (96 transitive VPs (V + Obj) representing all Vendler's classes and 42 intransitive verbs, being 21 ACH and 21 ACT). 20 native Italian-speaking students performed the test in a web-based format, each of them saw all the stimuli. The procedure was inspired by the pilot study in Bonnotte (2008). Per each event, participants were asked to choose one of four pictures, one representative of each ET (figure 1).

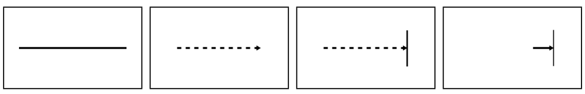


Figure 1: The long continuous line depicts a state that lasts in time, the long dashed arrow depicts a process that develops over a certain period of time, the long dashed arrow ending with a vertical dash depicts a process that develops over a certain period of time and leads to a result, the short arrow ending with a vertical dash depicts a change of state.

[1]Crowdsourcing has been increasingly popular also in linguistics (Snow et al., 2008), allowing for lower economic and logistic costs

Figure 2: ACC (*to peel*), ACH (*to break*), ACT (*to ski*), STA (*to float*).

Experiment 2: Experiment 2 was conducted with the same modality as Experiment 1, but for English. An effort was made to translate the materials for Experiment 1 into English, taking particular care that each English stimulus showed the same ET and the same low degree of ambiguity of its Italian correlate. Materials for Experiment 2 were 134 predicates (96 transitive VPs (24 ACC, 24 ACH, 24 ACT, 24 STA) and 38 intransitive verbs, being 19 ACH and 19 ACT). 10 of the transitive VPs (2 ACC, 4 ACH, 4 ACT) were also presented together with the particle "up" ("up verbs", e.g. "use the materials"/"use up the materials"); so the total number of stimuli was 144. Our intuition was that the particle added an extra element of telicity to ACT VPs such as "use the materials", or simply made the telicity of ACH and ACC verbs more prominent (e.g. "lock the box"/"lock up the box"). 24 native English speakers took part in Experiment 2; as it is usual for crowdsourcing experiments, not all the participants annotated every stimulus. The minimum number of participants annotating each stimulus was 16, maximum 22, with a mean of 18.

Experiment 3: Experiment 3 was conducted with the same modality as Experiments 1 and 2, but picture stimuli was used instead of word stimuli: 111 pictures (19 ACC, 40 ACH, 40 ACT, 12 STA) were selected from the IPNP database (Bates et al., 2000), see Figure 2. 20 native Italian-speaking students took part in Experiment 3, each of them saw all the stimuli.

Experiment 4: Experiment 4 was conducted with the same modality and stimuli as Experiments 3, but participants were native speakers of English (42). The minimum number of participants annotating each stimulus was 10, maximum 16, with a mean of 13.6.

2.2 Results

Agreement: Agreement was computed with Krippendorff's α and with weighted α_w(Krippendorff, 2004). The latter was com-

		base version				"up" version			
item	ET	ACC	ACH	ACT	STA	ACC	ACH	ACT	STA
draw [up] the map	ACC	6	3	8	0	10	4	4	0
dry [up] the cutlery	ACC	17	0	0	0	17	2	0	0
lock [up] the box	ACH	13	1	3	0	15	3	1	0
swallow [up] the syrup	ACH	12	1	5	0	14	1	3	0
tear [up] the table cloth	ACH	3	13	0	0	4	13	0	0
wake [up] the doorman	ACH	7	10	1	0	12	6	1	0
beat [up] the wife	ACT	5	11	2	0	16	1	2	0
eat [up] the strawberries	ACT	6	0	10	1	15	1	2	0
use [up] the materials	ACT	10	0	5	2	12	1	2	2
wait [up] for the verdict	ACT	4	15	0	0	7	11	0	0
		83	54	34	3	122	43	15	2

Table 2: Answers given for "up" verbs

puted in order to modulate disagreement across categories which are not equally distant[2].

Agreement values were above chance and reasonably good for Experiment 1 ($\alpha = .35; \alpha_w = .43$) and Experiment 2 ($\alpha = .46; \alpha_w = .53$), since the subjects were naive to linguistics and ET classification and no sentence context was given; agreement was lower for Experiment 3 ($\alpha = .22; \alpha_w = .31$) and Experiment 4 ($\alpha = .28; \alpha_w = .39$).

Accuracy: Accuracy values are reported in table 3 (please note that for Experiment 2 "up verbs" were excluded from accuracy computation).

A binomial logistic regression analysis ($correct\_answer \sim ET * valency * sem\_class$) for Experiment 1 yielded a significant effect of ET (binomial $p < 0.05$), a highly significant effect of valency and semantic class[3] (binomial $p < 0.001$), a significant interaction ET*valency and valency*sem_class (binomial $p < 0.05$) and a highly significant interaction ET*sem_class and ET*valency*sem_class (binomial $p < 0.001$). The same analysis for Experiment 2 yielded a highly significant effect of ET, valency and semantic class and semantic class with significant interactions ET*valency (binomial $p < 0.005$) and ET*sem_class (binomial $p < 0.001$). A binomial logistic regression analysis ($accuracy \sim ET$) for both Experiment 3 and 4 yielded a highly significant effect of ET on accuracy (binomial $p < 0.001$).

Certain ETs seem to be easier to identify than others. In particular, within the transitive VPs, ACCs seem easier than ACTs, probably due to their being more prototypically transitive in Italian and English, and ACHs and ACTs seem easier to identify when intransitive (as in Italian and English ACHs and ACTs are more prototypically intransitive).

Also, it seems that the semantic class of the predicate might play an important role in leading the annotators' choices in ET classification. Please note that a straightforward correspondence between ETs and semantic classes (e.g. motion verbs → ACT) was when possible avoided: a special effort was made when building the stimuli, in order to have, within each ET class, representatives of different semantic classes, and, within each semantic class, representatives of different ETs.

As to the 10 transitive VPs (2 ACC, 4 ACH, 4 ACT) which also appeared a second time with the particle "up" (see table 2), the contribution of the particle to the ET of the VPs strengthen their telicity, making it more prominent (for ACC and ACH items) or by changing the value of the RES feature (ACT answers go from 34 for the base version to only 15 for the "up" version).

	all	ACC	ACH	ACT	STA
Exp 1 (it, verbs)	0.63	0.76	0.66	0.61	0.48
Exp 1, transitives	0.59		0.57	0.53	
Exp 1, intransitives			0.76	0.69	
Exp 2 (en, verbs)	0.68	0.81	0.66	0.72	0.51
Exp 2, transitives	0.64		0.60	0.64	
Exp 2, intransitives	0.78		0.73	0.82	
Exp 3 (it, pictures)	0.42	0.34	0.54	0.60	0.34
Exp 4 (en, pictures)	0.54	0.68	0.54	0.50	0.48
MaxEnt	0.85	0.89	0.90	0.74	0.78
SOM	0.50	0.86	0.47	0.50	0.27

Table 3: Accuracy values

[2]Disagreement weights were arranged according to the number of features shared by the ET categories: a disagreement between ACH and ACC, which only differ for the feature of [+/−RES], is not as bad as the one between ACH and STA, which differ for three features ([+/−DUR], [+/−DYN], [+/−RES]).

[3]WordNet top-nodes were used as semantic class labels.

Disagreement with the gold standard: An item-wise analysis showed that, despite our effort to select non-polysemous stimuli (e.g. *passeggiare*, "to stroll", ACT; *montare un gioco*, "to assemble a toy", ACC), the items upon which the participants agreed the least with the gold standard actually allowed for multiple ET interpretations. Consider the following examples from Experiments 1 and 2:

- *formare una fila*, "to form a queue", potentially ACH/STA ambiguous;
- *scegliere il disco*, "to choose the recorder", arguably unspecified for [+/−DUR] (ACC/ACH);
- *conceive the theory*, arguably unspecified for [+/−DUR] (ACC/ACH);
- *tumble*, ACH reading or ACT (iterative) reading;

Some lexical differences emerged between Italian VPs and their English correlate:

- *impiegare i materiali*, "to use the materials" was classified as ACC ($[+RES]$) in Italian, but as ACT ($[-RES]$) in English;
- *precipitare*, "to tumble", was classified as ACH by our Italian participants, but its English correlate seemed to have a more durative (iterative) ACT reading.
- the picture for *to crawl*, was correctly classified as an ACT by English speakers, but 8 out of 20 Italian speakers gave a STA answer; interestingly enough, the speaker of a language lacking of a compact verb for *to crawl* as Italian have also selected a stative reading for the picture;
- *precipitare*, "to tumble", is classified as ACH by our Italian participants, but its English correlate seems to have a more durative (iterative) ACT reading.

Agreement and accuracy were lower for Experiments 3 and 4, which used picture stimuli: a picture offers a sample of reality from which only some parts can be selected. For example, consider the pictures for *to bounce* and *to salute*, both of which showed low accuracy values both for speakers of Italian and English (< 3):

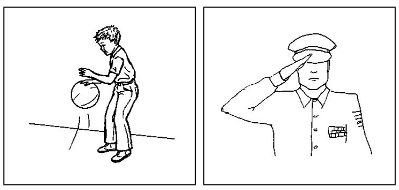


Figure 3: (*to bounce*), (*to salute*).

The picture for *to bounce* was originally labelled as ACH, but the participants interpreted it as an ACT (i.e. repeated acts of bouncing), the picture for *to stand* (STA) was interpreted as ACH (*to stand up*). Also in picture classification tasks ET classes seem far from being comparably evident to metalinguistic judgements.

feature set	distributional feature
adverbial	- temporal adverbs (e.g. in X time, for X time) - intentional adverbs (e.g. deliberately) - frequency adverbs (e.g. rarely, often) - iterative adverbs (e.g. X times)
morphological	- present tense - imperfect tense - future tense - simple past - perfect tenses - progressive periphrasis
syntactic and argument structure	- absence of arguments besides the subj. - presence of direct object, indirect obj. - presence of indirect obj. - presence of a locative argument - presence of a complement sentence - passive diatesis - number, animacy and definiteness of subj. and direct obj.

Table 4: Features

3 Corpus-based models of Event Types

Results from experiment 1 on speakers' metalinguistic judgements of ETs have been compared with the performance of computational models of ET classification trained with linguistically-motivated features extracted from Italian corpora: MaxEnt and SOM from Zarcone and Lenci (2008). MaxEnt is a supervised model which performs ET classification with Maximum Entropy classifiers, SOM is a self-organizing map which identifies ET clusters is an unsupervised way. See accuracy values in table 3[4].

3.1 Linguistically-motivated features

The features used to train the corpus-based models are very well-known in the linguistic literature for being (positively or negatively) correlated with particular event types (Dowty, 1979; Bertinetto, 1986; Pustejovsky, 1995; Rothstein, 2004). Extracted features include the following (see table 4 for a complete list):

adverbial features - they are among the main "event type" diagnostics in ET literature, but they are not very frequent in corpora data;

[4] Accuracy was higher for MaxEnt, but its coverage is limited to only 28 verbs; accuracy for SOM raised to 0.73 when lumping ACH and ACC into a single telic class.

morphological features - although actionality and aspect are independent categories, it is possible to observe typical correlations between some event types and specific aspectual values (Comrie, 1976). This set of features includes verb morphological tense-aspectual values;

syntactic and argument structure features - they include verb morphosyntactic, syntactic and semantic features of verb arguments, which are typically held responsible for ET shifts.

3.2 Corpus-based models vs. behavioral studies

A significant effect of ET on accuracy was yielded by a binomial logistic regression analysis for MaxEnt (binomial $p < 0.001$): significant differences in the pairwise comparisons clearly show a picture where [+RES] ETs (ACC and ACH) are easier to recognize than [−RES] ETs [5] - this seems to be the case also for Experiment 1.

No effect of ET on accuracy was found for SOM (binomial $p > 0.1$), but pairwise comparisons yielded a significant difference between ACC and STA accuracy ($z = -2.17; p < 0.05$). The distance between ACC and STA is comparable to the one found in Experiment 1 and 2: ACC are again the easiest to identify, STA the most difficult. SOM does not perform well on ACH, and this could be due to the sparseness of linguistic indicators for ACH (e.g. "in x time", punctual temporal indications).

Results from MaxEnt seem to mirror the ones from Experiment 1, showing that [+RES] classes (ACC and ACH) are more prominent and more easily identifiable. Such difference seems to be purely linguistic, since it does not show in Experiment 3. The convergence between the metalinguistic study and the computational models is coherent with the idea that the characterization of ET as "linguistic objects" is strongly related with their corpus distribution. Not only can distributional data capture semantic classes such as ETs, but it seems also to be the case that ET classes which have a clearer distributional characterizations are also easier for the speaker's to identify.

[5]Significancies for pairwise comparisons yielded by the binomial logistic regression analysis: ACC > ACT; $z = -6.69, p < 0.001$; ACC > STA; $z = -5.66, p < 0.001$; ACH > ACT, $z = -8, p < 0.001$; ACH > STA, $z = -6.96, p < 0.001$

Similar comparisons between Experiment 2 and computational models trained on English corpora are ongoing.

4 Future experiments

We presented above-chance results from behavioral studies and corpus-based models in event type classification with pictures and lexical items for English and Italian. Materials for the corpus studies and the behavioral studies presented here are not homogeneous: the stimuli for the behavioral experiments were first selected to match criteria for on-line psycholinguistic studies, whereas the corpus-based models were trained with highly frequent verb items, in order to limit the sparseness of the distributional vectors. There is ongoing work to train corpus-based models with a state-of-the-art dependency corpus of Italian (Baroni et al., 2004; Bosco et al., 2009) and to evaluate them using the same dataset of the behavioral experiments presented here. As in Zarcone and Lenci (2008), the contribution of each feature set (adverbial, morphological, syntactic) will be evaluated by running different experiments with different feature sets.

Another battery of experiments is planned to test metalinguistic judgements on small video clips, which promise to be a useful tool in the investigation of event representations, and to better convey features like DUR or RES which are not easily delivered by a picture stimulus.

It has been suggested (Embodied Cognition Framework, Haggard et al. (2007)) that semantic representations are not purely amodal, but rather grounded in our sensorimotor perception. Cross-modal and intra-linguistic differences can provide useful insights to better grasp the very nature of ETs, and to better understand to what extent they are a purely linguistic phenomenon or to what extent they provide us with schemes to interpret reality.

References

F. Antinucci and R. Miller. 1976. How children talk about what happened. *Journal of Child Language*, 3:167–189.

Marco Baroni, Silvia Bernardini, Federica Comastri, Lorenzo Piccioni, Alessandra Volpi, Guy Aston, and Marco Mazzoleni. 2004. Introducing the la Repubblica corpus: a large, annotated, TEI(XML)-compliant corpus of newspaper italian.

In Maria Teresa Lino, Maria Francisca Xavier, Ftima Ferreira, Rute Costa, and Raquel Silva, editors, *Proceedings of LREC 2004*, pages 1771–1774, Lisbona. ELDA.

E. Bates, K. Federmeier, D. Herron, and et al. 2000. Introducing the CRL International Picture-Naming Project (CRL-IPNP). *Center for Research in Language Newsletter*, 12.

P.M. Bertinetto. 1986. *Tempo, Aspetto e Azione nel verbo italiano. Il sistema dell'indicativo*. Accademia della Crusca, Firenze.

I. Bonnotte. 2008. The role of semantic features in verb processing. *Journal of Psycholinguistic Research*, 37:199–217.

Cristina Bosco, Simonetta Montemagni, Alessandro Mazzei, Vincenzo Lombardo, Felice DellOrletta, and Alessandro Lenci. 2009. Evalita '09 Parsing Task: comparing dependency parsers and treebanks. In *Proceedings of EVALITA 2009*, Reggio Emilia, Italy.

O. Bott. 2008. *The Processing of Events*. Ph.D. thesis, University of Tübingen. Ph.D. Dissertation.

Bernard Comrie. 1976. *Aspect: An Introduction to Verbal Aspect and Related Problems*. Cambridge University Press.

D. Dowty. 1979. *Word meaning and Montague Grammar. The semantics of verbs and times in generative semantics and in Montague's PTQ*. Reidel, Dordrecht.

C. Finocchiaro and G. Miceli. 2002. Verb actionality in aphasia: data from two aphasic subjects. *Folia linguistica*, 36:335–357.

S. Gennari and D. Poeppel. 2002. Events versus states: Empirical correlates of lexical classes. In *Proceedings of CogSci2002*.

P. Haggard, Y. Rossetti, and M. Kawato, editors. 2007. *Sensorimotor foundations of higher cognition*. Oxford University Press, Oxford, UK.

S. Im and J. Pustejovsky. 2010. Annotating Lexically Entailed Subevents for Textual Inference Tasks. In *Proceedings of FLAIRS 2010*.

K. Krippendorff. 2004. *Content Analysys: An introduction to its methodology (2nd ed.)*. Sage, Thousand Oaks, CA.

P. Li and Y. Shirai. 2000. *The Acquisition of Lexical and Grammatical Aspect*. Mouton, New York.

V. Lucchesi. 1971. Fra grammatica e vocabolario. studio sull'aspetto del verbo italiano. *Studi di grammatica italiana*, 1:179–270.

J. Pustejovsky. 1995. *The Generative Lexicon*. MIT Press, Cambridge MA.

S. Rothstein. 2004. *Structuring Events - A Study in the Semantics of Lexical Aspect*. Blackwell Publishing.

R. Snow, B. O'Connor, D. Jurafsky, and A.Y. Ng. 2008. Cheap and fast - but is it good? evaluating non-expert annotations for natural language tasks. In *Proceedings of EMNLP 2008*.

Z. Vendler. 1967. *Linguistics in Philosophy*. Cornell University Press, Ithaca, NY.

H. J. Verkuyl. 1972. *On the Compositional Nature of the Aspects*. D. Reidel Publishing Company, Dordrecht.

A. Zarcone and A. Lenci. 2008. Computational models of event type classification in context. In Andreas Witt, Felix Sasaki, Elke Teich, Nicoletta Calzolari, and Peter Wittenburg, editors, *Proceedings of LREC 2008*, Marrakesh. ELDA.

A. Zarcone and A. Lenci. 2010. Priming effects on event types classication: Effects of word and picture stimuli. Poster presented at CogSci 2010. Portland.

Basis for the annotation of EPEC-RolSem

Izaskun Aldezabal
IXA NLP Group
Basque Philology Department
University of the Basque Country
Donostia, Spain
izaskun.aldezabal@ehu.es

Abstract

In this paper I present the linguistic resources used when annotating the Reference Corpus for the Processing of Basque EPEC, in terms of semantic roles, argument structure and verb senses (EPEC-RolSem). When facing the annotation at any level, some crucial decisions have to be made, such as the model to be adopted and the criteria for the adaption of such model. Among other reasons, the fact of having the resources I am presenting here has led us to select the PropBank/Verbnet style model (Palmer *et al.*, 2005, Kipper 2005). Concretely, these resources are: the translation of all the verbs in Levin (1993) into Basque and an in-house database with syntactic/semantic subcategorization frames (ssf) for Basque verbs (EADB–Data Base for Basque Verbs), similar to the mentioned models. By means of the first resource and based on the Levin's class, we have linked the Basque verbs with the related PropBank/Verbnet verbs, getting all the corresponding information. On the other hand, the ssf-s of the EADB have been very useful to associate the appropriate English verb sense as well as to define an entry in the PropBank/Verbnet style for the Basque Verb.

1 Introduction

In the Ixa research group[1], the Reference Corpus for the Processing of Basque EPEC is being tagged at many linguistic levels, starting from morphosyntax to semantics, including some pragmatic features (Aduriz *et al.*, 2006). In each level, certain models and tagging manuals have to be developed for the annotators. In the case of semantics, the most difficult task is to establish criteria to define senses in a coherent and understandable way to facilitate the annotation process. Our previous work on semantics, when treating nouns, has mainly focused on Wordnet fine-grained senses (Felbaum, 1998), having as a result the Basque Wordnet (EusWordnet) (Pociello et al, forthcoming) and the Basque Semcor (EuSemcor) (Agirre et al., 2006a). Nevertheless, for the annotation of verbs, this fine-grained orientation has been questioned as some works point out (Ide and Véronis, 1998). This way, our data-base for Basque verbs (EADB) has been built with a more general perspective: although senses are defined for each specific verb, they are thought to be valid across verbs, based on Levin's (1993) methodology but mainly following Vázquez et al.'s (2000) alternation criteria. Consequently, more coarse-grained senses, similar to cognitive categories, are proposed for each verb entry. In addition, with the aim of defining alternations (either general or language specific ones), the syntactic realizations of the arguments in each sense are also taken into account (see section 2). With all this in mind, the PropBank/Verbnet style model (Palmer *et al.*, 2005, Kipper 2005) was thought to be a suitable one for adopting for Basque, as shown in Agirre et al.(2006b).

Other reasons have also persuaded us to choose the PropBank/Verbnet model:

1. The PropBank project starts from a syntactically annotated corpus, as we do.
2. Given the VerbNet lexicon and the annotations in PropBank, many implicit decisions according to problematic issues like argument/adjunct selection for distinguishing each verb senses are settled by examples, and

[1] ixa.si.ehu.es

seem therefore easier to replicate when we tag the Basque data. Moreover both (PropBank and Verbnet) resources are complementary as each one has appropriate and different linguistic information for defining verbs and for learning them automatically (Merlo and Van der Plas, 2009).

3. The PropBank model is being developed in other languages such as Chinese (Palmer and Xue, 2003), Spanish and Catalan (Civit et al., 2005a), Dutch (Monachesi et al, 2007), French (Van der Plas et al., 2010) and Russian (Civit et al., 2005b). Having corpora in different languages annotated following the same model makes it possible to carry out crosslingual studies, as it is demonstrated in Korhonen et al. (2010).
4. In the Verb Index[2], the information regarding PropBank and Verbnet is linked for many verbs. There is also information about other models such as Framenet (Baker et al, 1998), Wordnet (Fellbaum, 1998), Ontonotes (Hovy et al, 2006). This way, we could enrich Basque verbal models with the richer information currently available for English.

In this paper I present the main linguistic resources used for the predicate labelling of the EPEC corpus. In section 2, I explain the work carried out to translate the Levin's (1993) English verbs into Basque and I show the linking with the PropBank/Verbnet information. In section 3, I describe the syntactic-semantic frames (ssf) used to define verb entries in the EADB and the way of adopting the entry in the PropBank/Verbnet style. Finally, in section 4, the current situation and future work are outlined.

2 Translation of Levin's (1993) verbs into Basque and linking them to PropBank/Verbnet

Levin (1993) has been a reference to analyze verbs in other languages. She claims that the distinctive behavior of verb classes with respect to the diathesis alternations arises from their meaning: "once such a class is identified its members can be examined to isolate the meaning component they have in common. Thus, the diathesis alternations can be used to provide a probe into the elements entering into the lexical representation of word meaning" (Levin, 1993:14).

Many works have been carried out to compare the alternations she proposes for English with the ones existing in other languages (Jones et al., 1994; Taulé, 1995; Saint-Dizier 1995), Basque among them (Aldezabal, 1998, 2004). The studies carried out for Basque show that from the 80 Levin's alternations 24 are found in Basque. One reason for that is that some of the alternations in Levin are specific to a few English verbs. It has also to be pointed out that the dialectal variation was not considered, and some works reveal that, the dative alternation (or similar to it) appears in the dialect of the North part (Etxepare and Fernandez, 2011). However, those alternations that occur in specific dialects seem to be more a shade of the meaning of the sentence and they do not seem to be so useful for doing semantic classes (following Levin's methodology). Besides, in other languages also occur that many alternations do not exist. However, the alternations that are found in both languages are relevant enough for doing big classes, what we precisely do in our EADB (see section 3).

Anyway, when comparing the alternations only the verbs in the examples of the alternations were taken into account. All the verbs are found in the second part of Levin's book, where she describes the semantic classes resulting from the shared alternations. Therefore, in order to make a complete comparison, all of them were studied. The first task for that was to identify the equivalent verb in Basque, and then to ensure the differences and similarities both at alternation and class level.

2.1 Translation criteria and some examples

The translations in each class were done based on the *Morris* dictionary (Morris, 1998) and applying two general criteria:

- First of all, the meaning of the Levin's semantic class was considered.
- Then, the most syntactically similar equivalent(s) was/were selected.

In many cases, the verb in the class and the alternations involved are shared in both languages. For instance, most of the verbs in the 45.1 "*break* verbs" class with its prototypical causative/inchoative alternation can be translated without any difficulty; many of the verbs in the 9.1 "*Put* verbs" class do not either show any difficulty to be translated. Only we find the fact that for one English verb we can use more than one equivalent. For instance, the verb `break` in the 45.1 "*Break* Verbs" class is translated with the three synonyms `hautsi`, `puskatu`, `apurtu`,

[2] http://verbs.colorado.edu/verb-index/index.php

because they three mean the same (regarding the class meaning) and admit the causative/inchoative alternation without showing differences at this level. These cases are not problematic since each of the Basque verbs will be linked to the *break* verb when annotating the corpus.

However, as we went going down in the subclasses, some mismatches were found. These are illustrated in this section.

- Some of the syntactic properties are shared. For instance, the verbs `tell` and `say` which differ between them in the different behavior when admitting the dative alternation (`tell` admits (I tell sb sth / I tell sth to sb)) while `say` does not (*I say sb sth), are expressed with the same equivalent verb in Basque: `esan`. In Basque, there is not a different verb that reflects this syntactic variation, and the valence properties are the same as in the both English verbs. In those cases, both English verbs should be assigned to the Basque `esan` verb when annotating the corpus.

- A single word is used in English while two are necessary in Basque. These are mostly those verbs that have a manner or an instrument meaning incorporated, such us many of verbs into the "*funnel* verbs", "*wipe* verbs", "*spray/load* verbs", "*drive* verbs", "*poison* verbs", and "verbs of instrument of communication", among others. E.g.: `ladle`: `burduntzaliaz bota`, literally meaning 'to throw with a ladle'.

In these cases it may happen that the concept expressed by the verb that lexicalizes the manner in English is not a lexicalized concept in Basque. The example above represents that case. Therefore, these verbs should have to be considered and analyzed into the single verb (`bota throw`) class (17.1), where manner is going to be a possible adjunct. These cases are more difficult to solve when annotating: the annotator should realize that the non lexicalized adjunct + verb expression in Basque has to be annotated with the appropriate single word in English.

- The same verb is used in Basque for different verbs in different classes, but the object candidates must be specified as in English. For instance, Verbs in the 13.4.2 "*Equip* verbs" class in Basque are translated with the same verb as in 13.1 "give" or 13.2 "contribute" verb classes (for example, `charge` (13.4.2): `ardura eman` -> lit. 'to give (13.1) the charge'). That is, in Basque such distinction does not exist ('to charge somebody with a task' but not *'to give somebody with a task'). However, in order to provide the equivalent for `charge`, the object (`zeregin baten ardura`: 'charge of a task') must be equally specified in Basque. In these cases, when annotating the corpus the verb *charge* should be used when in the Basque sentence appear "*give the charge (of a task)*" and it should also be considered a multiword.

2.2 Linking to PropBank/Verbnet

Taking into account all these phenomena, we are able to say *a priory* that when linking the Propbank/Verbnet equivalent to the Basque verb, the argument structures (at least at valence level) of English and Basque verbs are not going to be the same in some cases, and, as a consequence, neither the alternations involved on them.

Moreover, when the concepts are not lexicalized in Basque, there will be an element that will be appearing as an apart adjunct (and not as an argument) in Basque, while in English there will not be a syntactic counterpart (but it will be incorporated in the verb).

In any case, the information obtained from the linking regarding the sense and rolesets will be very helpful in the process of building the Basque verb entry with the PropBank/Verbnet scheme (although classes are not shared). In table 1 a list of some of the verbs after the linking based on Levin's classes is shown.[3]

glue	22.4	*erantsi, kolatu*
go	47.7	*joan*
go	51.1	*joan*
gobble	38	*glu-glu egin*
gobble	39.3	*irentsi*

Table 1: the link between the PropBank and Basque verbs based on Levin's (1993) class.

3 The EADB: data-base of syntactic-semantic frames (ssf) for the Basque verbs

Following the methodology that Levin suggests in her work, the crucial task is to detect those alternations that are semantically sensitive and then find the semantic components that would be in the lexical representation of the verbs.

For this task, and taking a revised point of view of the alternation concept which is also

[3] It has to be noted that the Levin's classes have been revised in PropBank/Verbnet. Consequently some verbs remained without any link (Aldezabal et al, 2010).

adopted in other works (Vázquez et al., 2000; Rebolledo, 2002), I studied 100 Basque verbs basing on real corpus examples (Aldezabal, 2004).

I concluded that each verb has one or more prototypical frames to express any of the general semantic values appearing when analyzing verbs in general. These semantic values are not senses in the way they appear in the dictionaries, but basic cognitive categories or general predicate types which can serve to propose big classes at semantic level (such as *change of state*, *change of position*, *activity of an entity*, *creation of an entity*, *assignment of an attribute*, *exchange of an entity*, *situation of an entity* and so on). This semantic information is expressed by general semantic roles (or semantic components) coherently combined (that is, for a verb to express the general predicate *change of state*, at least an *affected theme* must be contain; or for a verb to express the general predicate *creation of an entity*, a *created theme* must be contain, and so on)[4]. This way, some verbs share the capability to represent the same general predicate. However, it does not mean neither they should have the same syntactic frames (although it happens in many cases), nor they share the same alternations (although it also happens in many cases).

Based on that assumption, I defined a number of syntactic-semantic frames (ssf) for each verb. Each ssf is formed by semantic roles and the declension case that syntactically realizes this role. The ssf-s that have the same semantic roles define a verbal coarse-grained sense and are considered syntactic variants of an alternation. Different sets of semantic roles reflect different senses. This is similar to the PropBank model, where each of the syntactic variants (similar to a frame) pertains to a verbal sense (similar to a roleset).

In Table 2 we can see an example of the ssf-s for the verb `esan`. It has two senses and the first one contains two syntactic variants. The first sense can be translated as 'tell/say' as in Levin's 37 "Verbs of communication" class, and the second sense as 'call', as in Levin's 29 "Verbs with Predicative Complements" class.

`esan-1` (= 'tell/say'): Activity (communication) of an entity. Two arguments in two syntactic variants:
`esan-1.1`: arg1_ERG[5], arg2_ABS
`esan-1.2`: arg1_ERG, arg2_COMP
`esan-2` (= 'call'): Assignment of an attribute. Three arguments in a single syntactic realization:
`esan-2`: arg1_ERG, arg2_ABS, arg3_DAT

Table 2. Syntactic-semantic frames for the verb `esan` (='tell/say/call') as provided by the EADB lexicon.

These ssf-s together with the information obtained from the link to PropBank/Verbnet are a robust basis to define the new lexical entry with the PropBank/Verbnet scheme and to go on tagging the EPEC corpus in such framework.

Table 3 shows the adopted PropBank/Verbnet entry for the verb *esan*.

Basque verb: esan	
say.01/tell.01	***call.01***
Arg0: Agent *(ERG)* Arg1: Topic *(ABS/COMP)* Arg2: Recipient *(DAT)* Arg3: Attributive *(INS[6]/-i buruz/ -i erreferentzia eginez[7]/...)*	Arg0: Agent *(ERG)* Arg1: Theme *(DAT)* Arg2: Predicate *(ABS)*

Table 3: The PropBank/Verbnet style entry for the verb *esan*.

4 Current situation and future lines

We have already annotated a sample of sentences for each of the 100 verbs including in the EADB. During the tagging process some adjustments had to be made, because of differences both at multiword level and at valence level. For instance, in some verbs of motion an *extend* argument is defined for the English verb while in Basque it does not exist.

Besides, the annotation has been evaluated and one of the most significant conclusions has been that before annotating, taggers must clearly understand the entries that have been adapted to the PropBank/Verbnet model. In addition, it must be also taken into account that multiword expressions are problematic and that it is necessary to decide what to do with those cases. Moreover, in order to avoid confusions with modifiers, it is important to provide some information or guidelines, although we know that some things will remain unsolved since they are subjective.

At present, we are planning to automatize the annotation-process taking into account the lexi-

[4] I propose 13 general predicates and 21 semantic roles in total.

[5] ERG: ergative declension case; ABS: absolutive declension case; COMP: completive clause; DAT: dative declension case.

[6] The instrumental declension case

[7] These are complex prepositions meaning 'regarding', 'with respect to'…

con resulting from the annotated corpus. As a first step, we will detect the univocal role_case pairs, and then we will automatically annotate the occurrences of the corpus, including its corresponding verb sense. For the automatic annotation of new verbs, class based cross studies will be carried out.

References

Aduriz I., Aranzabe M.J., Arriola J. M., Atutxa A., Díaz de Ilarraza A., Ezeiza N., Gojenola K., Oronoz M., Soroa A., Urizar R. 2006. Methodology and steps towards the construction of EPEC, a corpus of written Basque tagged at morphological and syntactic levels for the automatic processing. *Corpus Linguistics Around the World*. Book series: Language and Computers. Vol 56 (pag 1- 15). Andrew Wilson, Paul Rayson, and Dawn Archer. Rodopi (eds.). Netherlands.

Agirre E., Aldezabal I., Etxeberria J., Iruskieta M., Izagirre E., Mendizabal K., Pociello E. 2006. (Agirre et al 2006a). A methodology for the joint development of the Basque WordNet and Semcor. *Proceedings of the 5th International Conference on Language Resources and Evaluations (LREC)*. Genoa (Italy).

Agirre E., Aldezabal I., Etxeberria J., Pociello E. 2006. (Agirre et al 2006b). A Preliminary Study for Building the Basque PropBank. *Proceedings of the 5th International Conference on Language Resources and Evaluations (LREC)*. Genoa (Italy).

Aldezabal I., Aranzabe M., Díaz de Ilarraza A., Estarrona A. 2010. Building the Basque PropBank. Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odjik, Stelios Piperidis, Mike Rosner and Daniel Tapias (eds.). *Proceedings of the Seventh Conference on International Language Resources and Evaluation* (LREC 2010). European Language Resources Association (ELRA). Valletta.

Aldezabal I. 2004. *Aditz-azpikategorizazioaren azterketa. 100 aditzen azterketa zehatza, Levin (1993) oinarri harturik eta metodo automatikoak baliatuz*. PhD thesis. Basque Philology Department Leioa. UPV-EHU.

Aldezabal I. 1998. Levin's verb classes and Basque. A comparative approach. Colloquium series of the Department of Computer Science at UMIACS (University of Maryland Institute for Advanced Computer Studies).

Baker, C.F., Fillmore, C.J., Lowe, J.B. (1998). The Berkeley FrameNet project. In *Proceedings of the COLING-ACL*. Montreal, Canada.

Etxepare, R. and Fernandez B. (to appear in 2011). Microparameters in Dative Constructions. Oxford Studies in Comparative Syntax. Oxford University Press. Under contract.

Fellbaum C. 1998. *Wordnet, an electronic lexical database*. MIT Press.

Hovy E., Marcus M., Palmer M., Ramshaw L. and Weischedel R. 2006. OntoNotes: The 90% Solution. *Proceedings of HLT/NAACL*. New York.

Ide N. and Véronis J. 1998. "Word Sense Disambiguation: State of the art". *Computational Linguistics*, 1998, 24 (1).

Jones D., Berwick R., Cho F., Khan Z., Kohl K., Nomura N, Radhakrishnan A., Sauerland U. & Ulicny B. 1994. *Verb Classes and Alternations in Bangla, German, English, and Korean*. Massachussetts Institute of Technology center for Biological and Computational Learning and the Artificial Intelligence Laboratory.

Kipper, K. 2005. Verbnet: *A broad-coverage, comprehensive verb lexicon*. Ph.D. thesis, University of Pennsylvnia.

Levin B. 1993. *English Verb Classes and Alternations. A preliminary Investigation*. Chicago and London. The University of Chicago Press.

Korhonen A., Sun L, Poibeau T. and Messiant C. 2010. Investigating the cross-linguistic potential of Verbnet-style classification. *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010).* Beijing.

Merlo P. and Van der Plas L. 2009. Abstraction and Generalization in Semantic Role Labels: PropBank, Verbnet or both? *Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP*. Suntec, Singapore.

Monachesi G., Stevens G., Trapman J. 2007. adding semantic role annotation to a corpus of written Dutch. *Proceedings of the Lingusitcs Annotation Workshop (LAW)*. Prague, Czech repuplic.

Morris M.. 1998. Morris Hiztegia.

Palmer M., Gildea D., Kingsbury P. 2005. The Proposition Bank: A Corpus Annotated with Semantic Roles. In *Computational Linguistics Journal*. 31:1.

Pociello E., Agirre E. and Aldezabal I. (to appear in Language Resources and Evaluation (LRE)). The Basque Lexical Knowledge Base: the Basque Wordnet.

Rebolledo M. 2002. *Estructura sintáctica y significado verbal*. Dissertation. University of Santiago de Compostela.

Saint-Dizier P. 1995. A semantic classification of French verbs based on B. Levin's approach. Research report. IRIT.

Taulé M. 1995. Representación verbal en una Base de Conocimiento Léxico. PhD. Barcelona

Van der Plas L., Samardzic T. and Merlo P. Cross-lingual Validity of PropBank in the Manual Annotation of French. *Proceedings of the Fouth Lingusitic Annotation Workshop, ACL 2010*. Uppsala, Sweden.

Vázquez G., Fernández A., Martí M. A. 2000. *Clasificación verbal. Alternancias de diátesis*. Quaderns de Sintagma 3. Edicions de la Universitat de Lleida.

Verb Classes as Evaluativity Functor Classes

Pranav Anand and **Kevin Reschke**
University of California, Santa Cruz
Santa Cruz, CA, U.S.A.
{panand, kreschke}@ucsc.edu

1 Evaluation and Composition

In the past decade, the semantics of evaluative language – a hallmark of the generative semantics debates (see Fillmore (1985)) – has received renewed attention in both formal and computational linguistics, though in different guises: appraisal (Martin and White, 2005), expressive content (Potts, 2005), sentiment (Pang and Lee, 2008), and valuation (Jackendoff, 2007). This work has largely focused on evaluativity at the lexical level, i.e., the evaluative stance encoded in a particular word, or holistic judgments of phrasal level stance (e.g.,what is the author's sentiment in this sentence?). Largely absent from these discussions is the role of compositional interpretation in connecting lexical evaluativity to phrasal level stance. When compositionality has been investigated, it is in the explication of polarity preservers/inverters (e.g., copulas/negation) or the attempt to extend such operators to the event domain (Moilanen and Pulman (2007), Nasukawa and Yi (2003)).

While preservation or inversion may be valid for logical operators, we argue that evaluations of events are more complex. As the determination of truth conditions for event predicates relies on a knowledge of the event participants, so too evaluative stance towards an event is a product of the evaluative stances an assessor bears towards the predicate's participants. We thus propose that, for the purpose of determining evaluative stance towards events, verbs and other predicates of events should be analyzed as functors (or, mappings) from n-ary evaluative tuples to an evaluative value. The task of determining event evaluativity thus reduces to the problem of determining the particular mapping from n-ary tuples to evaluative values that a given verb induces. This problem is complex (and potentially context-dependent), but we claim that verbal predicates fall into evaluative functor *classes* based on their lexical entailments about their arguments. In particular, we identify three entailment types as prominent predictors of functor class: POSSESSION by a participant, EXISTENCE of a participant, and AFFECTEDNESS of a participant. We justify these claims by human annotation of a generated corpus as well as coding of naturally occurring text in the Gigaword corpus (Parker et al., 2009).

The paper is organized as follows. In section 2, we outline the theoretical machinery behind treating event-level evaluativity in terms of n-ary functors. Section 3 discusses how verbal predicates show commonalities based on the three entailment types above; it additionally demonstrates that the evaluative stance towards verbs of change of state is determinable in terms of the result state alone. Finally, section 4 presents our empirical assessment of the predictions of section 3. Section 5 concludes with a discussion about how to extend the machinery in section 2 to deal with more subtle differences in evaluative stance.

2 Arity of Evaluativity

Wilson (1975) discusses a contrast in *verbs of withholding* (1); this is the product of contradictory presuppositions regarding the desires of the protagonist (Gazdar, 1979). Events of *deprivation* differ from those of *sparing* in terms of whether the withholdee desired the outcome in question:

(1) a. She deprived him of a day at the seaside.
b. She spared him a day at the seaside.

But there is a further evaluative component distinguishing (1a) and (1b) – in a sense, (1b) is, if not infelicitous, somehow pragmatically marked, given the intuition that speakers have a tendency to describe events as instances of sparing insofar as they are positive outcomes from the speakers' perspectives.[1] And, indeed, in the human-annotated MPQA Subjectivity Lexicon (Wilson et al., 2005) *deprive* and *spare* are specified as strongly negative and positive, respectively. In contrast, the semi-supervised SentiWordnet lexicon (Esuli and Sebastiani, 2006) marks both terms as neutral ("objective"). This surprising disagreement, we argue, is the result of the fact that sparing/deprivation events are not *en masse* positive or negative; nor are they non-evaluative. Rather, particular events of sparing or deprivation are modulated by the author's stances towards the event participants, as with the protagonists in (2). The deprivation of someone one feels positively towards (e.g., an ally) is typically evaluated as negative, and vice versa towards those one feels negatively associated with (e.g., an enemy):

(2) a. My {ally, enemy} was deprived shelter.

b. My {ally, enemy} was spared a dangerous mission.

A similar effect is noted by Nasukawa and Yi (2003) for predicates such as *have/lack* and Moilanen and Pulman (2007) for *fail*. The approaches in these papers consider only the direct object of a verb in determining overall evaluation, meaning that *lacking* and *failing* events are uniformly negative. Empirically, however, evaluation of an event of *lacking shelter* depends on the perceptions of the grammatical subject in an assessor's mind: an enemy's lack of shelter is arguably perceived as better than an ally's. Thus, compositional computation of the evaluation of an event requires the evaluations of all arguments, not merely the internal ones. More precisely, if $\mathcal{E}$ is the domain for evaluativity, an n-ary verb V induces an n-ary evaluativity functor:

(3) $E_V : \mathcal{E}^n \rightarrow \mathcal{E}$

Assuming that $\mathcal{E} = \{-, +\}$, the contrasts between withholding and possession verbs may be represented by the functions given in Table 1. As noted above, E_{have} and E_{lack} are opposites (equivalent to XNOR and XOR, respectively). However, $E_{deprive}$ and E_{spare} not. Rather, they are each partial functions of $E_{withhold}$, representing

x	y	E_{hve}	E_{lck}	E_{wthld}	E_{dprv}	E_{spr}
+	+	+	-	-	-	#
+	-	-	+	+	#	+
-	+	-	+	+	+	#
-	-	+	-	-	#	-
x have/lack *y*						
a withhold/deprive/spare *x* of *y*						

Table 1: Functors for verbs of possession and withholding.

the fact that *deprive* is an infelicitous description when event in question is assessed positively and *spare* infelicitous when the event is assessed negatively. Note additionally, that the agent for verbs of withholding does not affect event-level evaluativity (and thus $E_{withhold}$ and E_{lack} produce identical outcomes for the same x and y despite an arity difference), as one can determine the event-level evaluativity even in the absence of the assessor's stance towards the agent. This is not a universal characteristic of verbal functors. For example, E_{visit}, $E_{entertain}$, and E_{meet} are all sensitive to agent evaluation: the identity of the one who visits, entertains, or meets is just as crucial as the identity of the one visited, entertained, or met for computing how one would feel about the event.

3 Evaluativity and Entailment Class

We propose that the contrast in sensitivity towards agentivity between $E_{withhold}$ and E_{visit} is not accidental. Rather, it, as well as the general character of a verb's evaluativity may be understood on the basis of lexical entailments of the verb. Verbs of withholding, for example, are result verbs (Levin and Rappaport, 1995), entailing the existence of an unspecified causal event as well as a consequent lack of possession. Correspondingly, when felicitous, they are identical to E_{lack}. We propose that this insensitivity to the agent follows from a more general principle on changes of state (excepting AFFECTEDNESS, below):

(4) CHANGE IRRELEVANCE: The evaluation of a change of state is equivalent to the evaluation of the end state.

CHANGE IRRELEVANCE thus correctly predicts that the evaluativity of verbs of profit and loss (e.g., *profit from*) are insensitive to the causing event.

[1]The preferences of withholdee give rise to an additional, distinct evaluative stance the withholdee bears towards the event of withholding. The derivation of such event participant evaluativity is beyond the scope of this paper, though it too is expressible via the general machinery we introduce.

x	E_{xst}	E_{nxst}	E_{pstv}	E_{ngtv}
+	+	-	+	-
-	-	+	-	+

x is existing/non-existent
x has a positive/negative property

Table 2: Functors for the basic entailment states of EXISTENCE and AFFECTEDNESS.

A direct consequence of CHANGE IRRELEVANCE is that determining the evaluation of a complex event may be reduced to determining the evaluation of a state. While the factors affecting evaluation of a state are complex, we propose the following list of characteristic entailments:

(5) a. EXISTENCE: A participant has/lacks existence.
b. POSSESSION: One participant has/lacks possession of another.
c. AFFECTEDNESS: A participant has a positive/negative property.

Thus, the evaluativity functor of an event whose result state is one involving POSSESSION will either be identical to E_{have} or its negation (E_{lack}). A similar situation holds for events resulting in states involving EXISTENCE and AFFECTEDNESS, whose corresponding functors are provided in Table 2. In particular, we predict that the change of state verb classes of creation/destruction, gain/lose possession, and benefit/injury will correspond to the basic functors in Tables 1 & 2 as follows: a) creation events result in states involving EXISTENCE, and hence their evaluation will be identical to $E_{existing}$; b) gain events result in states involving POSSESSION, and hence their functors will be determined by E_{have}; finally, benefit events result in a positive property, and hence will follow $E_{positive}$. The antonymic classes of each of these will have functors determined by $E_{non\text{-}existent}$, E_{lack}, and $E_{negative}$, respectively. The predictions are summarized in Table 3.

Although there are class-based commonalities, it is important to acknowledge lexical idiosyncrasies that are directly evaluative. In (6), where *abuse* seems infelicitous regardless of the injury to a negatively evaluated individual, unlike *injure*. This contrast reflects the negative evaluation lexically encoded in *abuse*, which conflicts with the strongly positive evaluation of *defenseless child* – abuse events are inherently bad, and it is difficult to think of defenseless children committing them.

(6) The defenseless child {#abused, injured, ?tortured} the monster.

VERB CLASS	RESULT STATE	FUNCTOR
creation	existence	$E_{existing}$
destruction	existence	$E_{non\text{-}existent}$
gain	possession	E_{have}
loss	possession	E_{lack}
benefit	affectedness	$E_{positive}$
injury	affectedness	$E_{negative}$

Table 3: Predicted functors for 6 change of state verb classes.

We represent verbs like *abuse* as constant functors over their argument tuples. In large part because they are constant (and hence, most expressible in terms of simple lexical inferences), these functors have received the greatest attention in both the sentiment and expressive content literature. While that attention is deserving, the claim here is that they are part of a larger combinatoric system which may be obscured by focusing only on limiting cases.

4 Empirical Assessment

We conducted two empirical assessments of the predictions summarized in Table 3: an annotation study on constructed sentences and a corpus investigation of the arguments correlating with overt markers of evaluativity.

4.1 Annotation Study

In order to test the claims of the theory directly while controlling for variables of interest, we conducted an annotation study on 6480 constructed sentences. The sentences combined 48 predicates from the injury, benefit, destruction, creation, and transfer classes in FrameNet (Ruppenhofer et al., 2005) with 18 participants bearing canonical positive, neutral, and negative evaluations. Example sentences are provided in (7), and the complete list of verbs and nominals used are listed in Tables 5 and 6, respectively.

(7) a. The {hero, man, villain} {comforted, assaulted} the {child, monster}.
b. The {hero, man, villain} {assembled, defaced} the {cathedral, building, torture chamber}.

Twelve annotators were instructed to indicate overall evaluative stance (positive, negative, neutral) the author had towards the event described in each sentence. Each annotator had 10% overlap with each of two other annotators (i.e., each annotated 486 sentences uniquely, 54 with one annotator, and 54 with another).

An ordinal logistic regression was fit to the data, using the participant (agent, object, goal) polarities and verb classes, as well as random intercepts for annotator and verb. Significant interactions are as follows: Both positive affected arguments showed significant interaction with the injury/destruction classes ($p < 2.22 \times 10^{-16}$); inter-annotator agreement replicated this (Cohen's $\kappa = 0.92$); that is, killing was judged more positive when the entity losing existence was an enemy and judged more negative when it was an ally. Neutral affected arguments showed behavior similar to positive arguments ($p < 2.22 \times 10^{-16}$), suggesting a principal of charity with respect to events of harm. Transfer verbs showed a weak sensitivity to possession. Table 4 illustrates the correlation between object evaluativity and annotated event evaluativity for dyadic verbs; the cells predicted by Table 3 are in bold. While gaining possession for positive possessors (e.g., a hero gaining a valuable watch) tracked the value of the object ($p < 1.34 \times 10^{-2}$, $\kappa = 0.89$), negative possessors showed less inter-annotator agreement ($p < 0.09$, $\kappa = 0.68$), with a tendency for neutral evaluation.

We also compared four putative constant negative functors: E_{abuse}, $E_{assault}$, E_{murder}, and E_{rape}. E_{murder} showed no difference from E_{kill}, but the remainder showed a strong negative evaluation in all +agent or +patient cases ($\kappa = 0.92$). However, in cases with a negative patient and a non-positive agent, there was less consensus ($\kappa = 0.68$), with a slight preference for a positive valuation. We suggest that the additional inference in these constant functors is negativity toward the agent. When the agent is otherwise positive, the negative response expresses disappointment in the agent, but not the outcome. When the agent's behavior is deemed less relevant (e.g., cases where the agent is otherwise negatively judged), the basic characteristics of the injury class reappear.

4.2 Corpus Evaluation

To assess whether the predictions in Table 3 are attested in naturally occurring text, we searched the one billion word Gigaword corpus (Parker et al., 2009) for ten target verbs. Event-level evaluativity was approximated by considering predicates extracted from three sentence frames which indicate stance directly: emotive factives (*X was happy/sad that* ϕ), polar adverbs (*Thankfully/Unfortunately,* ϕ), and promises/threats (*X promised/threatened to* ϕ). The target predicates were verbs of benefit (*help*, *cure*, *protect*, *reward*), injury (*kill*, *murder*, *execute*, *assault*, *injure*), and destruction (*destroy*). This procedure yielded approximately 6200 matches. Of these, we selected a sample ($n = 690$) for manual inspection.

OBJ	EVENT	BEN/CREAT	INJ/DESTR
-	-	**449**	125
-	n	128	74
-	+	73	**449**
n	-	17	563
n	n	**161**	**68**
n	+	470	19
+	-	28	**588**
+	n	105	16
+	+	**518**	44

Table 4: Dyadic Verb Annotations

In order to test the accuracy of our theory, we examined matches for the evaluativity of the verb object. According to the predictions of Table 3, object evaluativity should be deducible from the verb class and event-level evaluativity; thus, if the match in question involves a verb of destruction embedded in a positive frame (e.g., *Thankfully, they destroyed* ...), we predict that the object should be perceived as negative by the author. Table 7 summarizes our judgments on object evaluativity (positive, or +obj and negative, or -obj) across the three verb classes and evaluative contexts (positive, or +c and negative, or -c); the cells predicted by Table 3 are in bold. Example positive, negative, and neutral objects are given in Table 8. Objects with no obvious polarity and non-entity objects (e.g. *nobody*) were marked 'other.'

In general, the predictions of the theory are attested. The evaluativity of a benefit clause tends to match that the participant, while evaluativity of an injury clause tends to oppose it.[2] Likewise,

[2] The injury/+obj/+c counts are inflated. Half arose from context polarity misclassification; in remaining cases, context polarity was positive because it was contrastive – e.g., *Fortunately, the official was only injured in his hand.*

INJURY	abused, arrested, assaulted, beat up, executed, injured, insulted, killed, murdered, raped, scratched, undermined
BENEFIT	aided, comforted, cured, educated, helped, pardoned, protected, resuscitated, rewarded, strengthened, supported
DESTRUCTION	broke apart, crumbled, defaced, demolished, desecrated, destroyed, eliminated, fractured, obliterate, pulverized, shattered, wrecked
CREATION	assembled, brought forth, created, fabricated, faked, fashioned, generated, made, pieced together, produced, sculpted, synthesized
TRANSFER	accepted, acquired, bought, delivered, donated, gave, passed on, procured, sold, stole, surrendered, took

Table 5: Annotation Study Verbs

ANIMACY	POSITIVE	NEUTRAL	NEGATIVE
animate	child, hero, friend	co-worker, man, middle-aged individual	enemy, monster, villain
inanimate	cathedral, lovely silks, universal prosperity	building, cloth, general silence	torture-chamber, disgusting rags, universal poverty

Table 6: Annotation Study Nominals

	benefit		injury		destruction	
	+c	–c	+c	–c	+c	–c
+obj.	**49**	0	18	**216**	2	**20**
–obj.	1	**8**	**15**	2	**39**	1
other	27	10	150	115	7	10

+c: *thankfully; happy; promised*
+c: *unfortunately; sad; threatened*

Table 7: Token frequencies by verb-class, object, & context in Gigaword sample ($n = 690$).

obj.	examples
+	(indexicals) *us*; (concepts) *free speech*; (allies) *French partisans*; (valued objects) *treasures*; (esteemed individuals) *champion*.
–	(concepts) *incompetence*; (enemies) *bin Laden*; (illegal possessions) *field of poppy*; (weapons) *Iraqi Scuds*.

Table 8: Example judgments of evaluativity in verb objects

the evaluativity of a destruction clause generally opposes the polarity of the the participant's existence. For example, three-fourths of –obj/+c cases were instances of disarmament. The existence of weapons stockpiles is negative, thus their destruction (which results in loss of existence) is positive.

5 Further Directions

We have argued that event-level evaluativity should be considered in terms of verbal functors, and have shown that this allows us to capture entailment class based generalizations that appear distributionally valid. Given the sensitivities of the transfer annotations to not only polarity of evaluativity towards objects, but also the domain of evaluation, it would be wise to model the multiple dimensions of valuation that may be relevant within a judgment. Thus, the ontologies in both Jackendoff (2007) and Martin and White (2005) distinguish between ethical, teleological, and aesthetic evaluations. Hence, a gain of something aesthetically valuable may not give rise to the same evaluative intuition as something teleologically valuable. While it is straightforward to implement such distinctions via a multidimensional evaluative domain (i.e., $\mathcal{E} \cong \{-, 0, +\}^n$), the important and difficult work will be to determine how mappings across these finer-grained domains correlate with the semantic properties of the verbs they correspond to.

References

Andrea Esuli and Fabrizio Sebastiani. 2006. Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of LREC-06*.

Charles Fillmore. 1985. Frames and the semantics of understanding. *Quaderni di Semantica*, 6(2):222–254.

G. Gazdar. 1979. *Pragmatics: Implicature, Presupposition and Logical Form*. Academic Press.

Ray Jackendoff. 2007. *Language, consciousness, culture*. MIT Press.

B. Levin and M. Rappaport. 1995. *Unaccusativity*. MIT Press.

J. R. Martin and P. R. R. White. 2005. *Language of Evaluation: Appraisal in English*. Palgrave Macmillan.

Karo Moilanen and Stephen Pulman. 2007. Sentiment composition. In *Proceedings of RANLP 2007*.

T. Nasukawa and J. Yi. 2003. Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd international conference on Knowledge capture*.

Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135.

Robert Parker, David Graff, Junbo Kong, Ke Chen, and Kazuaki Maeda. 2009. English gigagword, 4th ed. Technical report, The LDC Corpus Catalog.

Chris Potts. 2005. *The Logic of Conventional Implicature*. Oxford University Press.

Josef Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, and Christopher R. Johnson. 2005. Framenet ii: Extended theory and practice. Technical report, ICSI Technical Report.

T. Wilson, J. Weibe, and P. Hoffman. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of EMNLP-05*.

D. Wilson. 1975. *Presupposition and Non-Truth Conditional Semantics*. Academic Press.

A Maximum Entropy Approach To Disambiguating VerbNet Classes

Lin Chen
University of Illinois at Chicago
Chicago, IL, USA
lin@chenlin.net

Barbara Di Eugenio
University of Illinois at Chicago
Chicago, IL, USA
bdieugen@uic.edu

Abstract

This paper focuses on verb sense disambiguation cast as inferring the VerbNet class to which a verb belongs. To train three different supervised learning models –Maximum Entropy (MaxEnt), Naive Bayes and Decision Tree– we used lexical, co-occurrence and typed-dependency features. For each model, we built three classifiers: one single classifier for all verbs, one single classifier for polysemous verbs only, and an ensemble of classifiers, one per each polysemous verb. Among those algorithms, Naive Bayes performs surprisingly badly. In general, MaxEnt models perform better, but Decision Trees models are competitive. Our best results are obtained with classifier ensembles.

1 Introduction

Our research group has long been involved in research on the interpretation and generation of instructional texts. Not only do we believe that verbs provide a crucial component of the semantics of such texts; we also have shown that verb-based semantics helps achieve more accurate discourse parsing (Subba and Di Eugenio, 2009). For our work on discourse parsing, we developed a new resource, the HomeRepair corpus, which contains 176 documents for a total of 53,250 words. It was manually annotated with rhetorical relations and *quasi-automatically* annotated with semantics. It was parsed with LCFLEX (Rosé and Lavie, 2000), which we integrated with VerbNet (Kipper et al., 2008) and with CoreLex, a noun lexicon (Buitelaar, 1998) (see (Subba et al., 2006) for details).

VerbNet (VN) is currently the largest English verb semantics resource. In VN, verbs are grouped in classes and subclasses. Each VN class is completely described by thematic roles, selectional restrictions on the arguments, and frames consisting of a syntactic description and semantic predicates – see the class *remove-10.1* in Figure 1. Our parser was integrated with VerbNet 2.1, which covered 3445 different verbs, for a total of 4656 verb senses, grouped in 191 first level classes. [1]

The *quasi*-automatic quality of the semantic annotation of our corpus is due to manual disambiguation of the correct interpretation among several LCFLEX may return. Some alternative interpretations are due to syntactic ambiguities, but others, to lack of verb sense disambiguation. For example, in the sentence *you may have to cut some tiles*, *cut* is mapped to two distinct VN classes, *BUILD-26.1* and the correct *CUT-21.1*.

Our work builds on much previous work on verb sense disambiguation. Verb sense disambiguation is a subtask within word sense disambiguation, but we do not have room here to review that vast literature. As concerns verb sense disambiguation, a first distinction concerns what counts as a verb sense: some of the work, e.g. (Dang and Palmer, 2005; Dligach and Palmer, 2008; Banerjee and Pedersen, 2010), focuses on verb senses variously derived from WordNet senses, not on VN class disambiguation. Other work, e.g. (Lapata and Brew, 2004), uses Levin's verb class definitions, which in turn are the foundations of VerbNet class definitions, but result in a different classification problem. If we now turn to VN class disambiguation, distinctions in approaches concern the specific models used, the features those models are built from, and / or the corpora that are employed. Previous work on VN class disambiguation (Girju et al., 2005; Abend et al., 2008) has focused almost exclusively on standard corpora such as PropBank; more importantly, it has used no relational information between a verb and its arguments, whereas we use typed dependencies here.

[1] VerbNet 3.1, the latest version, contains 3769 different verbs, for a total of 5257 verb senses, grouped in 274 classes.

CLASS: remove-10.1
PARENT: -
MEMBERS: abstract, cull, delete, disgorge, dislodge, disengage, draw, eject, eliminate, eradicate, remove ...
THEMATIC ROLES: Agent Theme Source
SELECTIONAL RESTRICTIONS: Agent[+int_control OR +organization] Theme[] Source[+location]
FRAMES:

Transitive	Agent V Theme	cause(Agent, E) $\wedge$ $\neg$location(start(E), Theme, ?Source) $\wedge$ location(end(E), Theme, ?Source)
Transitive (+ Source PP)	Agent V Theme Prep[+src] Source	cause(Agent, E) $\wedge$ $\neg$location(start(E), Theme, Source) $\wedge$ location(end(E), Theme, Source)

Figure 1: The class remove-10.1 from VerbNet

Name	**Description**	**Sentence**	**Extracted Relation**
dobj	direct object	They win the lottery	dobj(win, lottery)
iobj	indirect object	She gave me a raise	iobj(gave, me)
prep/prepc	prepositional modifier	I saw a cat with a telescope	prep(saw, with)
prt	phrasal verb particle	They shut down the station	prt(shut, down)
nsubj	nominal subject	Clinton defeated Dole	nsubj(defeated, Clinton)
nsubjpass	passive nominal subject	Dole was defeated by Clinton	nsubjpass(defeated, Dole)
xsubj	controlling subject	Tom likes to eat fish	xsubj(eat, Tom)

Table 1: Typed Dependency Examples

In this paper, in section 2 we describe the three supervised learning approaches we experimented with, using three types of feature sets (section 2.1), and developing three different classifiers per model. We ran experiments on four datasets (section 3), and results can be found in section 4. As discussed in section 5, Naive Bayes performs surprisingly badly in all conditions. MaxEnt models always perform better than Decision Trees on manually built datasets such as VerbNet itself and WordNet; however, on our own HomeRepair corpus, Decision Trees perform better when a single classifier for all verbs is built, most likely because the VN class training data is somewhat noisy. Our best results are obtained with an ensemble of classifiers, one per polysemous verb.

2 Methodology

In this work, we were mainly interested in exploring the purported strength of the MaxEnt classification algorithm, with respect to more traditional models such as Naive Bayes and Decision Tree (DT) classifiers. MaxEnt is a uniform model, which makes no assumptions in addition to what we know from the data. It also has the strong capability to combine multiple and dependent knowledge sources, as opposed to the independence assumption underlying Naive Bayes. MaxEnt has been widely used in NLP and proven to be effective and efficient. Our hypothesis that Naive Bayes would perform poorly was borne out; however, the performance of the DT classifiers is competitive with that of MaxEnt, as we will discuss below.

We recast the VN class disambiguation problem as a classification problem, where a tuple (Sentence, Verb) needs to be assigned to the correct VN class. We devised two different classification models:

- **Single Classifier Model**: Train each classifier on all the verbs in the dataset.
- **Per-verb Classifier Model**: Train one classifier per each verb in the training set. Given a tuple (Sentence, Verb), we only use that verb's classifier to choose its VN class.

2.1 Features

We build our classification models using the following three types of features:

- **Lexical**: The word's base form and its POS tag.
- **Co-occurrence**: The words and POS tags which appear around the target verb in a window size of 5 (the window size of 5 was determined experimentally).
- **Typed Dependency**: All the dependencies where the verb to be disambiguated partic-

ipates, derived by means of the Stanford parser (de Marneffe and Manning, 2008).

We use lexical features and co-occurrence features as in (Girju et al., 2005; Abend et al., 2008); co-occurrence features are used to approximate collocations, since the collocation of a word can help decide its sense (Yarowsky, 1993). We add typed dependencies, since, compared to co-occurrence, grammatical relations capture more specific relations between the verb and other words in the sentence, and encode some of the structure of the sentence. We parse the sentences by means of the Stanford parser (de Marneffe and Manning, 2008) and the dependencies related to the verb to be disambiguated are extracted as part of the feature space. We use all dependencies available, some of which are presented in Table 1 with illustrative examples.

3 Data Sets

Our datasets are composed of all instances defined as follows: (Sentence, Verb, VN Class). We built three data sets according to where the sentences come from, plus a fourth that combines the other three. Whereas the goal of our work is to fully automatize parsing our HomeRepair corpus, other datasets are used to validate our approach. Each sentence is POS-tagged and parsed with the Stanford Dependency parser, to derive all the features we discussed earlier.

- **VerbNet**: For each frame in every VN class, VN provides one example sentence. The sentence is paired with the specific verb it contains. Clearly, this dataset should, and will, give rise to the most accurate results.
- **WordNet**: When VN provides members of a VN class, it also gives WordNet sense mappings when applicable. For example, the verb "instruct" in the VN class "advise-37.9" is mapped to WordNet sense "instruct%2:32:01". In turn, WordNet provides illustrative sentences as examples for each word sense. We extracted the example sentences to construct a dataset, but we excluded cases where VN provides multiple WordNet sense mappings.
- **HomeRepair**: the portion of the HomeRepair corpus that was used to evaluate the discourse parser in (Subba and Di Eugenio, 2009). As noted earlier, the VN class was obtained via LCFLEX and manual disambiguation of VN classes. However the data was parsed again with the Stanford parser to obtain Typed Dependencies.

Finally, we combined the 3 datasets above to build a larger dataset. The sizes of those datasets are listed in Table 2. In that table, *Instance* gives the number of (Sentence, Verb, VN Class) tuples; *Verb* is the number of distinct verbs in the dataset; *Class* is the total number of distinct VN classes in that dataset. Note that for VerbNet, the number of verbs in Table 2 is much lower than the number of distinct verbs we mentioned above. This is due to the fact that, for each (sub)class, VerbNet uses only one representative of the (sub)class in all the examples for that (sub)class.

4 Experiments

We used OpenNLP Tools,[2] an open source Java NLP library that provides a collection of basic text processing tools for tasks like sentence detection, tokenization, part-of-speech tagging, text chunking, named entity recognition, co-reference resolution and tree parsing. OpenNLP also includes the MaxEnt[3] Java package for Maximum Entropy modeling. We employed the data mining tool package Weka (Hall et al., 2009) for Naive Bayes and Decision Tree classifications (Weka's J48 implementation was used for Decision Trees). The JWNL (Java WordNet Library)[4] was used to extract the WordNet dataset.

For each approach, we conducted three sets of experiments, for each dataset we described in Section 3. In all experiments we used exactly the same features by building a converter to convert the extracted features for MaxEnt to Weka's Attribute-Relation File Format(.arff) data format. In all experiments, the MaxEnt models were Generalized Iterative Scaling(GIS) models trained through 100 iterations and with no cut off. All the accuracies are calculated with 10-fold cross validation. Our baseline model assigns a (Verb, Sentence) tuple to the majority class of the verb in the training data set.

The first set of experiments used the entire datasets, no matter whether a verb is polysemous or not; results are shown in Table 2. Weka failed to

[2] http://opennlp.sourceforge.net/
[3] http://maxent.sourceforge.net/
[4] http://sourceforge.net/projects/jwordnet/

Dataset	Instance	Verb	Class	Baseline	NaiveBayes	J48	MaxEnt
VerbNet	838	310	265	0.9078	0.3059	0.8091	**0.9558**
WordNet	1586	1108	224	0.8432	0.1883	0.4678	**0.8877**
HomeRepair	2111	293	127	0.8115	0.4424	**0.8423**	0.7335
Combined	5633	1523	329	0.7800	N.A.	0.7264	**0.8528**

Table 2: Single Classifier on All Data

Dataset	Instance	Verb	Class	Baseline	NaiveBayes	J48	MaxEnt
VerbNet	431	151	140	0.8385	0.3364	0.7819	**0.9077**
WordNet	516	257	146	0.5354	0.1686	0.3817	**0.6375**
HomeRepair	1353	154	98	0.5458	0.4804	**0.8064**	0.6776
Combined	2300	418	233	0.5083	0.3561	0.7130	**0.7283**

Table 3: Single Classifier on Polysemous Verbs

generate any result when we ran Naive Bayes on the combined data set, since it ran out of memory even after we assigned to it 2 GB of memory, the maximum amount we had available.

The second set of experiments is restricted to only polysemous verbs (see Table 3). The degree of attested polysemy for the three datasets hovers just above 2, with VerbNet the lowest at 2.08, and the combined set the highest at 2.24 (the same VN class inventory is used in each set).

The third set of experiments uses the per-verb classifier model. In each iteration, for every verb, we select all the instances that include that verb, and split them according to 10-fold validation. For verbs which have less than 10 instances, we randomly choose one of them as testing instance, and use the others for training. We ran this set of experiments only on those polysemous verbs for which there are at least 3 instances of each pair (Verb, VN Class). Results are shown in Table 4.

5 Discussion

Not surprisingly, in all experiments, results on VerbNet are always very high for all classification models. This is due to the consistency of VN data, because VerbNet always uses the same verb to give examples for a VN class.

Baseline gave very high accuracy in experiment set 1 (see Table 2). This is not surprising, since the complete data set contains a large portion of verbs which are not polysemous.

Naive Bayes became the real baseline, since it always performs worst, and by far, in all the experiments. We believe it is because of the nature of posterior probabilities, and lack of independence among features. For co-occurrence features and Typed Dependency features, the feature value space is too big. Additionally, the overlap of values among different co-occurrence features violates the assumption of independence underlying Naive Bayes.

In almost every case, MaxEnt models perform better than the other models. In most cases, χ^2 shows that these differences are significant at the $p \leq 0.05$ level.

MaxEnt did worse than J48 is the single classifier experiments on the HomeRepair data, both on all verbs and on polysemous verbs only (Tables 2 and 3). As we noted earlier, in HomeRepair the VN class data was obtained by employing the LCFLEX parser, and then manual choice of the correct parse when more than one was returned. Both the parser itself and the manual disambiguation introduce noise: there are similar sentences with the same verb, where the two verb instances are assigned to two different VN classes. Please note that in Table 3, whereas MaxEnt performs better than J48 on the combined set, the difference is not significant. In general, J48 performs poorly on the WordNet set when a single classifier is trained (Tables 2 and 3).

We obtained our most promising result with the per-verb classifier ensemble, and on the HomeRepair / combined corpora (Table 4). Whereas the accuracy drops considerably with respect to VerbNet in the other experiments, it does not here. We note however that in this table, the difference in performance between MaxEnt and baseline on VerbNet is not significant (all other differences are).

It is not possible to draw a real comparison

Dataset	Instance	Verb	Class	Baseline	NaiveBayes	J48	MaxEnt
VerbNet	277	49	52	0.9429	0.3704	0.9259	**0.9667**
WordNet	158	30	33	0.6750	0.3333	0.7095	**0.8378**
HomeRepair	1221	81	59	0.7764	0.4194	0.8065	**0.8956**
Combined	1858	164	145	0.7113	0.2752	0.6833	**0.8986**

Table 4: Per-Verb Classifier on Polysemous Verbs

with work in the literature because we use different datasets. However, at a high level we note that (Girju et al., 2005) uses data derived from PropBank, but finds that only about 4% of verbs are polysemous, and on this set their best model achieves around 80% accuracy. (Abend et al., 2008) performs at around 92% when tested on the Wall Street Journal, but when the model is applied to medical tests, it falls to 55%. Because VerbNet is domain independent, we expect our per-verb classifier trained on the combined datasets to be accurate on other domains as well. This claim clearly needs to be tested on other datasets.

6 Future Work

As just stated, we intend to explore the applicability of our models, specifically the ensemble of classifiers trained on the whole dataset, to other corpora.

One of the remaining issues is handling unseen verbs in the training data. We believe our single classifier model will be able to handle it, but we need to design experiments to evaluate the performances.

Another issue is how to generalize the Typed Dependency features we employ. Because the dependency arguments we extracted are not generalized, when the pre-labeled training data set is small, the extracted features will be hard to match incoming examples. One promising approach is to generalize the arguments to the dependencies. For example, we could use CoreLex (Buitelaar, 1998) to generalize nouns to CoreLex classes. Another way to generalize Typed Dependencies is to use Dynamic Dependency Neighbors as employed by (Dligach and Palmer, 2008).

Acknowledgments

This work is supported by award IIS-0905593, National Science Foundation. Thanks to three anonymous reviewers for their suggestions.

References

Omri Abend, Roi Reichart, and Ari Rappoport. 2008. A supervised algorithm for verb disambiguation into verbnet classes. In *Proceedings of the 22nd International Conference on Computational Linguistics*, volume 1. Association for Computational Linguistics.

Satanjeev Banerjee and Ted Pedersen. 2010. An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet. In *Computational Linguistics and Intelligent Text Processing*, volume 2276 of *Lecture Notes in Computer Science*, pages 117–171. Springer Verlag.

Paul Buitelaar. 1998. *CoreLex: Systematic Polysemy and Underspecification*. Ph.D. thesis, Computer Science, Brandeis University, February.

Hoa T. Dang and Martha Palmer. 2005. The Role of Semantic Roles in Disambiguating Verb Senses. In *Proceedings of the 43rd Annual Meeting of the Association f or Computational Linguistics*, pages 42–49, Ann Arbor MI.

Marie-Catherine de Marneffe and Christopher D. Manning. 2008. Stanford typed dependencies manual. Technical report, Stanford University.

D. Dligach and M. Palmer. 2008. Novel semantic features for verb sense disambiguation. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Short Papers*, pages 29–32. Association for Computational Linguistics.

Roxana Girju, Dan Roth, and Mark Sammons. 2005. Token-level disambiguation of verbnet classes. In *Proceedings of The Interdisciplinary Workshop on Verb Features and Verb Classes*.

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. 2009. The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18.

Karin Kipper, Anna Korhonen, Neville Ryant, and Martha Palmer. 2008. A Large-Scale Classification of English Verbs. *Journal of Language Resources and Evaluation*, 42(1):21–40.

Mirella Lapata and Chris Brew. 2004. Verb class disambiguation using informative priors. *Computational Linguistics*, 30(1):45–73.

Carolyn Penstein Rosé and Alon Lavie. 2000. Balancing robustness and efficiency in unification-augmented context-free parsers for large practical applications. In J.-C. Junqua and G. van Noord, editors, *Robustness in Language and Speech Technology*. Kluwer Academic Press.

Rajen Subba and Barbara Di Eugenio. 2009. An effective discourse parser that uses rich linguistic information. In *North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT)*, Boulder, CO, June.

Rajen Subba, Barbara Di Eugenio, and Elena Terenzi. 2006. Building lexical resources for PrincPar, a large coverage parser that generates principled semantic representations. In *LREC06, the fifth International Conference on Language Resources and Evaluation*, pages 327–332, Genova, Italy, May.

David Yarowsky. 1993. One sense per collocation. In *Proceedings of the workshop on Human Language Technology*, pages 266–271. Association for Computational Linguistics.

Argument linearization in the production of German and Dutch verbs

Kees de Schepper
Radboud University Nijmegen
k.deschepper@let.ru.nl

Monique Lamers
VU University Amsterdam
mja.lamers@let.vu.nl

Abstract

In this paper we will compare the results of a German and a Dutch production experiment on argument linearization patterns (subject-before-object, object-before-subject and passive). We distinguish between several factors that have an effect on the outcome (agent vs. experiencer, dative vs. accusative, animate vs. inanimate) and conclude that SubjectFirst, AgentFirst, ExperiencerFirst and DativeFirst are all important principles, but AnimateFirst is not, surprisingly.

1 Introduction

One linguistic phenomenon that is intrinsically tied to verbs is *argument linearization*. Three well-known linearization patterns are: (i) *subject-before-object (SbO)* as in *John wrote this book*; (ii) *object-before-subject (ObS)* as in *This book, John wrote*; and (iii) *passive* as in *This book was written by John*. Studies on the perception of linearization patterns have identified several important factors: animacy (animate or inanimate), case (nominative, accusative, dative) and semantic roles (agent, patient or experiencer). This paper will present a novel view on linearization from the production perspective brought about by the results of two sentence production studies, one in German and one in Dutch, which shall disentangle the above-mentioned factors.

Although Dutch and German are two closely related languages, they differ in many aspects. Dutch, with no overt case marking of full noun phrases, exhibits a strong preference for SbO sentences (e.g. Lamers, 2005). For German, a language with overt case marking of full noun phrases, SbO preference seems to be less robust. Psycholinguistic perception studies have shown that for sentences with verbs that assign dative case the ObS order is preferred (e.g. Bornkessel & Schlesewesky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009).

2 Method

By using different types of verbs the influence of case marking from other factors that might influence the linearization of the arguments was isolated. In both studies participants were asked to construct a sentence using the words (two arguments and a verb) provided in a prompt (see Ferreira, 1994). The four types of verbs and their characteristics can be found in Table 1. The experiment featured six different verbs per verb type.

Verb type	Restrictions	Passivize	Case on object	German/Dutch example
Standard Agentive	Animate subject	Yes	Accusative	*kritisieren/bekritiseren* ‘to criticize’
Caustive Psych	Animate object	Yes	Accusative	*verblüffen/verbazen* ‘to amaze’
Unaccusative Psych	Animate object	No	Dative	*gefallen/bevallen* ‘to please’
Dative Agentive	Animate object	No	Dative	*schaden/-* ‘to damage’

Table 1: Verb types and their characteristics (Agentive Dative verbs do not exist in Dutch)

3 Results

The results of the two studies show not many differences between Dutch and German. Overall, more SbO than ObS sentences were produced. Stimuli with Causative Psych verbs resulted in more passive constructions than with Standard Agentive verbs. ObS structures were most frequent with Unaccusative Psych verbs. See Table 2 and Table 3 for an overview.

However, one difference is that in German prompts with Unaccusative Psych verbs resulted in more ObS sentences than in Dutch. Furthermore, for German two additional issues could be tested: the effect of argument animacy and the class of Dative Agentive verbs. The effect argument animacy was surprisingly small. Only for Dative Agentive verbs there was an important difference: a prompt with one animate and one animate argument resulted in more object-initial sentences than a prompt with two animate subjects.

Verb type	**Example**	**Argument animacy**	**Subject-initial**	**Passive**	**Object-initial**	**Other**
Standard Agentive	‘to criticize’	1 animate 1inanimate	79	16	0	6
Causative Psych	‘to amaze’	1 animate 1inanimate	60	27	2	11
Unaccusative Psych	‘to please’	1 animate 1inanimate	61		24	14

Table 2: Dutch linearization patterns (in percentages) for a number of input combinations

Verb type	**Example**	**Argument animacy**	**Subject-initial**	**Passive**	**Object-initial**	**Other**
Standard Agentive	‘to criticize’	1 animate 1inanimate	84	4	2	9
		2 animates	83	3	1	13
Causative Psych	‘to amaze’	1 animate 1inanimate	60	22	10	8
		2 animates	67	20	4	10
Unaccusative Psych	‘to delight’	1 animate 1inanimate	37		54	8
		2 animates	39		47	14
Dative Agentive	‘to damage’	1 animate 1inanimate	65	1	28	6
		2 animates	88	2	3	8

Table 3: German linearization patterns (in percentages) for a number of input combinations

4 Discussion

To explain the differences in patterns between sentences with different types of verbs on the one hand, and between the two languages on the other hand, we follow a multifactorial approach as proposed by Primus (1999, 2009; see also Lamers, to appear). In her approach argument realization results from the interplay of multiple factors. These factors give us several competing prominence principles (e.g. SubjectFirst, AnimateFirst, and AgentFirst).

Our results show that SubjectFirst must be a very strong principle: overall there are less ObS than SbO sentences. We are also in need of a No-Passivization principle, as there are not many passive sentences overall either.

The Causative Psych verbs show that an AnimateFirst principle cannot be very strong: the differences between verbs with two animates and verbs with one animate and one inanimate here is negligible. Instead, an ExperiencerFirst principle seems to be relevant, as a great number of passives are produced for Causative Psych verbs, and after passivization the Experiencer argument is in front. Yet, the NoPassivization principle is still stronger, because the number of subject-initial sentences produces is still rather large.

ExperiencerFirst is also at work in Unaccusative Psych verbs. Here this principle results in an increased number of object-initial sentences, as passivization is ungrammatical for these verbs. Furthermore, there is a difference between German and Dutch: in the German study far more object-initial sentences were constructed. We propose that morphological case is the cause of this difference. In German objects of Unaccusative Psych verbs are overtly case-marked. Hence, if a dative argument is fronted the dative signals to the addressee that a non-subject is fronted. Therefore it is easier in German to front the object of an Unaccusative Psych verb. This analysis does not hold for accusative arguments. Probably, dative arguments like to be fronted because they are prototypically animate.

This DativeFirst principle also applies to Dative Agentive verbs. However, since the objects of these verbs do not have the Experiencer role, the ExperiencerFirst principle does not apply, resulting in smaller numbers/occurrences of object-initial sentences. Furthermore, when the subject of an Dative Agentive verb is animate, it is virtually always seen as having the Agent role. Finding almost no ObS structures may thus be explained in terms of the preference to place the Agent in sentence initial position (AgentFirst principle).

Surprisingly, AgentFirst applies to Standard Agentive verbs, but not to Causative Psych verbs. The subject of Causative Psych verbs is seldom seen as an Agent – not even when this argument is animate – and therefore AgentFirst does not apply, see Broekhuis (in prep.). Our data supports this analysis: where 100% of the passives with German Standard Agentive verbs are unambiguously agentive (i.e. use *werden* ‘to become’ as the passive auxiliary), only 9% of the passives with German Causative Psych verbs are.

Finally, Table 4 provides an overview of the way in which each verb diverges from the standard pattern of producing an active, subject-initial sentence.

5 Conclusion

In sum, we conclude that there are several important principles in choosing a linearization pattern in the production of German and Dutch: SubjectFirst, ExperiencerFirst, DativeFirst and AgentFirst, for example. Surprisingly, the AnimateFirst principle seems to be of little influence on the linearization process. Note, however, that AgentFirst, ExperiencerFirst and DativeFirst all apply to obligatorily animate arguments. As such, animacy *is* of great importance.

Verb type	Example	Argument animacy	German	Dutch
Standard Agentive	*‘to criticize’*	1 animate 1inanimate	$\chi_{\text{AgentFirst}}$	$\chi_{\text{AgentFirst}}$
		2 animates	$\chi_{\text{AgentFirst}}$	
Causative Psych	*‘to amaze’*	1 animate 1inanimate	$\surd_{\text{ExperiencerFirst}}$	$\surd_{\text{ExperiencerFirst}}$
		2 animates	$\surd_{\text{ExperiencerFirst}}$	
Unaccusative Psych	*‘to please’*	1 animate 1inanimate	$\surd_{\text{ExperiencerFirst}}$ $\surd_{\text{DativeFirst}}$	$\surd_{\text{ExperiencerFirst}}$
		2 animates	$\surd_{\text{ExperiencerFirst}}$ $\surd_{\text{DativeFirst}}$	
Dative Agentive	*‘to damage’*	1 animate 1inanimate	$\surd_{\text{DativeFirst}}$	
		2 animates	$\chi_{\text{AgentFirst}}$	

Table 4: Divergence from producing active, subject-before-object (SbO) sentences in German and Dutch (χ = no divergence; $\surd$ = minor divergence; empty cell = not investigated)

References

Bornkessel, I. & M. Schlesewsky. 2006. The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. *Psychological Review*, 113.787-821.

Bornkessel-Schelswesky, I. & M. Schlesewsky. 2009. The role of prominence information in real time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass*. 3/1, 19-58.

Broekhuis, H. In preparation. The subject of causative object experiencer verbs.

Ferreira, F. 1994. Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33, 615-636.

Lamers, M. 2005. Resolving subject-object ambiguities with and without case: evidence from ERPs. *Case-marking: competition and variation*, ed. by M. Amberber and H. de Hoop, 251-293: Elsevier.

Lamers, M. To appear. Argument linearization in Dutch: a multi-factorial approach.

Primus, B. 1999. *Cases and thematic roles - ergative, accusative and active*. Tübingen: Niemeyer.

Primus, B. 2009. Case, grammatical relations, and semantic roles. *The Oxford handbook of case*, ed. by A. Malchukov and A. Spencer, 261-75. Oxford: Oxford University Press.

Class/Subclass	**Unique verb occurrences in two biology papers.**
Discourse verbs *Thompson and Ye: 'Textual verbs'*	address, base, depict, describe, mention, note, report, represent, review, **show**, study, **suggest**, term
Research Verbs *Thompson and Ye: 'Research verbs'*	
Investigation *Thomas and Hawes: 'Procedural verbs';* *Williams: 'Methods'*	**compare**, **demonstrate**, **detect, determine**, **elucidate**, **evaluate**, **examine**, **exclude**, exemplify, expose, extend, find, **identify**, investigate, pinpoint, mimic, **remain**, **require**, require, shed [light], start identify, strengthen, **substantiate**, test, verify
Procedure *Thomas and Hawes: 'Procedural verbs';* *Williams: 'Methods'*	accumulate, **activate**, adapt, administer, affix, allow recover, analyze, anesthetize, annotate, base, calculate, characterize, clone, **compare**, conduct, conform, contain, connect, conserve, consist, construct, **control**, cotransfect, correspond, create, derive, **determine**, develop, dissolve, divide, drill, employ, enrich, **evaluate**, **express**, find, follow, frozen, generate, handle, harbour, house, immortalize, **impair**, implant, include, infuse, insert, introduce, involve, keep, leave, localize, look, lose, lower, make, minimize, mix, model, mount, mutate, obtain, overcome, overlap, perform, permit, place, possess, present, prevent, purchase, reduce, remove, replace, resemble restrain, retract, section, serve, share, spend, stabilize, synthesize, take, transduce, transfect, **use**
Observation *Thomas and Hawes: 'Objective verbs'* *Williams: 'Observation verbs'*	**characterize**, **compare**, **correlate**, detect, detect, **express**, **find**, **identify**, monitored, note, observe, see, **show**
Sensemaking Verbs	
Prediction *Thomas and Hawes: 'Pre-experiment verbs'*	**elucidate**, hypothesize, **involve**, point to, predict, propose, provide [indication], raise, **remain**, seem, **suggest**
Interpretation *Thomas and Hawes: 'Post-experiment verbs'*	associate, conclude, conjecture, **demonstrate**, **exclude**, explain, **implicate**, indicate, provide, provide [evidence], reveal, **show**, stress, **substantiate**, **suggest**, support, underpin
Comparison *Hyland: 'Evaluative verbs'*	**compare**, confirm, expect, provide, underpin, validate
Cognition *Thomas and Hawes: 'Cognition verbs'/Biber: 'Mental verbs'*	choose, concern, decide, emphasize, **examine**, **exclude**, infer, judge, know, **remain**, take [advantage of]
Object Properties and Relations	
Cause and Effect *Williams: 'Cause and Effect'*	abolish, abrogate, accelerate, act, affect, allow, attenuate, block, bypass, cancel, cause, circumvent, collaborate, confer, connect, contribute, **control, correlate**, degrade, depend, disinhibit, disrupt, encode, enhance, exert, **express**, facilitate, fail [to express], fail [to discriminate], have [an effect], **impair**, **implicate**, improve, induce, inhibit, **involve**, lead, make [resistant to], mediate, modify, neutralize, numb, obtain, **overcome**, participate, permit, play [a role], predict, prevent, provoke, reduce, reflect, regulate, reinforce, relate, replace, **require**, result, reverse, **show**, silence, stimulate, suppress, target, undergo, underlie, **use**, yield
Change and Growth *Williams: 'Change and Growth'*	amplify, cease, continue, disrupt, downregulate, exert, expand, **express**, grow, increase, mimick, proliferate, reach, **remain**, **show**, **spend**
Properties *Williams: 'Defining verbs' (is a subset)*	accumulate, **activate**, **characterize**, conform, conserve, consist, contain, correspond, divide, enrich, exist, express, **find**, harbour, have, **impair**, include, **involve**, localize, lose, **overcome**, overlap, possess, resemble, share, **spend**, stabilize

Appendix. Instances of the verb class taxonomy for two full-text biology papers Voorhoeve and Louiseau).
On the left, our taxonomy; in italics, overlap with other taxonomies. On the right, verbs found in two full-text biology papers, classified according to this taxonomy. Bold indicates that verbs occur in more than one category.

Extracting prototypes from corpus data: a distributional account of representing near-synonymous verbs

Dagmar Divjak
University of Sheffield
Sheffield, United Kingdom
d.divjak@sheffield.ac.uk

Antti Arppe
University of Helsinki
Helsinki, Finland
antti.arppe@helsinki.fi

Abstract

In cognitive linguistics, *prototype* theory is currently one of the dominant views of how linguistic categories are stored and represented as cognitive structures in the brain. Yet two problems arise: Cognitive linguistics is a usage-based theory but has thus far not attempted to show how prototypes can be observed in usage in a systematic way. Furthermore, the bulk of the research done has focused on prototypes for nouns, denoting tangible objects, rather than verbs that denote intangible events. In this paper, we simulate how abstract prototypes for verbs could be formed using statistical learning mechanisms that track frequency distributions in input on the basis of actual usage as observed in corpus data.

1 Introduction

Nearly four decades ago Eleanor Rosch (1973 and later work) demonstrated the inadequacy of necessary and sufficient attributes for item classification. Instead, she presented a prototype approach to categorization, a probabilistic feature approach with instances displaying different degrees of representativity and similarity to a prototype. That prototype representation of a category is generally taken to be a generalization or abstraction of a class of instances falling into the same category.

In cognitive linguistics, *prototype* theory is one of the dominant views of how linguistic categories are stored and represented as cognitive structures in the brain (Taylor 1995). Yet, although cognitive linguistics actively promotes itself as a usage-based theory, thus far it has not been shown how prototypes can be observed in actual usage in a systematic and cognitively realistic way.

Moreover, the bulk of experimental and linguistic research done on prototype categorization has concentrated on nouns (Pulman 1983). A basic difference between nouns and verbs is that, typically, nouns describe items that are stable in time and therefore independent of that dimension, whereas verbs describe items that are neither stable in nor independent of time. In addition, nouns typically denote tangible objects, whereas verbs name intangible events. And thirdly, verbs render relational concepts, which implies that they are more susceptible to their meanings being influenced by the concepts they relate. This implies that prototypical situations are partly determined by the elements verbs co-occur with. It is precisely this contextual element that we aim to exploit in our corpus-based quest for a cognitively realistic and systematic procedure of extracting verbal prototypes from language use.

2 Methodology

We do so by statistically modeling large annotated datasets of exemplars and gradually reducing exemplars while abstracting properties. To this end, we build upon the results of the application of a multivariate statistical technique, *polytomous logistic regression* (see e.g. Arppe 2008) according to the *one-vs-rest* heuristic (Rifkin & Klautau 2004) which was used to study the contextual similarities and differences of two sets of Russian and Finnish near-synonyms expressing TRY and THINK. These two synonym sets from two typologically distinct languages have been selected for the practical reason that they have been the object of recent large-scale corpus-based studies (Arppe 2008 and Divjak 2010) exploring the phenomenon near-synonymy from different angles, which have produced extensive datasets for further analyses such as the one presented here.

2.1 Data

Data on the six most frequent Russian verbs that express TRY when combined with an infinitive, i.e. *probovat', pytat'sja, starat'sja, silit'sja, norovit', poryvat'sja,* were extracted from the Amsterdam Russian Corpus, the Russian National Corpus and (selected) Internet pages. In all, there were 1,351 occurrences of this syntactically homogenous category (i.e. all verbs share the same argument structure). Depending on the frequency of the verb, between 100 and 250 examples were annotated per verb.

For Finnish, the four most frequent synonyms meaning 'think, reflect, ponder, consider', i.e. *ajatella, miettiä, pohtia, harkita*, were extracted from two months of newspaper text (Helsingin Sanomat 1995) and six months of Internet newsgroup discussion (SFNET 2002-2003). In all, there were 3,404 occurrences of this syntactically non-homogenous category (i.e. not all verbs share exactly the same argument structure), with frequencies ranging from 1,492 for the most common one *ajatella* to 387 for the rarer *harkita*.

2.2 Annotation

For Russian, the 1,351 examples were tagged using the annotation scheme from Divjak & Gries (2006). This scheme captures all information provided at the sentence level by tagging for morphological properties of the finite verb and the infinitive (tense, aspect, mode), syntactic properties of the sentences (sentence type, clause type) and semantic properties of the infinitive (semantic type of subject, properties of the event denoted by the infinitive, controllability of the infinitive action) as well as optional elements (adverbs, particles, negation). The final tagset contains 14 variables amounting to 87 variable categories. This annotation scheme thus contains all elements ecountered within sentence boundaries and can, as such, be transferred to the annotation of other verbs, e.g. verbs expressing INTENTION (Divjak 2006, 2010) or RESULT (Divjak 2003, 2010).

For Finnish, the 3,404 examples were first morphologically and syntactically analyzed using an implementation of the Functional-Dependency Grammar (FDG) parser (Tapanainen & Järvinen 1997) for Finnish, after which all the instances of the studied verbs together with all their relevant associated context (not limited merely to obligatory syntactic arguments) were manually checked, corrected and supplemented with semantic subclassifications. The morphological level of analysis of the node verb covered subtypes of infinitive and participle, non-finite case, number and possessive suffix (indicating person and number), polarity, voice, mood, simplex tense, and finite person-number, whereas that of the entire verb chain of which the THINK verb was part of concerned polarity, voice, mood, an aggregate of person and number marking for both finite or non-finite verb forms, and surface-syntactic role. The syntactic argument types follow those of the FDG formalism, and the semantic and structural subtyping was a combination of various schemes including WordNet (Miller et al. 1990), several prior Finnish studies (Pajunen 2001, Kangasniemi 1992 and Flint 1980) and an evidence-based bottom-up classification procedure suggested by Hanks (1996).

Although the two analysis schemes have different starting points (i.e. an argument structurally homogenous category for Russian versus an argument structurally varied category for Finnish) and, as a result, operate with a different set of analytical categories, they are nevertheless similar in trying to grasp the immediate context in its entirety. Moreover, using such two distinct schemes is a test of the overall robustness of the statistical modelling and analysis, provided we are able to produce effectively similar results.

2.3 Statistical modeling

We modeled the annotated corpus data using polytomous logistic regression (see e.g. Arppe 2008).[1] The one-vs-rest heuristic (Rifkin & Klautau 2004) distinguishes each member of the set without requiring a baseline category and directly provides lexeme-specific odds with respect to selected variables (representing linguistic properties). It models probabilities of occurrence given a particular context. The variable parameters it estimates can be naturally interpreted as *odds* (Harrell 2001). As a simple selection rule, the verb receiving the highest estimated probability

[1] Since *multinomial logistic regression* is often used to refer in effect to only a particular heuristic out of many possible ones, i.e. where a set of (binary) baseline models are fitted simultaneously and in relation to each other with a given algorithm, we use the term *polytomous logistic regression* modeling as an umbrella concept for any heuristic tackling polytomous (i.e. more than two alternatives) outcomes as long as it is based on logistic regression analysis, regardless of how the polytomous setting is broken down into a set of binary models and whether these component binary models are separately or simultaneously fitted (for an overview, see Arppe 2008).

is picked for any given context representing a cluster of properties, i.e., $arg_{Verb}\, max[P(Verb|Context)]$. The highest estimated probability is not necessarily always close to P=1.0 or even P>0.5 but can range from slightly over $1/n$ (n indicating the overall number of outcomes) to 1.0. Moreover, since the constituent binary logistic regression models are fit separately with the one-vs-rest heuristic, the sums of their instance-wise probabilities are not always exactly $\sum P=1.0$. Therefore, the verb-specific probabilities for each instance in both data sets are adjusted to satisfy this condition by dividing, instance-wise, each original lexeme-specific probability estimate by the sum of these estimates for that particular instance.

The original variable sets were pruned since the number of variables allowed in logistic regression is maximally 1/10 of the frequency of the rarest outcome. The selection strategy we adopted for the Russian TRY lexemes was to retain variables with a broad dispersion among the verbs. We required the overall frequency of the variable in the data to be at least 45 and to occur at least twice with all verbs. Additional technical restrictions excluded one variable for each fully complementary case (e.g. the aspect of the verb form) as well as variables with mutual pairwise association statistic Uncertainty Co-Efficient (Theil 1970) UC>0.5 (i.e. knowledge of one variable decreases more than ½ of the uncertainty concerning the other). In the end, 18 property variables remained.

For the Finnish THINK lexemes, a minimum overall frequency was required, in this case set at $n \geq 24$. Pair-wise associations of individual properties were likewise carefully evaluated using UC, but due to the heterogeneity of the argument structure of the Finnish THINK verbs, occurrence with all four verbs was not required. Semantic subtypes were included only for the most frequent syntactic argument types, and many contextual property variables were lumped together, whenever possible and appropriate. In the end, 46 linguistic property variables were chosen for the full model, of which 10 were morphological, concerning the entire verb chain, 10 simple syntactic arguments (without any semantic subtypes), 20 combinations of syntactic arguments with semantic and structural subclassifications, and 6 semantic characterizations of the entire verb chains.

2.4 Model performance

In the case of the six Russian TRY verbs, 51.7% of all cases were correctly predicted (i.e. Recall) according to the prediction rule of selecting the verb with the highest estimated probability. The Recall rate for the four Finnish THINK verbs was 64.6%. Comparing these percentages to the 52.7% correct answers the average non-English US college applicant provided in a 4-way choice between semantically related verbs such as *imposed, believed, requested and correlated* (Landauer and Dumais 1997) confirms that the statistical models perform at a rate comparable to that of human beings.

3 Results

3.1 Property-wise verb-specific odds

The one-vs-rest analysis technique has two key attractive characteristics as stepping stones towards showing how prototype formation may be achieved on the basis of usage data.

Firstly, a model created with polytomous logistic regression provides probability estimates for the (proportional) occurrence of an outcome, such as a verb within some synonym set, given the contextual occurrence of some combination of linguistic properties incorporated in the model. Secondly, and crucially, the *one-vs-rest* heuristic can be understood to highlight those properties which distinguish the individual outcome classes (in this case the near-synonymous verbs) from all the rest (within the same set), in natural terms as *odds*. Individual odds (parameter values) which are greater than 1.0 for some property and the singled-out verb can be interpreted to reflect the increased chances of occurrence of this verb when the property in question is present in the context. Conversely, odds less than 1.0 denote a decreased chance of the occurrence of this verb in such a context. As an example case, take the Russian *probovat'*, for which the property-wise odds are shown in Table 1.

PROPERTY/VERB	ODDS
(Intercept)	1:22
CLAUSE.MAIN	**3.4:1**
FINITE.ASPECT_PERFECTIVE	**29:1**
FINITE.MOOD_GERUND	1:8.3
FINITE.MOOD_INDICATIVE	1:2.8
FINITE.TENSE_PAST	(1:1)
INF….ASPECT_IMPERFECTIVE	**6.1:1**
INF….CONTROL_HIGH	(1:1.2)
INF….SEM_COMMUNICATION	**2.1:1**
INF….SEM_EXCHANGE	(1.4:1)
INF….SEM_METAPH…_MOTION	(1.5:1)
INF….SEM_METAPH…_PHYSICAL_EXCHANGE	(1:1.3)
INF….SEM_METAPH…_PHYSICAL_OTHER	(1.3:1)
INF….SEM_MOTION	(1.7:1)
INF….SEM_MOTION_OTHER	(2.6:1)
INF….SEM_PHYSICAL	**3.9:1**
INF….SEM_PHYSICAL_OTHER	**2.5:1**
SENTENCE.DECLARATIVE	1:2.8
SUBJECT.SEM_ANIMATE_HUMAN	(1.5:1)

Table 1: Odds for/against Russian *probovat'* (Odds in parentheses are non-significant)

3.2 Aggregating properties as a prototype

In the model, those verb-specific linguistic properties – *per definition* abstract generalizations – which have significant odds in favor of a verb can be aggregated to construct an abstraction which as a whole embodies and represents the prototype of each verb, when contrasted with the rest of the verbs in either near-synonym set.

For the Russian TRY verbs, out of a total of 1,351 individual property combinations, 660 combinations of a distinct verb plus a context type can be distinguished (reducing to 296 if the outcome verb is ignored), leading to 20 permissible property combinations with significantly favorable odds, and ultimately to as few aggregates of properties with such strongly favorable odds as there are verbs. For *probovat'*, the set of such properties with significant odds in favor of this verb occurring when they are evident in the context are boldfaced in Table 1. Note that only one of the three semantic characterizations of the infinitive can possibly be observed at the same time. Thus, the aggregate of properties in fact represents three permissible property combinations.

For the Finnish THINK verbs, out of a total of 3,404 individual property combinations, 2,196 combinations of distinct property clusters with (one of the) verbs can be identified, which reduces only slightly to 1,908 if the outcome verb is ignored. This is a result of the heterogeneity of the allowed argument structures of the Finnish THINK verbs (versus the syntactic homogeneity of the Russian TRY verbs), as well as the greater overall number of properties included in the analysis. Due to this syntactic heterogeneity and optionality of many arguments and properties, practically only a lower bound can be estimated of altogether at least 51 permissible combinations of properties with significant favorable odds for the four THINK lexemes, distributed as follows: *ajatella* (32), *miettiä* (8), *pohtia* (10), and *harkita* (1).

4 Discussion

The aggregated properties with significant odds in favor of a verb are, as a whole, manifestations of (the core of) a prototype for a verb. It is plausible to interpret the above properties for *probovat'* as conveying the notion of telling someone to *try* (using the perfective aspect hence signaling the *attempt* should be taken to its natural conclusion and with limitations imposed on the time or effort invested), and carry out a physical action, to manipulate someone or something, or to communicate (using the imperfective, i.e. without insisting that the attempted *action* be taken to its natural end). This interpretation of *probovat'* explains why this verb is typically characterized as an "experimental attempt" (Apresjan et al. 1999), and why it is the most frequently used TRY verb in mother-child interaction (Stoll corpus, see Divjak & Gries 2006).

This definition has been distilled from the extracted estimated odds over properties that predict which of the near-synonymous alternatives is most likely to be selected given a specific linguistic context. Over the past decade, numerous studies have been published supporting the claim that infants are equipped with powerful statistical language learning mechanisms. If speakers model input statistically, as is assumed by statistical learning (cf. Saffran et al. 1996), they may be operating with similar prototypes as the regression technique outputs.

Nevertheless, a caveat needs to be expressed. We have aimed to model produced language systematically by means of a statistical heuristic, regression analysis, yet the heuristic by which this model is constructed and the constituent binary logistic regression models and mathematical algorithms by which they are optimized to fit to the data were not designed to mimic cognitive behavior. The resulting model fits descriptions that linguists feel are appropriate for the data, but the underlying mechanics of regression analysis lacks cognitive grounding other than the fact that human beings seem able to detect statistical regularities in input.

References

Apresjan, J. D. 1999, *Новый объяснительный словарь синонимов русского языка. Vol. I.* Moskva: Škola "Jazyki Russkoj Kul'tury".

Arppe, A. 2008. *Univariate, bivariate, and multivariate methods in corpus-based lexicography – A study of synonymy*. Publications of the Dep't of General Linguistics, University of Helsinki, 44.

Divjak, D. 2003. Слова о словах. К вопросу о неточных синонимах «умудриться», «ухитриться», «исхитриться» и «изловчиться. [Words on Words. The Case of Four Near Synonyms] In: Soldatjenkova, T., Waegemans, E. For East is East. Liber Amicorum Wojciech Skalmowski. Leuven-Paris: Peeters, 345-365. [Orientalia Lovaniensia Analecta].

Divjak, D. 2006. Ways of Intending: Delineating and Structuring Near-Synonyms. In: Gries, St. & Stefanowitsch, A. (eds.) Corpora in cognitive linguistics. Corpus-based Approaches to Syntax and Lexis. Berlin-New York: Mouton de Gruyter, 19-56. [Trends in Linguistics]

Divjak, D. 2010 (in press). *Structuring the Lexicon. A Clustered Model for Near-Synonymy.* Berlin/New York: Mouton de Gruyter. Cognitive Linguistics Research [CLR] 43.

Divjak, D. & S. Th. Gries. 2006. Ways of Trying in Russian: Clustering Behavioral Profiles. *Corpus Linguistics and Linguistic Theory*, 2 (1): 23–60.

Flint, A. 1980. *Semantic Structure in the Finnish Lexicon: Verbs of Possibility and Sufficiency*. Helsinki: Suomalaisen Kirjallisuuden Seura (SKST 360).

Hanks, P. 1996. Contextual Dependency and Lexical Sets. *International Journal of Corpus Linguistics*, 1:1, 75-98.

Harrell, F. 2001. *Regression Modeling Strategies. With Applications to Linear Models, Logistic Regression and Survival Analysis*. New York: Springer-Verlag.

Kangasniemi, H. 1992. *Modal Expressions in Finnish*. Studia Fennica, Linguistica 2. Helsinki: Suomalaisen Kirjallisuuden Seura.

Landauer, T. & S. Dumais. 1997. A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psych. Review*, 104 (2): 211–240.

Miller, G. A. 1990. Nouns in WordNet: a lexical inheritance system. (revised August 1993). *International Journal of Lexicography*, 3:4, 245–264.

Pajunen, A. 2001. *Argumenttirakenne: Asiaintilojen luokitus ja verbien käyttäytyminen suomen kielessä*. Suomi 187. Helsinki: Suomalaisen Kirjallisuuden Seura.

Pulman, S. G. 1983. *Word Meaning and Belief.* London: Croom Helm.

Rifkin, R. & A. Klautau. 2004. In Defense of One-Vs-All Classification. *Journal of Machine Learning Research*, 101-141.

Rosch, E. 1973. Natural Categories. *Cognitive Psychology* 4: 328–350.

Saffran, Jenny R, Richard N. Aslin and Elissa L. Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274: 1926-1928.

Tapanainen, P. & T. Järvinen. 1997. A non-projective dependency parser. *Proceedings of the 5th Conference on Applied Natural Language Processing*. Assoc. of Computational Linguistics, 64–71.

Taylor, J. R. 1995. *Linguistic Categorization: Prototypes in Linguistic Theory. 2nd Edition.* Oxford: Clarendon Press.

Theil, Henri 1970. On the Estimation of Relationships Involving Qualitative Variables. *The American Journal of Sociology*, 76:1, 03-154.

Usage frequencies of complement-taking verbs in Spanish

Paola E. Dussias[1 2], Alejandra Marful[3], Chip Gerfen[1 2], María Teresa Bajo Molina[3]
[1]Penn State University, [2]Center for Language Science, [3]Universidad de Granada

Abstract

Verb bias—or the tendency of a verb to appear with a certain type of complement—has been employed in psycholinguistic literature as a tool to test competing models of sentence processing. To date, the vast majority of sentence processing research involving verb bias has been conducted predominantly with monolingual English speakers. To test the generality of competing theories of sentence comprehension, it is important to conduct cross-linguistic studies of sentence processing and to add data from other languages to theories of sentence comprehension. Given this, it is critical for the field to develop verb bias estimates from speakers of languages other than English. Here, we report the results of a norming study for 135 Spanish verbs. One important goal of this study was to determine whether verb bias estimates remain stable over time, a question that to our knowledge has not been investigated. Our results demonstrate that individual verbs show significant similarities in their verb bias across the three years of data collection. To facilitate cross-linguistic work, we compare our verb bias results with those provided by monolingual English speakers in a previous norming study conducted by Garnsey, Lotocky, Pearlmutter, & Myers (1997).

1 Introduction

Verb bias norms have been critical for conducting studies in which the predictions of various models of sentence processing are tested. However, the interpretation of the results of such studies has been contentious. Findings that have been taken to reflect early influence of lexical information on syntactic decisions can also be explained in terms of reanalysis processes. This scenario is complicated by the recent evidence suggesting that different statistical analyses produce competing results (Kennison, 2009). To determine which sentence processing model best characterizes the cognitive architecture that underlies the ability to construct syntactic representations in real time, we need to combine existing theoretically sophisticated experimentation with on-line methods and statistical analyses that allow us to unambiguously distinguish earlier- from later-stage processes of syntactic parsing. Until such methods are developed, converging evidence from cross-linguistic studies of monolingual sentence processing are critical for the construction of models of syntactic processing and for empirically testing the claims of each model. Up to now, such testing has come primarily from studies with monolingual speakers, and predominantly from studies with monolingual English speakers. Because verb bias provides a crucial testing ground for competing theories of sentence processing, it is critical for the field to develop verb bias norms in numerous languages. As our knowledge about the factors that modulate syntactic parsing expands, we need an increasingly rich set of norms in order to probe the emerging theoretical questions raised by the different models of sentence processing.

With this in mind, we report the results of a norming study in which the usage frequencies of 135 Spanish verbs were collected. Using verb bias data derived from normative studies to test competing models of sentence comprehension presupposes that norming results are consistent across time. To our knowledge, no study to date has explicitly addressed this question. Ensuring that verb specific biases are robust is important because past studies have shown the verb bias estimates are affected by a number of variables, including the method used to gather the data (e.g., sentence completion tasks or corpus-based approaches), the method used to compute verb bias (e.g., absolute frequency or relative frequency), and the specific senses of a verb (e.g., when 'conclude' is followed by a direct object, it usually means 'to bring to an end.' However, if followed by a subordinate clause, it can mean 'to arrive at an end by reasoning') (for an extensive discussion, see Gahl et al., 2004; Hare, McRae & Elman, 2004). In the present study, we examine the stability of verb bias estimates over time by collecting Spanish verb norms across three years.

2. Method

2.1 Participants

A total of 575 monolingual speakers of Peninsular Spanish participated in the norming study. Participants were recruited over the course of three years. They were students in the Department of Psychology at the University of Granada (Spain) and received course credit for their participation. All participants reported having minimal or no knowledge of a second language. 2

3

2.2 Materials

Eighty-one verbs were selected from a list of 100 English verbs used in a norming study conducted by Garnsey, Lotocky, Pearlmutter, & Myers (1997). The 81 verbs were translated into Spanish using the Collins Dictionary of Español-Inglés/English-Spanish (2000) and the resulting translations were subsequently verified by a Spanish-English translator. An additional 54 Spanish cognate verbs, selected from Nash's (1993) Spanish-English cognate dictionary, were also normed. Using these 135 verbs (81 + 54), two lists were created. List 1 contained the 81 target verbs and 49 fillers (e.g., dative verbs, intransitive verbs, and verbs that subcategorized for prepositional phrases) of similar length and number of syllables. Fillers were included in order to discourage participants from limiting their completions. List 2 included the 54 cognate verbs (useful to conduct research with bilinguals, given that cognates have a special status in the bilingual lexicon) and a subset of 46 verbs from the 81 verbs included in List 1. Because List 1 and List 2 would be administered to different groups of monolingual Spanish participants, the 46 verbs in List 2 were included to check for consistency in the participants' responses between the two lists.

Twenty randomized files were created, 10 for each list. Each file contained the target and filler verbs in their past tense form embedded in a sentence fragment headed by a subject (always a proper name).

2.3 Procedure

Usage frequencies were obtained using the sentence completion task described in Garnsey, Pearlmutter, Myers, & Lotocky (1997). Participants were instructed to read a sentence fragment silently and to fill in a completion by hand next to the corresponding verb. They were told that there were no constraints on the length of their completions and that the resulting sentence needed to be grammatically correct and semantically plausible. No other instructions were provided.

Data collection took place over the course of three years. Data for List 1 were collected twice during Year 1 and twice during Year 2. List 1 was administered to a total of 464 Spanish monolingual speakers. List 2 was administered to an additional group of 111 monolingual Spanish speakers during Year 3.

3. Results

Participants' responses to a verb were coded in three categories: Direct Object (DO) completions, Sentential Complement (SC) completions and Other completions. This last category included prepositional phrases, infinitivals, and completions headed by relative pronouns such as *lo que* (that which). For our analyses, we focused mainly on the DO/SC classification because of its theoretical importance in current sentence processing literature. Average responses in each category were computed. It was often the case that participants failed to provide a completion for a given verb, particularly if it occurred towards the bottom of the list. For these cases as well as for cases in which the responses were illegible, the trial was coded as missing, and the number of participants included for the particular verb was reduced by 1.

Because norms for the 81 Spanish verbs List 1 were collected at different times, Pearson r correlations were computed between DO average responses and between SC responses to determine whether the completions provided for each verb were consistent across time. Results showed a significant and positive correlation ($p < .0001$). In addition, we conducted a second correlation analysis that compared responses to the 46 verb entries that were common to List 1 and List 2. When responses were compared across the three different years in which the data were collected (i.e. Year 1 and Year 2 for List 1 and Year 3 for List 2), the results again showed a significant and positive correlation ($p < .0001$). Taken together, the findings suggest that participants' responses were highly consistent across time. Therefore, in subsequent analyses, we collapsed the responses for each verb.

Following a criteria frequently used in psycholinguistic verb bias studies (Trueswell et al., 1993; Garnsey, Pearlmutter, Myers, & Lotocky, 1997), a verb was classified as DO-bias if it was used at least twice as often with a direct object completion as with a sentential complement completion and as SC-bias if there were at least twice as many sentential complement completions as direct object ones. Verbs were classified as EQ-bias if the difference between DO and SC completions was not greater than 15%. Remaining verbs were classified as No Bias. We chose a coding method that relied on relative frequencies (e.g., Garnsey, Lotocky, Pearlmutter, & Myers, 1997; Trueswell et al., 1993) rather than absolute frequencies (e.g., Merlo, 1994) for two reasons. First, we wanted to generate a set of verb bias norms in Spanish that would be comparable to existing English norms. Second, Gahl et al. (2004, Experiment 4) presented evidence suggesting that only the relative criterion for classifying verb bias could account for some of the results reported in a number of psycholinguistic studies. Therefore, we opted for coding our verbs using the relative method.

From the total 135 verbs normed, 50% were DO-bias, 23% were EQ-bias, 16% were SC-bias, and 11% were No-bias. Correlation analyses comparing the results of the 81 Spanish verbs to the equivalent English translations from the Garnsey et al. norms were positively significant (DO average, $r = .44$, $p < .0001$ and SC average, $r = .41$, $p < .001$). Establishing whether there are cross-linguistic differences in verb bias, particularly between SC-biased and DO-biased verbs in Spanish and English, is critical not only for conducting cross-linguistic studies, but also for identifying how the lexicon and grammar of a bilingual's two languages produce mutual influences and how competition between the two linguistic systems is resolved. Hence, a 2 (Language: Spanish vs. English) x 2 (Type: DO vs. SC) ANOVA on the average responses was computed. Neither a main effect of language ($F < 1$) nor an interaction between language and type ($F < 1$) were found. However, the results showed a main effect of type [$F(1,80) = 10.453$, $p < .01$] such that DO responses ($M = .46$, $SD = 32$) were more frequent than SC responses ($M = .31$, $SD = 23$), as shown by comparing responses in Spanish [$t(80) = 2.72$, $p < .01$], in English [$t(80) = 2.81$, $p < .01$], and when comparing Spanish to English [$t(80) = 2.88$, $p < .01$] and English to Spanish [$t(80) = - 3.15$ $p < .01$].

When bias alone was taken into account, 49% of the verbs showed a different bias in Spanish and English. However, the vast majority of these involved cases in which a verb changed between EQ and either DO or SC bias. Reverses in bias between Spanish and English also occurred, but were considerably fewer.

4. Conclusion

A significant feature of our findings is the correlation in participants' responses across the different times data were collected. The implication of this finding for theory-building is encouraging because a highly reliable set of norms strengthens experimental findings relevant to theoretical issues being debated in the sentence processing literature. Another important result is the distribution of verb subcategorization frequencies between Spanish and English--only about half of the verbs normed share bias in the two languages. The availability of a corpus containing verbs with same and different bias will be of use to scholars who wish to conduct cross-linguistic studies of sentence processing. It will also enable researchers interested in bilingualism to examine lexical effects on sentence processing when bilinguals read in each of their two languages.

References

Gahl, S., Jurafsky, D., & Roland, D. (2004). Verb subcategorization frequencies: American English corpus data, methodological studies, and cross-corpus comparisons. *Behavior Research Methods, Instruments, & Computers*, **36**, 432-443.

Garnsey, S. M., Lotocky, M. A., Pearlmutter, N. J., & Myers, E. M. (1997). *Argument structure frequency biases for 100 sentence-complement-taking verbs.* (Unpublished manuscript, University of Illinois at Urbana-Champaign).

Hare, M., McRae, K., & Elman, J. L. (2003). Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Lan*guage, **48**, 281-303.

Kennison, S. M. (2009). The use of verb information in parsing: Different statistical analyses lead to contradictory conclusions. *Journal of Psycholinguistic Research*, **38**, 363-378

Merlo, P. (1994). A corpus-based analysis of verb continua tion frequencies for syntactic processing. *Journal of Psycholinguistic Research*, **4**, 435-447.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence-processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 528-553.

Semantics vs. Syntax in Verb Structure Acquisition

Polina Eysmont
State University of Aerospace Instrumentation
polina272@hotmail.com

Abstract

The paper deals with the problem of verb argument acquisition and verb realization in the speech of children and adults, which is one of the most discussed, but still vaguely understood problems in Russian grammar. In their reports both children and adults used about 800 verbs, that were classified according to their semantics and their argument structure. The paper discusses the results of the analysis of adults' and children's usage in all different types of verb argument structures. It proves that both groups tend to fulfill all possible arguments of verbs of speech. However, it is only subject position that is usually fulfilled in two-participant verbs of motion while the second argument position is often omitted. It also shows that semantics are acquired earlier than syntactic rules, even in relation to simple sentence structure, and that children acquire syntactic structure by analogy with other verbs of the same semantics.

1 Introduction

The aim of the paper is to discuss some problems relating to the acquisition of verb argument structure and their representation in children's narratives in comparison with adults in connection with verb semantic classification. The history of verb argument studies and their acquisition is very long, but it should be noticed that it is mostly in connection with the earliest periods of language acquisition, when children are not expected to produce any long monologues. At the same time we should not forget that narrative as a coherent account of some sequence of events provides usage of special sentence structures and the omission of different arguments.

It should also be mentioned that the study of Russian language acquisition has unfortunately not paid sufficient attention to the problem, mostly discussing different morphological peculiarities of Russian language acquisition.

Nevertheless, it has been shown that children acquire 2-argument verbs with subject and object by the age of 2 (Özçalişkan, Goldin-Meadow 2005; Lidz, Gleitman 2004) and children tend to omit subjects and objects, especially in pro-drop languages such as Italian or Russian (Schmitz, Müller 2003; Tseitlin 2000). The acquisition of verb-argument structures depends on their semantics and are acquired generally by analogy with other verbs of the same semantics (motion, speech, perception, etc.) (Gropen, Pinker, Hollander & Goldberg 1991; Tseitlin 2009).

2 Material

The paper represents some conclusions which are based on the data of two similar experiments – one with adults as the subjects, and the other – with children. A 4-minute cartoon was chosen to elicit verbal reports of the Ss. The adult experiment was run in three sessions with 2 month intervals in order to give subjects enough time to forget the plot of this cartoon and their own narratives. The experiment with children was run in two sessions with 2 month intervals. The same group of Ss. participated in all experimental sessions: 20 students, aged 19-25, and 17 children, aged 7-8. 94 original narratives have been recorded, and the total duration of all the records is about 6,5 hours.

In their narratives, the subjects used 788 verbs in total, while children used only 245. All these verbs were classified according to their argument characteristics. A 1- argument verb may have 3 slots, namely the verb itself, its only actant and a circonstant[1], a 2- argument verb may have 4 slots: the same plus the second actant; a 3-argument verb – 5: 3 actants, the verb and a circonstant[2].

If one of the actant slots is left empty, the utterance is elliptic. In the research two types of

[1] In the analysis we count one slot for circonstants regardless of their number.

[2] In the analysis first actant is indicated as S in order to distinguish it from other actant and because it usually coincides with syntactic subject.

elliptic structures are distinguished. *Ellipsis* is an incomplete syntactic structure in which it is possible to restore the deleted items, thus recovering the original meaning and filling the gap in the syntactic structure. There are rather semantic or pragmatic reasons for omitting this component of the structure (for example: *and X begins to destroy them / and X throws at / now X is throwing a ball,* where *X* stands for omitted subject).

The other type is a *reduction* where the restoration of the gap, relying on the semantic and syntactic context, is impossible or may lead to a different meaning. This is because there are restrictions on a syntactic manifestation of a semantic argument, i.e. an element is omitted for purely syntactic reasons and is determined by grammar (for example, infinitives or impersonal sentences in Russian).

3 Discussion

The paper discusses different semantic groups of verbs as examples of coordination between semantic and syntactic argument structures. The first group represents the verbs of speech. The verbs of this group were not frequent either in adults' or in children's texts. Among all 55 verbs of speech 52% of the total were 3-argument verbs, whilst 2-argument verbs constituted 40% of the total. As for their syntactic structures, the 3-argument verbs are of two types: nominative, prepositional, prepositional instrumental (*besedovat'* 'talk to smbd about smth') and nominative, accusative, different prepositional cases (*obsuzhdat'* 'discuss smth with smbd').

The distribution of these argument structures in children's and adults' texts is presented in the diagram (Figure 1):

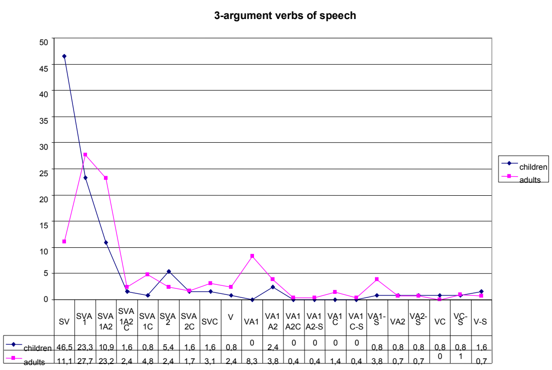


Figure 1. Distribution of argument structures in verbs of speech[3].

[3] Unfortunately, the article size does not allow to elaborate every argument structure here, see Eysmont 2008.

The data show, that both children and adults have used almost all possible structures, but children generally tended to fill the first position of the subject – the subject position is filled in 91,5% of all utterances, while adults filled it in only 76,5%. As for optional valences, children tended to omit them as well: more than 80% of all utterances were such structures as SV, SVA_1 and SVA_1A_2, while the same structures in adults' texts were only about 60%. So children, even using 3-argument verbs in their narratives tend to fill only 1 or two possible arguments. If we look at their syntactic representations, we'll see that children fill the second argument position with direct speech, whilst adults tend to reformulate it in indirect clauses or its nominalization. So, we may conclude that children have already acquired the semantic structure of a 3-argument verb and tend to fill most semantic valences of a speech event, but still drop out the syntactic ones.

The other group to be discussed in the paper is the group of verbs of motion. This group was the most frequent, and all of these verbs are 2-argument verbs, where second position refers to either destination or start point of the motion, and is to be filled with different prepositional cases or an adverb.

The distribution of argument structures in children's and adults' texts is presented in the following diagram (Figure 2):

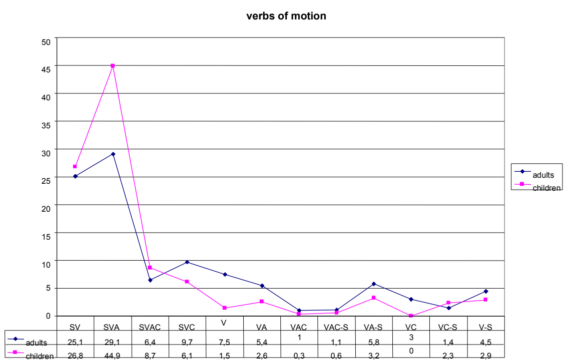


Figure 2. Distribution of argument structures in verbs of motion.

The diagram shows, that all possible structures have been used by both children and adults in their narratives and proves, that children have already acquired verbs of that kind and feel free using them in their speech. But at the same time, you may notice that the most frequent structure in children's narratives was the structure with both positions – subject and object – filled, while in the verbs of speech this structure followed the simpler structure with only one – subject position filled. It also demonstrates that children are free to use 2-argument verbs as easy as adults use 3-argument verbs.

So high number of verbs of motion in children's narratives may also be explained as a result of their semantic and syntactic structure: they are 2-argument verbs, and children acquire them by the age of 2 (Özçalişkan, Goldin-Meadow 2005; Lidz, Gleitman 2004). If so, we may suggest that although children are obvious to acquire syntactic structure on the analogy of other verbs of the same semantics, in some period of language development it begins to work in an opposite way: children tend to use in their speech verbs of the same syntactic structure as the verbs acquired before. In other words, having acquired a new syntactic structure, children try to apply it as often as it is possible.

As for syntactic structures, this analysis proves that there is no distinction in use between different 2-argument verbs either they are prototypical or non-prototypical, and their usage does not depend on their syntactic structures. I should also mention that there was no single mistake within the noun cases in children's narratives, although the previous studies by Tseitlin showed quite a big percentage of cases misuse in schoolchildren's narratives (2009).

4 Conclusion

The analysis of adults' narratives proved that adults use all possible types of argument structure – both semantic and syntactic, but tend to omit 'everything that is possible to be omitted', and use elliptic utterances as often as reduced ones. At the same time children tend to use complete semantic argument structures and do not omit subjects for elliptical reasons, but use quite a lot of reduced structures. Children fill most of subject and first object semantic positions independently of the number of verb arguments, but in 3-argument verbs they do it in a different way, and avoid using syntactically required structures (cf., direct vs. indirect speech). Verbs of motion are the most frequent, whereas verbs of speech are much more infrequently observed. It should be mentioned, that this situation cannot be interpreted as verbs of motion referring to concrete actions, which are known to be more important for little children. This is because the group which follows verbs of motion is that of verbs of perception, which are probably even more abstract than verbs of speech.

The analysis has also proved that semantics are acquired earlier than syntactic rules, even in relation to simple sentence structure. This result seems to be quite understandable and predictable, as the semantics of verbs and especially their argument structures represent the structures of real events and their perception does not require any specific language skills.

Children tend to fill subject position. This result was quite unexpected for two reasons. The first is the fact that Russian children start speaking by producing mostly elliptic structures, which they may be considered to have already acquired by the age of 7. The second reason is that Russian is a pro-drop language, and adults generally prove it in their narratives. So, these results probably demonstrate a kind of a non-pro-drop period in acquisition which has not been studied before.

The last point is that children, as have been shown acquire syntactic structure by the analogy of other verbs with the same semantics. Having studied this new syntactic structure they are able to use it widely in their speech with all the appropriate verbs.

References

Jeffry Lidz & Lila Gleitman. 2004. *Argument structure and the child's contribution to language learning. Trends in Cognitive Sciences*, 8(4): 157-161.

Jess Gropen, Steven Pinker, Michelle Hollander and Richard Goldberg. 1991. *Affectedness and Direct Objects: The role of Lexical Semantics in the Acquisition of Verb Argument Structure. Cognition, 4*: 153-195.

Katrin Schmitz & Natascha Müller. 2003. *Strong and clitic pronouns in monolingual and bilingual first language acquisition: comparing French and Italian. Arbeiten zur Mehrsprachigkeit 49*: 1-38.

Polina Eysmont. 2008. *Interconnection between information and argument structures in child narrative. Problems of speech production and speech perception*. Tcherepovets.

Şeyda Özçalişkan & Susan Goldin-Meadow. 2005 *Do parents lead their children by the hand? Journal of Child Language*. Vol. 32: 481-505.

Stella N. Tseitlin. 2000. *Language and a child. Linguistics of child language*. Saint-Petersburg.

Stella N. Tseitlin. 2009. *Speech errors and their prevention*. Moscow, URSS.

(Light) Verb Retrieval in a Case of Primary Progressive Aphasia

Franco Ludovico
Ca' Foscari University of Venice
franco.ludovico@gmail.com

Zampieri Elisa
Ca' Foscari University of Venice
zampy81@yahoo.it

Garzon Martina
IRCCS Ospedale S. Camillo, Lido di Venezia, Italy
martina.garzon@gmail.com

Meneghello Francesca
IRCCS Ospedale S. Camillo, Lido di Venezia, Italy
npsi@ospedalesancamillo.net

Cardinaletti Anna
Ca' Foscari University of Venice
cardin@unive.it

Semenza Carlo
Department of Neuroscienze, Università di Padova
carlo.semenza@unipd.it

Abstract

In this work we present a study of verb syntax in a case of Primary Progressive Aphasia, trying to empirically check two radical claims about Lexicon and syntax: a) verbs are a closed class of *light verbs* (Kayne, 2009); b) argument structure is a matter of syntactically driven operations (incorporation / conflation), in a constructionalist fashion (Hale and Keyser, 1993, 2002).

1 Introduction

In this paper, we present a case of Primary Progressive Aphasia (PPA), which seems to support the idea of verbs as a "closed class" recently proposed by Kayne (2009) - drawing inspiration from Hale and Keyser (1993; 2002) - in a paper in which he addresses many questions concerning antisymmetry and the Lexicon. PPA is a degenerative syndrome marked by progressive deterioration of language functions and relative preservation of other cognitive domain[1], firstly investigated by Mesulam (1982). On the basis of the nature of language impairment, patients with PPA have been subdivided into semantic, agrammatic/dysfluent and logopenic subtypes (Mesulam *et al.*, 2009). Following this classification, the semantic variant is characterized by poor single word comprehension but relatively well-preserved fluency and syntax; the agrammatic variant by poor syntax and fluency but relatively preserved word comprehension; and the logopenic subtype by preserved syntax and comprehension but variable fluency.

[1] Functional neuro-imaging studies on PPA have shown abnormalities mostly in the left anterior and posterior temporal lobe, with reduced language-related activations in Broca's and Wemicke's areas, and increased activations of the left posterior frontal cortex and right hemispheric regions (cf. Sonty *et. al.*, 2001).

2 Aim of the work

In order to empirically investigate Kayne's (2009) claim about a Lexicon in which only nouns can be considered as primitives and to test the proposal of an argumental structure without thematic roles as primitives, which derives thematic interpretation from syntactic position (*akin to* Hale & Keyser 1993, 2002), we ideally need the "mirror image" of an agrammatic speaker[2], namely a subject who has well-preserved functional morphology and, on the other hand, a deep anomia, affecting her lexicon. In other words, we need a subject that could trigger a sort of "transparency effect" in order to provide evidence that the noun *vs.* verb processing dissociation in aphasia[3] (an inflated topic in the literature) can be addressed, starting from the hypothesis that nouns are primitives, while verbs are a "syntactic product[4]", derived by incorporation of nouns into

[2] One of the hallmarks of agrammatic-type Broca's aphasia is a deficit in the production of functional morphology. Both free-standing function words and bound morphemes used to mark grammatical functions are impaired in this population, crosslinguistically (see Avrutin, 2000 and references cited there; see also the groundbreaking works of Berndt and Caramazza, 1978 and Goodglass, 1976).

[3] Several researchers have shown that verb production is more impaired than noun production in individuals with agrammatic aphasia (see Luzzatti and Chierchia, 2002 and references cited there; also see Miceli *et al.* 1989).

[4] Another possibility, pursued in the Distributed Morphology (DM) framework (Halle and Marantz, 1993) is to assume an "Underspecification" of lexical items, which basically prevents items of the Lexicon from needing to be fully specified relative to their contexts of occurrence. From a psycholinguistic viewpoint, Barner and Bale (2002), adopting DM paradigm, argue that a theory without prefixed lexical categories can provide a better account of creative language use and category-specific neurological defi-

verb positions[5]. An approximation of this ideal subject has been found in a patient affected by non-fluent PPA. Her grammatical features (e.g. tense and agreement markers) are well preserved, so that her language production appears to be almost exclusively damaged by severe anomia.
A preliminary probe to check if we are on the right track should come from raw hints in previous researches within the neurolinguistic literature. Our idea, in fact, leads to the following approximate consequence: agrammatic Broca's aphasics should be impaired on semantically light (functional) verbs, while pathologies which have anomia as one of their most salient feature (say, for example, Alzheimer disease) should lead to transparency effects in the Lexicon, relying on an increased rate of complex predicate/light verb constructions.
For Broca's agrammatism, for instance, a recent study by Barde *et al.* (2006) has detected greater difficulty producing verbs that have fewer semantic components (namely, *light verbs*) compared to verbs that have greater semantic weight; conversely, the "semantic complexity" of verbs seems to affect Alzheimer disease, but not agrammatic patients' performance (Kim and Thompson, 2004). Hence, these data seem to support our hypothesis of selective differential impairments within the verbal category.
Shifting on a "*bioprogrammatic* perspective", which basically follows Bickerton (1984, and subsequent related works), our data should find endorsements in the field of *language acquisition* and *language creation*, labeling under the language creation process, the formation of *pidgins* & *creoles* (see DeGraff, 1999). Leaving aside the debate on Creole genesis (cf. Lefevbre, 1998) the interesting fact here is that creoles heavily rely on light verb constructions. A paradigmatic example is Sranan, a creole language spoken as a *lingua franca* by approximately 300,000 people in Suriname (see Essegbey, 2004), which makes extensive usage of serial light verb constructions (cf. on serial verbs Baker, 1989; Carsten, 2002; Collins, 1997, 2002; Aboh, 2009). Other examples, just to name a few include Saramaccan, a creole spoken by about 24.000 people near the Saramacca and upper Suriname Rivers in Suriname (Veenstra, 1996; cf. also Aboh, 2005), and many other Caribbean creoles (e.g Leeward creole of Antigua and Barbuda, Jamaican creole, etc., see Durrleman-Tame, 2008).
Other *hints* come from language acquisition. During the past few years, language acquisition researches have reported learners' use of semantically empty, "dummy" verbs (e.g. for English, Dutch, German, etc.), such as the verb form 'is' in the Dutch (ungrammatical) example "Hij is doorrijden" (*He is drive*) or the ungrammatical sentence with "do" "ik doe ook praten" (*I do also talk*). These constructions resemble English *do-support* constructions where '*do'* lacks a specific meaning (Evers and Van Kampen, 2001; Van Kampen, 1997; Radford, 1990; Roeper, 1992; Zuckerman, 2001; see also Bottari *et al.* 1993; Lightfoot, 1999).
Other possible suggestions can come from *language contact*. For instance, interesting evidence comes from loan words in a typological perspective. Recent investigations have found out that cross-linguistically a (wide-spread) strategy to absorb loan words into a "native" Lexicon is a special derivation process involving a light verb to accommodate the item that has been borrowed (see, for example, Wichmann and Wohlgemuth, 2008).
Independently from the perspective that can be adopted to explain the light verb spreading in the contexts of language acquisition, language creation and language contact (parameter setting, underspecification, pragmatically based accounts, etc.), the facts outlined above let us think that the "light verb" issue is a matter worth being investigated within neurolinguistics, with special regard to language pathology.

2.1 Case Study

Our patient (BB) is a 59 right-handed Italian woman with 17 years of scholarity, tested for three months (April-July 2009), 2.3 years after the onset. Standard tests (B.A.D.A., AAT) showed no differences in her production of nouns vs. verbs, both highly impaired. Previous works on PPA reported greater impairment in the naming of verbs than nouns (Hillis *et al.* 2006). Another study, however, found no evidence of reduced verb production (Graham *et al.* 2004). It was noticed that PPA patients use a verbal vocabulary that is somewhat less specific than normal speakers, with a larger use of so called "light verbs" (Graham and Rochon, 2007).

cits, while also offering a natural solution to the "bootstrapping" problem in language acquisition.

[5] Hale and Keyser (1993) claim that lexical verbs derive from a syntactic operation of incorporation/conflation of nouns into verbal slots.

3 Ligh verbs and the "verb as a closed class hypothesis" in a typological perspective

While in many languages it has been observed that, for instance adjectives or adverbs can constitute a closed, often quite small class[6] of elements (Dixon, 2004), the claim that verbs can be a closed class may appear controversial. But, as observed in Cinque and Rizzi (2010: 58): "If Hale and Keyser's (1993) idea that most transitive and intransitive verbs are not primitive but result from the incorporation of a noun into a limited class of light/general purpose verbs ('do', 'give', 'take', 'put', 'hit', etc.), then even the class of primitive verbs may turn out to be closed and relatively small. This seems confirmed by the fact that some languages typically fail to incorporate the noun into the light verb so that most 'verbal meanings' are expressed as V+N periphrases".

In order to find out a possible definition, in the framework of LFG, Alsina, Bresnan and Sells (1997: 1) summarize these V+N "complex predicates" as: "predicates which are multi-headed; they are composed of more than one grammatical element (either morphemes or words), each of which contributes part of the information ordinarily associated with a head". Examples of languages in which verbs seem to be a closed (functional) class include Iranian languages, such as Persian and Kurdish, which rely almost exclusively on functional verb constructions. It has been argued that (simple) verbs in these languages form a closed class and most light verb/complex predicate constructions[7] do not have simple verb counterparts (Megerdoomian, 2002). A somewhat different instance of light verb construction is found in a number of Northern Australian languages[8]. Other examples of languages that adopt a "functional verbs" strategy are Urdu, Hindi (Butt, 1995), Amharic (Amberber, 2010) and some South-American languages (e.g. Mosetén (Sakel, 2007)). Given this theoretical and typological introduction, we can illustrate what we have found in our PPA patient.

[6] This fact is evident in those languages – such as, for example Yoruba (see Dixon, 2004; Cinque 2006) - in which adjectives cannot be used predicatively: the attributive only adjectives form a closed (functional) class of elements.

[7] Folli *et al.* (2005) have showed that Persian can be considered as a transparent instance of Hale & Keyser's "constructionalist" model.

[8] In these languages, the host element is not a nominal, but comes from an open class of "underived predicative elements", termed *coverbs* (Pawley, 2006).

4 Methods and materials

We collected five samples of BB spontaneous speech from March to July 2009 (1251 words altogether). The samples were recorded in a quiet room at San Camillo Hospital, Venice, at the presence of two examiners that BB knew very well. During BB speech production the examiners never interrupted her, excepting for some few words to encourage her when she seemed to be tired or frustrated.

Two people separately transcribed the recordings. The two transcriptions were compared, and the few controversial passages were listened by a third person who didn't know the previous results. Only if the third transcription tied in one of the two previous ones, so the passage was included in the transcription.

Four control subjects, two men and two women, have been involved in the experiment. They matched with the subject by age and years of instruction and they didn't have any physical, neurological or psychological problem.

We collected five samples of spontaneous speech from the control subjects too, and then we cut the samples in order to obtain approximately the same number produced by the patient. We faithfully transcribed the samples following the same method we had used for the patient.

All verbs occurrences were counted. By "occurrence" we mean every time a verb was necessary to avoid an ungrammatical sentence. In this way, also omissions were included in the number of occurrences. The repetition of a word was considered a single occurrence if it was used to express a single concept.

Then the number of verbs really produced by the subject was considered, and a percentage was obtained in relation with the number of words.

Finally, verbs were divided into three classes: lexical verbs, functional verbs and quasi functional verbs (for the hypothesis of the existence and relevance of quasi functional / semi-lexical categories refer to the works collected in Van Riemsdijk and Corver 2001). Following Cinque (2006) and Cardinaletti and Shlonsky (2004), we considered as functional verbs not only auxiliaries, but also volitional, modal and causative verbs and the light verb "fare". In particular, we separately counted the occurrences of the functional verb "fare" + NP, because BB quite systematically substituted unergative and transitive verbs by this "semantically lighter" construction.

Percentages of every verbal class were obtained in relation to the total number of verb produced by the subject.
Omissions were lso classified; we separately counted omissions of the entire VP (calculated on the number of occurrences), omissions of the functional verb, and omissions of the lexical part of a "compound" form (the last two were considered in relation to the number of verbs produced).
We also considered the type/token ratio relatively to lexical verbs, in order to establish whether BB's lexicon was poorer than which of the controls.

5 Results

Interestingly, the progressive erosion of the lexicon in this case of PPA left almost intact the functional domain above VP, assuming a "cartographic" architecture of this kind ([FPz [FPy [FPx [VP]]]]) (Cinque, 2006; Cardinaletti and Shlonsky, 2004).
If we compare BB to the control group, BB produced a satisfactory number of verbs in relation to her words production Nevertheless some crucial differences could be detected, compared to the control subjects, considering the classes of verb she used most often.
Firstly, BB deep anomia was confirmed by the low percentage of lexical verbs she produced in relation to the total number of verbs (12,7% vs 42,1% of the controls).
Moreover, the lexical part of "compound" forms in which the functional verb was preserved, was omitted in the 13,2% of cases. The control group omitted it only in the 0,7% of the times.
The patient however had no hesitation with volitional, modal, and causative verbs which we assume to be hosted in functional projections above VP[9], and which are the 40,2% of her verb production. Only 1,6% of errors/omissions, affecting this verbal class, has been detected in obligatory contexts.
A second challenging result is that unergative and transitive verbs (such as *spiegare*, to explain) have been quite systematically substituted, in their "heavy form", by a light-verb+N form (e.g. *fare una spiegazione*, to do an explanation).
BB used the construction FARE+NP in the 14,8% of times (including the contexts in which BB was unable to retrieve (omitted) the *paired* N; 6,4%). Control subjects hardly ever omitted the nominal part (we found only one case in the entire corpus) and used this construction the 5,4% of the times on the average.
A third striking fact is that unaccusative verbs are preserved (17,5% of correct distribution/retrieval), confirming previous neurolinguistic observations (Froud, 2006), about their (quasi)functional status (see also Cardinaletti and Giusti (2003); Zubizzareta and Oh (2007)).
Interestingly the control group produced the 13,4% of unaccusatives verbs on the average, showing that, not only BB has not difficulties in retrieving this kind of verbs, but also that she prefers to use them, instead of looking for a more specific one.
Hence, from a quantitative viewpoint a different ratio of performance between functional verbs (preserved) and lexical verbs (impaired) in a patient of PPA has been detected. Data from previous works (Kim and Thompson, 2004; Barde *et al.* 2006) seem to predict that similar investigations on broader populations of "purely" anomic patients may confirm our results. From a theoretical viewpoint, we can suggest that, if anomia (the salient feature of this case of PPA) affects lexical classes, and according to Kayne (2009), the only open (lexical) class is represented by nouns, the fact that anomia selectively spares functional (light) verbs and leads to the surface's retrieval of Hale and Keyser's L-syntax could be considered a neurolinguistical evidence that the noun-verb distinction may be understood as a consequence of antisymmetry[10]: verbs may be seen as a closed class (all functional, all light). The immediate retrieval of a light verb (e.g. *fare*: to do) is forced by anomia: BB uses the otherwise silent light verb to which nouns incorporate/conflate (e.g. adjoining to the light-verb or moving to a Spec-position related to the light-verb). Notice that the lexical items (nouns), as we have already seen above, don't easily resurface in BB speech (poor inventory, pauses, neologisms).
We will try to support this theoretically driven claim with further (and broader) experimental evidence, trying also to elaborate a model (roughly based on CoLFIS (*Corpus e Lessico di Frequenza dell'Italiano Scritto,* see Bertinetto *et al.* 2005) to account for possible frequency effects.

[9] No data concerning aspectual verbs emerged from BB spontaneous speech.

[10] In recent updates of his work, Kayne (2009) attributes different (opposed) configurational properties to nouns and verbs.

6 Conclusion

In this paper we have tried to empirically verify two radical claims about Lexicon & syntax: i) real verbs are only grammatical ones (Kayne, 2009); ii) argument structure is a matter of syntactically driven operations, in a constructionalist fashion (Hale and Keyser, 1993, 2002). We have find out that these interrelated hypotheses can be tested and successfully verified within clinical linguistics, more specifically working with anomic patients, whose syntax is not (heavily) impaired.

References

Enoch Aboh. 2009. Clause Structure and Verb Series. *Linguistic Inquiry*, 40: 1-33.

Enoch Aboh. 2005. Object shift, verb movement and verb reduplication. in Guglielmo Cinque and Richard Kayne (eds.). *The handbook of comparative syntax,* , 138-177. Oxford University Press, New York.

Alex Alsina, Joan Bresnan and Peter Sells (eds.). 1997. *Complex predicates*. CSLI, Stanford.

Sergey Avrutin. 2000. Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. In Yosef Grodzinsky, Lee Shapiro and David Swinney (eds). *Language and Brain: Representation and processing*. Academic Press. San Diego. CA.

Mengistu Amberber. 2010. The Structure of the Light Verb Construction in Amharic, In Mengistu Amberber, Brett Baker, and Mark Harvey (eds.) *Complex Predicates: Cross-linguistic Perspectives on Event Structure*, Cambridge University Press, Cambridge.

Baker, Mark. 1989. Object sharing and projection in serial verb constructions. *Linguistic Inquiry,* 20: 513-533.

Laura H.F. Barde Myrna F. Schwartz and Consuelo B. Boronat. 2006. Semantic weight and verb retrieval in aphasia. *Brain and Language*, 97: 266-278.

David Barner and Alan Bale. 2002. No nouns, no verbs: Psycholinguistic arguments in favor of lexical underspecification. *Lingua*, 112: 771-791.

Pier Marco Bertinetto, Cristina Burani, Alessandro Laudanna, Lucia Marconi, Daniela Ratti, Claudia Rolando, Anna Maria Thornton. 2005. Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS). http://linguistica.sns.it/CoLFIS/CoLFIS_home.htm

Derek Bickerton. 1984. The Language Bioprogram Hypothesis. *Behavioral and Brain Sciences*, 7: 173-188.

Piero Bottari, Paola Cipriani and Anna Maria Chilosi. 1993. Protosyntactic Devices in the Acquisition of Italian Free Morphology. *Language Acquisition*, 3:327-369.

Miriam Butt. 1995. *The Structure of Complex Predicates*, CSLI Publications, Stanford, CA.

Alfonso Caramazza and Rita Sloan Berndt. 1978. Semantic and Syntactic Processes in Aphasia: A Review of the Literature. *Psychological Bulletin*, 85: 898-918.

Anna Cardinaletti and Ur Shlonsky. 2004. Clitic Positions and Restructuring in Italian. *Linguistic Inquiry*, 35: 519-557.

Anna Cardinaletti and Giuliana Giusti. 2003. Motion Verbs as Functional Heads. In Christina Tortora, ed. *The Syntax of Italian Dialects*, Oxford University Press, Oxford: 31-49.

Carstens, Vicki. 2002. Antisymmetry and word order in serial constructions. *Language*, 78: 3-50.

Guglielmo Cinque. 2006. *Restructuring and Functional Heads.* Oxford University Press, Oxford.

Guglielmo Cinque and Luigi Rizzi. 2010. The cartography of syntactic structures. In Bernd Heine and Heiko Narrog (eds.), *Oxford Handbook of linguistic analysis*. Oxford University press, Oxford: 51-65.

Chris Collins. 1997. Argument sharing in serial verb constructions. *Linguistic Inquiry*, 28: 461-497.

Chris Collins. 2002. Multiple verb movement in ±Hoan. *Linguistic Inquiry,* 33:1 – 29.

Norbert Corver and Henk van Riemsdijk (eds.). 2001. *Semi-lexical Categories*, Mouton de Gruyter, Berlin.

Michel DeGraff, (ed.). 1999. *Language Creation and Language Change: Creolization, Diachrony and Development.* MIT Press, Cambridge, MA.

R. M. W. Dixon. 2004. Adjective classes in typological perspective. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.). *Adjective classes*. Oxford University Press, Oxford: 1-45.

Stephanie Durrleman-Tame. 2008. *The syntax of Jamaican Creole: a cartographic perspective.* John Benjamins, Amsterdam.

James Essegbey. 2004. Auxiliaries in serialising languages: on COME and GO verbs in Sranan and Ewe. *Lingua*, 114: 473-494.

Arnold Evers and Jacqueline van Kampen. 2001. E-language, I-language and the order of parameter setting. *UiL OTS Working Papers*.

Raffaella Folli, Heidi Harley and Simin Karimi. 2005. Determinants of event structure in Persian complex predicates. *Lingua* 115: 1365-1401.

Karen Froud. 2006. Unaccusativity as lexical argument reduction: evidence from aphasia. *Lingua*, 116: 1631-1650.

Harold Goodglass. 1976. Agrammatism. In H. Whitaker and H.A. Whitaker (Eds.), *Studies in neurolinguistics, Vol. 1*. Academic Press, New York.

Naida Graham, Karalyn Patterson and John Hodges. 2004. When more yields less: Speaking and writing deficits in nonfluent progressive aphasia. *Neurocase*, 10: 141–155.

Naida Graham and Elizabeth Rochon. 2007. Verb production in sentences by patients with non- fluent progressive aphasia. *Brain and Language*, 103: 69-70.

Kenneth Hale and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of grammatical relations. *The view from Building 20* in Kenneth Hale and Samuel Jay Keyser (eds.), MIT Press, Cambridge, MA.

Kenneth Hale and Samuel Jay Keyser. 2002. *Prolegomenon to a Theory Of Argument Structure,* MIT Press, Cambridge, MA.

Morris Halle and Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. *The view from Building 20,* in Kenneth Hale and Samuel Jay Keyser (eds.), MIT Press, Cambridge, MA.

Argye Hillis, Jennifer Heidler-Gary, Melissa Newhart, Shannon Chang, Lynda, Ken and Thomas Bak. 2006. Naming and comprehension in primary progressive aphasia: The influence of grammatical word class. *Aphasiology*, 20: 246–256.

Jacqueline van Kampen. 1997. *First Steps in Wh-movement,* Eburon, Delft.

Richard Kayne. 2009. Antisymmetry and the Lexicon., *Linguistic Variation Yearbook* 2008: 1–32.

Mikyong Kim and Cynthia K. Thompson. 2004. Verb deficits in Alzheimer's disease and agrammatism: Implications for lexical organization. *Brain and Language*, 88: 1-20

Claire Lefevbre. 1998. *Creole genesis and the acquisition of grammar: The case of Haitian Creole*. Cambridge University Press, Cambridge.

David Lightfoot. 1999. *The Development of Language: Acquisition, Change, and Evolution*. Blackwell, Malden, MA.

Claudio Luzzatti and Gennaro Chierchia. 2002. On the nature of selective deficits involving nouns and verbs. *Italian Journal of Linguistics*, 14: 43-71.

Karine Megerdoomian,. 2002. *Beyond Words and Phrases*. PhD Thesis, University of Southern California, Los Angeles.

Marsel Mesulam. 1982. Primary progressive aphasia without generalized dementia. *Annals of Neurology.* 11: 592–598.

Marsel Mesulam, Emily Rogalski, Christina Wieneke, Derin Cobia, Alfred Rademaker, Cynthia Thompson and Sandra Weintraub. 2009. Neurology of anomia in the semantic variant of primary progressive aphasia. *Brain*: 132; 2553–2565.

Gabriele Miceli, Maria Caterina Silveri, Cristina Romani, and Alfonso Caramazza. 1989. Variation in the pattern of omission and substitutions of grammatical morphemes in the spontaneous speech of so-called agrammatic patients. *Brain an Language,* 36: 447–492.

Andrew Pawley. 2006. Where have all the verbs gone? Remarks on the organisation of languages with small, closed verb classes. ms.

Andrew Radford. 1990. *Syntactic theory and the acquisition of English syntax*. Blackwell, Cambridge, MA.

Thomas Roeper. 1992. From the initial state to V2: Acquisition principles in action. In J.M. Meisel (Ed.), The acquisition of verb placement, Kluwer, Dordrecht: 333-371.

Jeanette Sakel,. 2007. The verbness markers of Mosetén complex predicates. Bernhard Wälchli and Matti Miestamo (eds.) *New trends in Typology*. Mouton de Gruyter, Berlin: 315-335.

Sreepadma Sonty, Marsel Mesulam, Cynthia Thompson, Nancy Johnson, Sandra Weintraub, Todd Parrish and Darren Gitelman. 2003. Primary progressive aphasia: PPA and the language network. *Annals of Neurology* 53: 35-49.

Veenstra, Tonjes. 1996. *Serial Verbs in Saramaccan: Predication and Creole Genesis*. HIL Dissertations, 17.

Søren Wichmann and Jan Wohlgemuth. 2008. Loan verbs in a typological perspective. In: ThomasStolz, Dik Bakker, and Rosa Salas Palomo (eds.), *Aspects of Language Contact*, Mouton de Gruyter, Berlin: 89-121.

Maria Luisa Zubizzareta and Eunjeong Oh. 2007. *On the Syntactic Composition of Manner of Motion*. MIT Press, Cambridge, MA.

Shalom Zuckerman. 2001. The Acquisition of "Optional" Movement. Phd dissertation, University of Groningen.

Semantic role annotation: From verb-specific roles to generalized semantic roles

José M. García-Miguel
University of Vigo
Vigo, Spain
gallego@uvigo.es

Gael Vaamonde
University of Vigo
Vigo, Spain
gaelv@uvigo.es

Abstract

This paper aims to present the semantic role annotation carried out on the ADESSE project, an online database with syntactic and semantic information for all the verbs and clauses in a corpus of Spanish. In ADESSE, several subsets of semantic roles have been taken into account, interrelated through different levels of generalization.

1 Introduction

To have at our disposal annotated corpus is an obvious necessity for descriptive or computational purposes. Nevertheless, in carrying out any annotation process, we are required to move between two poles: the consistency of the data and the granularity of the analysis. Undoubtedly, this divergence increases when we have to deal with semantics, and in particular, with semantic role annotation. A factor which plays an important a role on this discrepancy tend to be the procedure adopted: automatic versus manual. The first one ensures a more systematic but coarse-grained product (Gildea & Jurafsky, 2002); the second one allows more accuracy, but it must face greater complexities. From a different point of view, the users of a linguistic resource may need sometimes very broad categories ranging over a wide set of data, and others may more detailed distinctions. Like in other annotation task, also in semantic role annotation the starting point, the design and the intended users determine to a great extent the resulting product (Ellsworth et al. 2002). Nevertheless, there are also some attempts to define a standard based on some existing alternative approaches (cf. Petukhova & Bunt 2008). Some well-known projects of semantic role annotation haven taken different paths in their design: FrameNet (Fillmore et al. 2003) is designed as an ontology of situation types (frames) and participants in those situations (frame elements) [1]. PropBank (Palmer et al. 2005) has a verb-dependent model of description of semantic relations. In this project, arguments are numbered and defined depending on the valency potential of each particular verb sense. VerbNet (Kipper, 2006) approach to meaning is based in an extension of Levin(1993)'s verb classes.

Regarding Spanish language, the Spanish FrameNet[2] project (Subirats 2009) follows exactly the same methodology that the original. But other important resources and projects of semantic role annotation of Spanish corpora use a predefined set a semantic role labels irrespective of situation type. This is the case of AnCora (Martí *et al.*, 2007, Taulé *et al.*, 2008) , and SenSem (Castellón *et al.*, 2006).

In ADESSE, a linguistic resource for Spanish, an intermediary path has been taken trying to combine the specifics of verb-senses, like in PropBank, with some generalizations over process types or verb classes. Fine-grained annotation is achieved by appealing to different subsets of semantic roles, which arise as a result of different levels of generalization. The main design features of ADESSE have been described elsewhere (García-Miguel & Albertuz 2005, García-Miguel et al. 2010) and are briefly summarized in section 2. This paper aims to show a slightly more detailed description of the levels of semantic role annotation in ADESSE, and this is the purpose of section 3.

2 The ADESSE project

ADESSE (*Base de datos de Verbos, **A**lternancias de **D**iátesis y **E**squemas **S**intáctico-*

[1] http://framenet.icsi.berkeley.edu
[2] http://gemini.uab.es:9080/SFNsite/

S**emánticos del **E**spañol*)[3], a project being developed at the University of Vigo, is an online database providing detailed syntactic and semantic information about verbs and clauses from a Spanish corpus. ADESSE is an expanded version of BDS (B**ase de **D**atos **S**intácticos del español actual*), the syntactic analysis of a corpus of Spanish into a relational database. ADESSE takes a syntactically analyzed corpus to semantically annotate all and only the clauses in the corpus. In this respect, ADESSE is partly similar to a Treebank with syntactic and semantic annotation, although limited to argument structure. The manually annotated corpus has 1.5 million words, 159,000 clauses and 3,450 different verb lemmas. BDS contains grammatical features of verbs such as voice, tense and mood, and syntactic features of verb-arguments in the corpus, such as syntactic function, and phrase type. ADESSE has added semantic features such as verb sense, verb class and semantic role of arguments to make possible a detailed syntactic and semantic corpus-based characterization of verb valency. A fundamental goal of the project is to get a corpus-based description of verb valency in Spanish. The database includes, among other things, the syntactic function and the syntactic category for each core argument of each clause in the corpus, and semantic information about verb sense, semantic verb class for each verb sense, and semantic roles for each verb argument.

3 Semantic role annotation in ADESSE

Semantic annotation in ADESSE was primarily carried out for descriptive purposes, and follows always a bottom-up approach, starting from the data a trying to define a set of categories that can describe those data. This can explain why the cited project adopts a fine-grained annotation of semantic roles, compared with other similar resources for Spanish, like AnCora or SenSem. Unlike these projects, there is no just one set of roles for annotating arguments in ADESSE. Actually, we do not use any previous list of possible options. The strategy is an inductive one, taking verb meaning as the starting point and describing (types of) participants from each verb sense in an increasingly wide-ranging way. This strategy allows us to cover different levels of granularity and, at the same time, to establish generalizations about argument structure based on lexical verb meaning.

Taking all of this into account, role definition is made at three levels in ADESSE: verb-specific roles, class-specific roles, and generalized semantic roles.

3.1 Verb-specific roles

Verbs categorize types of situations and participants in those situations in a unique way, so at the extreme a distinct set of participant roles must be posited for each verb sense (cf. Langacker, 1991:284). Role definition in ADESSE is initially carried out on this maximally specific level. For each verb sense, we describe its valency potential, that is, the whole set of possible participants accepted with that verb, taking into account all the syntactic patterns recorded in the corpus (its valency realizations). The goal here is, on the one hand, to distinguish roles of participants co-occurring in the same syntactic pattern and, on the other, to trace equivalences between arguments of different syntactic patterns

For example, the verb *contar* 'to tell a happening' can be described by considering up to four arguments: A1: 'the one who tells something', A2: 'the thing told', A3: 'the one to whom something is told', and A4: 'the issue of what is told'. This allows us to describe examples like (1a), where the whole range of participants is expressed in a single clause, as well as (1b) or (1c), where only a subset of them is selected. (In these examples 1-2-3-4 stand for A1-A2-A3-A4)[4]:

(1) a. [1] <u>*Cuénta*</u>[*nos* 3] [*algo* 2] [*de Madrid* 4]
'Tell [_1] [us 3] [something 2] [about Madrid 4]'
b. [*El viejo* 1] <u>*cuenta*</u> [*su última treta* 2]
'[The old man 1] tells [his last ruse 2]'
c. *¡Ah, si* [*yo* 1] [*le* 3] <u>*contara*</u>*!*
'¡Oh, if [I 1] told [you 3]!'

The main problem in this process is to decide about the semantic equivalence between arguments of different syntactic patterns, and to decide if the examples are instances of the same verb sense. The general strategy has been to make as few verb sense distinctions as possible, reducing lexical entries are to a minimum.

Verb-specific description of semantic roles is also adopted in PropBank (Palmer et. al., 2005),

[3] http://adesse.uvigo.es/

[4] Note in passing that the database registers as arguments, not only full noun phrases and pronouns, but also clitics (*le*) and referents evoked by verb agreement like the A1 argument of (1a).

a project who aims to annotate a syntactically parsed corpus with information about argument structure. In this project, verbal arguments are labeled as numbered arguments, from Arg0 on.

Following the PropBank style, ADESSE also assigns a sequential number to each verbal argument: A0, A1, A2, … Nevertheless, there exist two important differences. The first one has to do with the scope of numbered arguments (we will turn to this question in section 3.3.). A second difference has to do with role labels. In PropBank, there is no semantic role label associated with each incrementally numbered argument, but only a brief description (generally, a formula of the type: 'V-er', 'thing V-ed') and, sometimes, the corresponding thematic role used in VerbNet (cf. Kipper et al., 2002).

In ADESSE, we usually do not suggest specific role labels on this level (but see Figure 2). If so, we would have to admit as many labels as existing slots for each verb recorded in the corpus[5]. However, our description of valency potential actually includes semantic role labels for each argument. In ADESSE, this information is directly inherited from the following more abstract level of representation, where types of situations and their corresponding types of participants must be considered.

3.2 Class-specific roles

Assuming that each situation is unique, the verbal lexicon of any language allow us to abstract commonalities from those partially different situations. With this idea in mind, one of the goals in ADESSE is to get a semantic classification of Spanish verbs by delimiting a set of possible conceptual classes or types of events. This is also a bottom-up process of grouping lexical entries. ADESSE's classification has an ontological basis and a hierarchical structure, with up to four levels at the present stage[6]. Each semantic class is associated with a set of semantic roles which are prototypical for the conceptual domain evoked, so that verbs belonging to the same class will share the same subset of semantic roles.

The conceptual basis adopted in ADESSE to characterize types of events and participants is reminiscent of FrameNet (Fillmore et al., 2003). However, there are important differences between both projects (García-Miguel & Albertuz 2004). ADESSE classes and subclasses are much more schematic than frames in FrameNet: the 63 verb classes of ADESSE (for approximately 4000 verb entries) cannot reflect the fine-grained distinctions of the more than 1000 frames defined in FrameNet. Nevertheless, FrameNet has frames at different levels of schematicity, and an elaborated system of inheritance relations between frames. More schematic frames, inherited or used by more specific ones, are most similar to ADESSE classes and subclasses.

Some of the labels used for these class-specific roles may fit with traditional thematic roles (e.g. agent, patient, instrument, location, etc.). Nevertheless, role labels in ADESSE where chosen by aiming at two factors: specificity (depending on the verbal class) and transparency (descriptive adequation). Some of them are stated in the following table:

Class	**A0**	**A1**	**A2**
Feeling		Emoter	Emoted
Perception	Causer	Perceiver	Perceived
Cognition	Causer	Cognizer	Content
Possession		Possessor	Possessed
Transfer	Donor	Final-poss.	Possessed
Change	Agent	Patient	

Table 1. Some class-specific roles in ADESSE

Verb-specific arguments inherit by default the labels from class-specific roles. For example, the valency potential of *prestar* 'to lend', which is classified as a verb of 'transfer', is semantically described by making reference to the set of roles associated with that class, that is: A0: 'Donor', A1: 'Final-Possessor', A2: 'Possessed' (see Figure 1). The same set of labels is used to semantically annotate the arguments of verbs like *dar* 'to give', *pagar* 'to pay', *vender* 'to sell', etc:

Transfer	Donor	Final-poss.	Possessed
↑	↓	↓	↓
prestar 'lend'	A0 lender	A1 borrower	A2 loan

Figure 1. Verb-specific roles of *prestar*, a verb of Transfer.

Up to now, ADESSE comprises a total of 196 class-specific roles spread over 63 different se-

[5] So far, there are 4,016 verb meanings and 9,758 verb-specific arguments in ADESSE, giving an average of 2,4 arguments per verb.

[6] The whole semantic classification can be consulted in http://adesse.uvigo.es/data/clases.php.

mantic classes[7]. Given that the semantic classification is hierarchical, with up to four levels of more general and more specific process types, class-specific roles allow us to cover and define types of participants at different levels of generalization. So, for example, the class of ‘change’ is subdivided in three subclasses: a) verbs of creation (e.g. *crear* ‘create’, *producir* ‘produce’), b) verbs of modification (*abrir* ‘open’, *romper* ‘break’), and c) verbs of destruction (*destruir* ‘destroy’, *eliminar* ‘erase’). Each subclass is associated with a different set of semantic roles: a) Creator and Creation, b.) Agent and Affected, c) Destroyer and Destroyed. But the more schematic class of ‘change’ neutralizes these semantic contrasts, abstracting the common properties of the mentioned roles into an Agent and a Patient. Likewise, the class 'Mental process' includes the classes Feeling, Perception, and Cognition so that the semantic roles Experiencer and Stimulus, associated to the Mental class must be seen as generalizations over the participant roles of the more specific process types. These and other similar cases of generalizations concerning class-specific roles are summarized in figure 2:

Verb-specific roles	**Class-specific roles**	
gustar.A1	Emoter	
ver.A1	Perceiver	Experiencer
saber.A1	Cognizer	
gustar.A2	Emoted	
ver.A2	Perceived	Stimulus
saber.A2	Content	
hacer.A1	Created	
romper.A1	Affected	Patient
destruir.A1	Destroyed	

Increasing generalization →

Figure 2. Semantic roles and levels of generalization

The set of relations between classes and class-specific roles in ADESSE is reminiscent of the network of inheritance relations between Frames and Frame Elements in FrameNet, although not as much fine-grained.

One might think that, by following this line of generalization, a maximally schematic level of representation could be achieved, so that we could get a limited set of semantic roles independently of process types.

As an equivalent of what is labeled ArgM in PropBank, we consider a small group of semantic roles for additional or secondary participants. These general roles (AG) are possible with verbs belonging to different semantic classes and allow to fully describe the valency potential of many verbs for which the inherited class-specific roles are not enough. The set labels used so far for these additional participants is: *Beneficiary, Location, Manner, Matter, Purpose, Reference, Attribute, Final State, Object, Means, Possessor, Facet, Company, Cause, Source, Role* .

However, for the more nuclear arguments, at the higher level of abstraction we must face a heterogeneous set of variables reflecting features of completely different semantic domains. Therefore, it is necessary to take into account the syntactic-semantic commonalities observed among the whole set of semantic roles.

3.3 Generalized semantic roles

There exist several linguistic theories which have dealt with a maximally schematic representation of argument linking (cf. Dowty 1991, Van Valin & LaPolla 1997, Croft 1998). Although different in many respects, all these proposals must be based on some kind of template or scale on which relative positions of arguments could be accounted for.

A usual way to do that is by starting from a logical decomposition of predicates based on Aktionsart distinctions, as proposed in RRG (cf. Van Valin & LaPolla, 1997). What these authors suggest is that all possible thematic relations can be summarized in only five distinctions, corresponding to the argument positions allowed by logical structure templates (Figure 3)[8]. As a result, a hierarchy is obtained from which two macro-roles are posited, Actor and Undergoer:

ACTOR → ← UNDERGOER

Arg of DO	1st arg of **do´** (x,…)	1st arg of **pred´** (x,y)	2nd of **pred´** (x,y)	Arg of **pred’** (x)

Figure 3. Actor-Undergoer hierarchy in RRG

Briefly, Actor macro-role fits with the subject of transitive and unergative verbs, while Under-

[7] An inventory much more bigger than the one used by AnCora (20 semantic role labels) or SenSem (32 semantic role labels)

[8] x and y are arguments of a predicate **pred'**. **do'** is a generalized activity predicate, and DO is a generalized causative predicate.

goer macro-role fits with the object of transitives and the subject of unaccusatives.

	Actor	Undergoer	[other]
KNOW	knower	thing known	
LEARN	learner	thing learned	
TEACH	teacher	thing learned learner	learner thing learned

Table 2. *Know, learn* and *teach* in RRG

A strategy based on correlative pointers to annotate predicate argument structures is used in PropBank: "An individual verb's semantic arguments are numbered, beginning with zero. For a particular verb, Arg0 is generally the argument exhibiting features of a Prototypical Agent (Dowty 1991), while Arg1 is a Prototypical Patient or Theme. No consistent generalizations can be made across verbs for the higher-numbered arguments, though an effort has been made to consistently define roles across members of VerbNet classes." (Palmer et al. 2005: 75). Therefore, in this project Arg0 is generally applied to the subject of transitive and unergative verbs, establishing similar correspondences to RRG (see Table 3).

	Arg0	Arg1	Arg2
KNOW	knower	thought	attributive
LEARN	learner	subject	teacher
TEACH	teacher	subject	learner

Table 3. *Know*, *learn* and *teach* in PropBank

Regarding ADESSE, we have already mentioned how verb arguments are incrementally numbered. However, beyond describing the valency potential of each verb, these numbered arguments can serve to represent generalizations from argument positions, in the way of variables in logical templates. In ADESSE, default pointers for arguments are chosen taking into account the following correspondences: A0=initiator or causer, A1=1st argument of **pred´**, A2=2nd argument of **pred´**. Schematically, we could trace the parallelisms between ADESSE hierarchy and the Actor-Undergoer hierarchy as follows:

A0	A1		A2
Arg of DO	1st arg of **do**´(x,…)	1st arg of **pred**´(x,y) or **pred**´(x)	2nd arg of **pred**´ (x,y)

Figure 4. ADESSE hierarchy versus Actor-Undergoer hierarchy

As can be deduced from Figure 4, in ADESSE A0 is reserved for the first argument of causatives, so that we can see more easily the correspondences between causatives and their non-causative counterpart (Table 4).

	A0	A1	A2
SABER 'know'		knower [Cognizer]	thought [Content]
APRENDER 'learn'		learner [Cognizer]	subject [Content]
ENSEÑAR 'teach'	teacher [Causer]	subject [Cognizer]	learner [Content]

Table 4. *Saber, aprender* & *enseñar* in ADESSE

That way, a greater coherence with lexical meaning and lexical relations is achieved, while linking of semantics and syntax is understood in terms of relative positions in the argument scale. As can be seen in Table 5, Subject is almost always higher than DObj in the hierarchy of GSRs

Subj - DObj (+ oblique) in Active Voice		
Subj=**A1**	DObj=**A2**	61%
Subj=**A0**	DObj=**A1**	25 %
Subj=**A0**	DObj=**A2**	3 %
Other		10%

Table 5. Linking of grammatical relations and arguments. Frequency in ADESSE

4. Conclusion

We have outlined a system for describing semantic roles at different levels of granularity. About 326K arguments of 159K clauses have been given annotation at one or more levels in the database. The frequency of each role index is given in Table 6.

index	more common class-specific role labels	N
A0	Causer, Agent, Donor, Assigner, …	31521
A1	Theme, Cognizer, Communicator, Perceiver, Affected, Possessor, …	156958
A2	Content, Perceived, Possessed, …	103103
A3	Goal, Addressee, Perceived-2, …	16414
A4/A5	Path, Content-2, Activity, Code, …	4566
AG	Beneficiary, Location, Reference, ..	13312

Table 6. Frequency of arguments in ADESSE

Acknowledgements

This work has been supported by the Spanish Ministry of Science and Innovation (BFF2002-01197, HUM2005-01573, FFI2008-01953) the Galician Government (PGIDIT03PXIC-30201PN), and the University of Vigo.

References

Irene Castellón, Ana Fernández-Montraveta, Gloria Vázquez, Laura Alonso Alemany & Joan Antoni Capilla. 2006. The SenSem Corpus: A Corpus Annotated at the Syntactic and Semantic Level. *Fifth International Conference on LREC*, 355-359

William Croft. 1998. Event Structure in Argument Linking. In Miriam Butt & Wilhelm Geuder (eds.), *The Projection of Arguments: Lexical and Compositional Factors*, Standford, Center for the Study of Language and Information

David Dowty. 1991. Thematic Proto-roles and Argument Selection. *Language*, 67(3), 547-619

M. Ellsworth / K. Erk / P. Kingsbury / S. Padó. 2004. PropBank, SALSA, and FrameNet: How Design Determines Product. In *Proceedings of LREC-2004*, Lisbon.

Charles J. Fillmore, Christopher R. Johnson & Miriam R. L. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography,* 16 (3):235-250.

José M. García-Miguel & Francisco Albertuz. 2005. Verbs, Semantic Classes and Semantic Roles in the ADESSE project. In K. Erk, A. Melinger & S. Schulte im Walde (eds): *Proceedings of Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes*, Saarbrücken, pp. 50-55.

José M. García-Miguel, Fita González Domínguez & Gael Vaamonde. 2010. ADESSE, a Database with Syntactic and Semantic Annotation of a Corpus of Spanish. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. (Valletta, Malta: Mai 2010). European Language Resources Association (ELRA)

Daniel Gildea & Daniel Jurafsky. 2002. Automatic Labeling of Semantic Roles. *Computational linguistics*, 28(3):245-288.

Karin Kipper, Martha Palmer & Owen Rambow. 2002. Extending PropBank with VerbNet Semantic Predicates. *Workshop on Applied Interlinguas*, Tiburon, CA

Ronald W. Langacker. 1991. *Foundations of Cognitive Grammar, Vol. II: Descriptive Application*, Standford, Standford University Press.

Beth Levin. 1993. *English Verb Classes and Alternations: a Preliminary Investigation*. Chicago / London: University of Chicago Press

Maria Antònia Martí, Mariona Taulé, Manu Bertran & Lluís Màrquez. 2007. *AnCora: Multilingual and Multilevel Annotated Corpora*, Draft. http://clic.ub.edu/ancora/ancora-corpus.pdf

Martha Palmer, Daniel Gildea & Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles, *Computational Linguistics*, 31(1), 71-106.

Volha Petukhova & Harry Bunt. 2008. LIRICS Semantic Role Annotation: Design and Evaluation of a Set of Data Categories. *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*. (Marrakech, Morocco: Mai 2008). European Language Resources Association (ELRA).

Carlos Subirats. 2009. Spanish Framenet: A frame-semantic analysis of the Spanish lexicon. In Hans Boas, ed. *Multilingual FrameNets in Computational Lexicography. Methods and Applications*. Berlin/New York: Mouton de Gruyter, pp. 135-162.

Mariona Taulé, Mª Antonia Martí & Marta Recasens. 2008. AnCora: Multilevel Annotated Corpora for Catalan and Spanish. *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*. (Marrakech, Morocco: Mai 2008). European Language Resources Association (ELRA).

Robert D. Van Valin & Randy J. LaPolla. 1997. *Syntax. Structure, meaning and function*. Cambridge, Cambridge University Press

On some factors licensing the locative alternation

Paul Hirschbühler
Université d'Ottawa.
phirsch@uottawa.ca

Marie Labelle
Université du Québec à Montréal
Labelle.marie@uqam.ca

Abstract[*]

We review some factors playing a role in licensing the Ground frame in the locative alternation, in particular in connection with German *be-* and Russian *za-*

1 Introduction

In English verbs like *load* enter in the two constructions exemplified in (1).

(1) a. John loaded hay on the truck. (Figure frame)
b. John loaded the truck with hay. (Ground frame)

This is called the *locative alternation*, and it exemplifies the constructional versatility of verbs. The locative alternation is found in many languages. Here we focus on some factors at play in allowing verbs to occur in the variant with the Location as direct object (the Ground frame).

2 Semantic factor : compatibility of the verb with the constructional meaning

In languages like French (F) or English (E), alternating verbs are morphologically identical in both constructions (1a & 1b). This suggests that, given a (neo)constructional perspective, the meaning of the verb is central for its ability to be integrated in the type of meaning associated with each construction.

2.1 German

It has been argued that in German, the *be-* prefix selects a locative argument, to be realized as a direct object (e.g. Wunderlich 1997, Brinkman 1997). However, Dewell (2004) and Iwata (2008 :153-155), show that the facts are not that simple, since some of the relevant verbs are licit in the Ground frame with or without the prefix (compare 2b with 2c).

(2) a. Die Mutter streicht Marmelade aufs Brot.
"The mother is spreading jam on the bread"
b. Die Mutter streicht ein Brot mit Marmelade.
"The mother is spreading a sandwich (open-faced) with jam.
c. Die Mutter bestreicht ein Brot mit Marmelade.
"The mother is (be-)spreading a piece of bread with jam.

According to these authors, there is a difference in Aktionsart between the prefixed and unprefixed verb in the Ground frame, which, simplifying matter, amounts to a contrast between activity (with *be-*) *versus* accomplishment (without *be-*): "the *be*-verb typically describes a sustained activity with no defined endpoint, in contrast to the simple verb in the location-as-object frame, which has an endpoint" (Iwata 2005 : 154). If this analysis is on the right track, we must conclude that, as in E/F, it is the lexical meaning of the verb that is relevant for its compatibility with the two frames, the prefix introducing aspectual features orthogonal to those associated with the frame.

This is not the end of the story, however, since it is also the case that some German verbs may appear in the Ground frame *only* when prefixed with *be-* (e.g. *werfen/bewerfen, streuten/bestreuten, schütten/beschütten*), while appearing in the Figure frame without *be-*:

(3) a. Bin auch für Kernkraft! Überlegt einmal, wieviel Kerne wir täglich in den Mülleimer werfen, …
"I am for nuclear power, too! Just think how many seeds (lit. nuclei) we throw into the garbage can everyday, …"
b. … Überlegt einmal, mit wievielen Kernen wir täglich den Mülleimer bewerfen, …
"… Just think with how many seeds a day we be-throw the garbage can, …"

[*] We are very thankful to Nina Kazanina and Sasha Simonenko for their help with the Russian examples.

Michaelis and Ruppenhofer (2001 : 32) note that the only meaning available for (3b) is one "in which the seeds are thrown at the outside of the garbage can rather than inside it", contrary to what is the case in (3a). The (b) example conforms to the aspectual generalization proposed by Iwata (it describes an activity); but as we saw, this by itself is unsufficient to license the Ground frame. If the meaning of *werfen* is equivalent in the relevant respects to that of *throw*, which does not alternate, we must conclude that what licenses the Ground frame is neither the lexical meaning of the verb on its own, nor an aspectual meaning associated with *be*-. Thus, *be*- plays a crucial role in the well-formedness of the Ground frame in at least some cases, and the factor at play is still unclear. One possibility suggested by Michaelis and Ruppenhofer is that *be*- licences the Ground frame and in addition requires the direct object to satisfy some requirement, like : "the goal argument must denote the exterior of an object" ibid., p. 32), "be planar" (ibid, page 48). When the verb meaning conflicts with the meaning of the construction, the constructional meaning overrides it (provided the characteristics of the event allows a reanalysis) (ibid. p. 49).

A somewhat different case is that of a verb like *giessen,* for which *pour/verser* is the approximate equivalent often given. French *verser* is restricted to the Figure frame, as is English *pour*, although a few examples in the Ground frame are attested. In German, unprefixed *giessen* appears in both frames. The two verbal forms are possible when the direct object is *Blumen* ('*to water the flowers',* Booij 1992, Kordoni 2003). Koch & Rosengren (1996 :19) observe a meaning difference between *giessen/begiessen* in the Ground frame, which relates to the notions of internal *vs* external change :

> "The form with *be*- means an 'outer equipment', whereas the non-prefixed form has a functional interpretation of 'inner equipment'. Thus,
>
> *Der Mann goß den Puddel mit Wasser
> 'The man poured the poodle with water')
>
> sounds funny, since it implies that the poodle may start growing like a flower when 'sprinkled'. Contrary to this
>
> Der Mann begoß den Pudel mit Wasser
> 'The man poured the poodle with water'
>
> makes perfect sense."

Leaving aside the aspectual difference there might be between *giessen* and *begiessen* in the Ground frame, the lexical meaning difference between the two verbs suggests that *giessen* has two distinct lexical entries. In the Figure frame, it means *pour*; in the Ground frame its meaning is closer to *water*. *Begiessen* appears to be the prefixed version of *giessen* in its *pour* interpretation. If so, one could again try to defend the view that *be*- selects a locative argument to be realized as direct object, separately from the aspectual import that it might have.

To conclude this section, Dewell's and Iwata's view on the contrast between presence or absence of *be*- with certain verbs in the Ground frame are very interesting : if the contribution of *be*- is purely aspectual, as suggested by Iwata, this raises the question of why the assumed equivalents of a number of E/F verbs that just occur in the Figure frame do alternate in German, whether prefixed or not. Does the meaning of the German verbs differ from that of the E/F verbs with which they are often equated, and in a way that would make the difference in behaviour expected? For those verbs that require the prefix *be*- to appear in the Ground frame, is the aspectual contribution sufficient to make the Ground licit, given the meaning attached to the construction and to the verbs?

2.2 Russian

Russian does not seem to be as versatile as German. The only verbs entering the Figure frame that we have seen mentioned as licit also in the Ground frame without a prefix are *mazat'* 'smear/spread' and *gruzit'* 'load' (but obviously, this could be a gap in the limited literature we have consulted), suggesting that their lexical meaning is of a type compatible with the meaning associated to each frame.

Apart from these two verbs, the presence of a prefix, generally *za*-, less often *o(b)*- or *u*-, seems to be required in order for a verb that occurs in the Figure frame to also appear in the Ground frame. This is illustrated in (4) and (5) for the verb *sypat'* 'pour', which expresses a manner of movement and requires *za-* to occur in the Ground frame (examples from Mezhevich 2003) :

(4) Oni sypali pezok v luzhu.
They poured-IMPF sand-ACC into puddle
'They were pouring the sand into the puddle'

(5) Oni za-sypali luzhu peskom.
They ZA-poured-PF puddle-ACC sand-INSTR
'They dried up the puddle by pouring sand into it'

It appears that, in general, the relevant prefixes, combined with Figure frame verbs, directly license the Ground frame. Olbishevska (2005) indeed suggests that the relevant prefixes do two things : 1) they introduce a (result) State; 2) they take two arguments, a Location realized as direct object, and a Figure, realized as an oblique in the instrumental case. Regarding the suggested argumental properties, this is very similar to the analysis proposed by Wunderlich and Brinkmann for German *be*-, and is not unexpected given that these prefixes are homonymous with prepositions expressing topological relations : *o(b),* a cognate of German *be*-, is glossed as *about, around*; *za* is glossed as *behind.* Traces of these meanings found in the prefixed verbs play a role in their distribution (Tsedryk 2006).

Mezhevich's view is partially different. With freely alternating verbs like *gruzit'*, *za*- does not appear to take the location as an argument since the Ground frame is allowed without it. Mezhevich argues that the role of *za*- is purely aspectual: unprefixed *gruzit'* in the Ground construction indicates that the activity denoted by the verb potentially can result in the location being filled, while the presence of the prefix indicates that the potential result has been reached. Loosely speaking, this is the opposite of the aspectual contrast mentioned in relation to the absence *vs* presence of *be*- in German (it might thus be interesting to look more closely at the similarities and differences between *o(b*) and *be*-). For her, the prefixed verb expresses a *change of state* related to the manner of motion expressed by the verbal root, and the location object is an argument of this complex verb. The prefix itself identifies the result subevent in a complex event structure. She considers the meaning of these prefixes as "too vague to express any specific state", they have "no specific semantic content" (p. 14). The result subevent is the source of the licensing of the Ground frame with these complex verbs.

Why doesn't Mezhevich assume that *za*-, when it appears in the Ground frame, selects a Location as an external argument in the result subevent, which is very tempting, given its meaning? Her reason is that "Russian prefixed verbs are not necessarily transitive" (p. 15), and presumably, because the *za*-prefixed locative verbs may appear in both the Ground and Figure frames, as shown below (examples from Nina Kazanina, pc).

(6) a. On zalil benzina v bak
He zalil petrol-GEN into tank-ACC
'He put gas in the tank'

(7) b. On zalil bak benzinom
He-NOM za-pour.PAST tank-ACC with gas
'He filled the tank with gas'

Olbishevska's analysis does not address the issue of why *za*-prefixed verbs may appear in the Figure frame, but Tsedryk (2006) does. Tsedryk (2006, chapter 4) argues that in the Ground frame, the verb selects as a complement a small clause headed by *za*-, considered a resultative (R) or low-applicative head. When used in the Ground frame, *za*- assigns two theta-roles, a Goal and a Theme (corresponding to the Ground and the Figure). The representation of the VP in the Ground frame construction is as in (9), where *za*- assigns the instrumental case to its complement, just as the preposition *za* does in (8) (Tsedryk p. 83) .

(8) Oni byli za kirpitchami
They were behind bricks.INSTR

(9) [$_{VP}$ lil [$_{RP}$ [$_{DP}$ bak]$_{GOAL\text{-}ACC}$ [$_{R'}$ *za*- [$_{DP}$ benzinom]$_{THEME\text{-}INSTR}$]]]

Let us turn to the Figure frame. In the construction without *za*-, the Location argument is the complement of a locative preposition, *v* in the case of (8). In that configuration, the verb *lit'* takes what Tsedryk calls a SYM(metrical) complement, where the Theme DP and the PP combine directly, without an intermediate head. The verb assigns the Theme role to SYM, but as SYM is not referential, the role percolates to the DP immediately dominated by SYM.

(10) [$_{VP}$ lil [$_{SYM}$ [$_{DP}$ benzin]$_{THEME\text{-}ACC}$ [$_{PP}$ v [$_{DP}$ bak]$_{GOAL\text{-}ACC}$]]]

For the Figure frame including the prefix *za*-, Tsedryk suggests that the result phrase headed by *za*- is the complement of the verb, as it is in the Ground frame construction. But this time, *za*- does not take its Location and Figure arguments directly. It takes a SYM phrase complement, just as unprefixed *lit'* does. This gives us a richer representation :

(11) [$_{VP}$ lil [$_{RP}$ za- [$_{SYM}$ [$_{DP}$ benzin]$_{THEME\text{-}ACC}$ [$_{PP}$ v [$_{DP}$ bak]$_{GOAL\text{-}ACC}$]]]]

Besides the obvious structural differences between (10) and (11), Tsedryk declares that in (10) the PP does not have a thematic relation with the verb itself, which supports only a Theme theta-role, while in (11) the PP has a the-

thematic relation with the verbal prefix. In (11), *bak* saturates the Goal theta-role of *v* as well as the Goal theta-role of *za-*. *Benzin* saturates the Theme theta-role of *za-* as well as the Theme theta-role of *lit'*. Thus, in Tsedryk's syntactic approach, *za-* may enter two constructions. The Ground frame results when *za-* takes an internal and an external referential argument; the Figure frame results when *za-* takes only a complex, non-referential internal argument, the SYM phrase. This unified account of *za-* in the Figure and Ground frame is attractive in that the arguments are realized in different ways depending on the syntax while at the same time satisfying the lexical requirements of verbs, prefixes, and the preposition in a flexible way. The insights of this analysis can certainly be expressed by linguists of various theoretical persuasions in their favorite framework.

Because of space limitations we have not discussed *ob-*, which as we have indicated, could be compared to German *be-*. For interesting observations about *ob-*, we refer, besides some nice observations in Tsedryk's thesis, to Lewandowsky (2010) on *ob-* in Polish.

3 Conclusion

Taking English and French as a starting point, where the distribution of verbs in the Figure and Ground frame appears to depend solely on the lexical information carried by the verb, we briefly reviewed factors proposed to account for the licensing of the Ground frame in German and in Russian. We showed that purely locative or purely aspectual approaches are insufficient, and suggested that Tsedryk's analysis of Russian *za-* was attractive because of the flexible way in which the lexical requirements of various functors could be satisfied. It may be extended to German *be-* if we assume that *be-* is a low applicative, but one that may only take its Theme and Location arguments directly, not take a SYM complement, as discussed in Hirschbühler and Mchombo (2006), that is, *be-* may not take a SYM phrase as its internal argument. We will it to further research to answer the question as to why *be-* and *za-* would differ in this way. In this connection, it would be interesting to see if *ob-* sides more with *be-* than with *za-*

Reference

Booij, Geert. 1992. Morphology, semantics and argument structure. In Roca, Iggy M. (ed.), *Thematic Structure: its role in grammar*: 46-64. Berlin & New York: Foris Publications.

Brinkmann, Ursula. 1997. The Locative Alternation in German. Its Structure and Acquisition. Amsterdam/Philadelphia: John Benjamins.

Paul Hirschbühler & Sam Mchombo. The location object construction in Romance and Bantu : applicatives or not? Poster presentation. The Bantu-Romance Connection: An ESF Exploratory Workshop. May 26-27, 2006. Department of Linguistics & Phonetics University of Leeds

Iwata, Seizi. 2008. *Locative alternation : a lexical constructional approach*. Amsterdam-Philadelphia : John Benjamins.

Kordoni, Valia. 2003. Valence alternations in German : an LMT analysis. *Proceedings of the LFG03 Conference*. Miriam Butt and Tracy Holloway King (Editors) CSLI Publications.

Lewandowski, Wojciech. 2010. The locative alternation and verbal prefixation in Slavic: a constructional view. Abstract. Universitat Autònoma de Barcelona

Michaelis, Laura A. & Josef Ruppenhofer. 2001. *Beyond Alternations. A Constructional Model of the German Applicative Pattern*. Stanford: CSLI Publications

Olbishevska, Olesya. 2005. Locative Alternation in Slavic: The Role of Prefixes. *Proceedings of the 2004 Annual Conference of the Canadian Linguistic Association*, CLA website http://www.chass.utoronto.ca/~cla-acl/, 12 pages.

Spencer, Andrew & Marina Zaretskaya. 1998. Verb prefixation in Russian as lexical subordination. *Linguistics*, 36. 1-39.

Tsedryk, Egor. 2006. Alternances ditransitives: Une étude de la symétrie en syntaxe. Thèse de Ph.D. The University of Western Ontario.

Wunderlich, Dieter. 1997. Cause and the structure of verbs. *Linguistic Inquiry* 28 : 2 7-68.

A Framework for Representing Event Semantics of Verb Word Senses

Chris Irwin Davis
University of Texas at Dallas
cid@hlt.utdallas.edu

Abstract

In this paper we present a framework for representing event semantics as a set of semantic entities connected by binary relations. In contrast to frame semantics, our use of a fixed small number of entities and relations facilitates easy decomposition of event semantics into constituent parts, as well as allowing for integration into other systems and resources that rely on binary relations. We map each event representation onto a WordNet verb synset or cluster of related synsets. Thus, event semantics may be indexed by WordNet verb entries. Next, we describe a high-level taxonomy for the categorization of events based upon the semantic roles of the verb arguments. Finally, we briefly discuss acquisition techniques.

1 Introduction

Semantic representation of text is an important aspect of text understanding, reasoning, and identifying inferences. *Semantic relations* are a succinct and formal way to represent semantics in text and are the building blocks for creating the semantic structure of a sentence. In general, semantic relations are unidirectional connections between two concepts or *entities*. For example, the noun phrase "car engine" entails a PART-WHOLE relation: the *engine* is a part of the *car*.

Fillmore (1968) introduced the notion of thematic roles giving a semantic label to the connection between a verb and its arguments. He proposed a set of nine roles: AGENT, EXPERIENCER, INSTRUMENT, OBJECT, SOURCE, GOAL, LOCATION, TIME and PATH (Fillmore 1971).

Recently, Helbig (2005) proposed a classification of semantic entities in order to formally define the sorts of concepts that are valid for a given relation. This classification is performed by an ontology of entities, which defines a hierarchy of concept types following a semantic criteria. For example, in our framework AGENT holds between animate concrete objects (my wife, the president) and situations (arrive, decide) . It is therefore inapplicable to talk about inanimate objects (chair, rock) or abstract objects (yesterday, pain) being the Agent of a Situation.

Much work has been done in the development of resources for the representation of semantics. WordNet (Fellbaum, 1998)[1], FrameNet (Baker et al., 1998)[2], and PropBank (Palmer et al., 2005)[3] are three of the most widely used resources within the research community during recent years.

Our approach is like FrameNet in that we have predefined roles to be filled within a given archetypal semantic representation. In this respect, our semantic entities may be likened to frame elements. However, our use of binary relations allows for the decomposition of an event into constituent parts and the integration with other resources that utilize binary relations.

Like PropBank, our approach is verb focused, and centers on verbs and their arguments. Though unlike PropBank, our framework is not so tightly coupled with the syntactic domain.

Each semantic representation of an event we define is mapped to a WordNet verb synset (or cluster of related verb synsets). In this way we encode our representation within its associated synset as an extension to WordNet. This approach is inspired by the logic form transforms (LFTs) of WordNet glosses that are part of Extended WordNet (Moldovan and Novischi, 2002)[4]. In a similar fashion, we encode our extensions to synsets as a separate but parallel resource to WordNet.

[1] http://wordnet.princeton.edu

[2] http://framenet.icsi.berkeley.edu

[3] http://verbs.colorado.edu/~mpalmer/projects/ace.html

[4] http://xwn.hlt.utdallas.edu

2 Semantic Entities and Relations

The upper ontology for semantic entities and the set of 26 binary relations that we use are taken from Blanco et al. (2010). That work is concerned with combining two semantic relations that share a common argument and inferring a third relation between the remaining two unshared arguments. By contrast, we are representing the semantics of an entire discrete event archetype as the potentially complex combination of several semantic relations.

2.1 Upper Ontology of Entities

The root of the semantic entity ontology is simply Entity, and refers to *anything about which we can say something*. A diagram of the upper ontology of entities is shown in Figure 1.

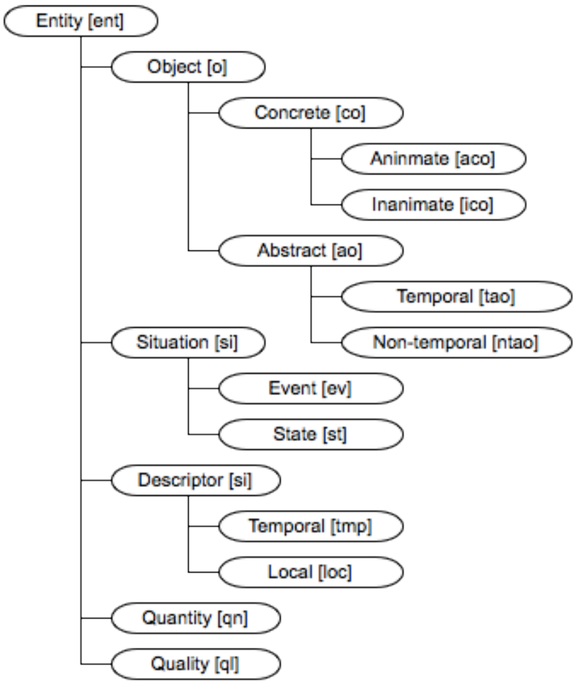


Figure 1. Entity upper ontology

We do not constrain semantic entities by part-of-speech or even to single lexical entries (an Entity may also be a syntactic phrase or clause). We do map WordNet noun synsets to entity types. This facilitates the identification of event arguments, either directly or through headword detection. The types of entity in the upper ontology are:

Objects can be either concrete or abstract. Concrete Objects are tangible things that exist in the physical universe. Abstract Objects are non-physical concepts that exist as a product of cognition. Concrete Objects may be either animate or inanimate. Animate Concrete Objects have agency, whereas Inanimate Concrete Objects do not. Abstract objects are either temporal or non-temporal. The former correlate with ideas regarding points or periods of time (e.g. *1984*, *tomorrow*), whereas the latter may be any other abstraction (e.g. *morality*, *illness*). Abstract objects can be perceived sensually (e.g., *pain*, *aroma*).

Situations are anything that occurs at a time and place. That is, if one can identify the time and location of an Entity, then it is a Situation. Events (e.g. *learn*, *dissolve*) imply a change in the status of other entities, States (e.g. *lying down*) do not. Although Situations can be expressed by either verbs or nouns, our framework focuses upon verbs.

Descriptors express temporal or spatial properties about entities. They may include an optional non-content word (e.g. a preposition) that indicates the temporal or spatial context in relation to another entity.

Qualities are qualitative properties than can be associated with entities. They can be either relative, (e.g. *wealthy*, *small*) or absolute, (e.g. *awake*, *invisible*).

Quantities are quantitative properties of entities (e.g., *750ml*, *a couple of dollars*).

2.2 Binary Semantic Relations.

A binary semantic relation is a relationship between two Entities that is expressed REL(x, y) and may be read "x is REL of y". We constrain the arguments of a binary semantic relation by entity type. In addition to having a more clear and concise definition for each semantic relation, defining the types of concepts that can be part of the DOMAIN and RANGE of a relation has several advantages: (a) helping to discard potential relations that do not hold. For example, inanimate objects cannot have INTENT. (b) aiding in the combining of semantic relations. By checking domain and range compatibilities, valid combinations of relations can be determined. The complete list of 26 relations is depicted in Table 1 along with DOMAIN and RANGE restrictions and examples of valid arguments.

3 Framework Description

In this section we describe the framework for combining the binary relations and semantic entities. We also expand the usage of three important entities from the upper ontology: State, Event, Non-temporal Abstract Object, and the ways they may interconnect in addition to binary relations.

Cluster	Relation Type	Abbreviation	Domain × Range	Examples
Reason	CAUSE	CAU	[si] × [si]	CAU (*virus*, *influenza*)
	JUSTIFICATION	JST	[si ∪ ntao] × [si]	JST (*it is illegal*, *not speeding*)
	INFLUENCE	IFL	[si] × [si]	IFL (*missing classes*, *poor grade*)
Goal	INTENT	INT	[si] × [aco]	INT (*teach*, *professor*)
	PURPOSE	PRP	[si ∪ ntao] × [si ∪ co ∪ ntao]	PRP (*storage*, *garage*)
Object Modifiers	VALUE	VAL	[ql] × [o ∪ si]	VAL (*smart*, *kids*)
	SOURCE	SRC	[loc ∪ ql ∪ ntao ∪ ico] × [o]	SRC (*Spanish*, *student*)
Syntactic Subjects	AGENT	AGT	[aco] × [si]	AGT (*John*, *bought*)
	EXPERIENCER	EXP	[o] × [si]	EXP (*John*, *heard*)
	INSTRUMENT	INS	[co ∪ ntao] × [si]	INS (*the hammer*, *broke*)
Direct Objects	THEME	THM	[o] × [si]	THM (*a car*, *bought*)
	TOPIC	TPC	[o ∪ si] × [si]	TPC (*agenda*, *discuss*)
	STIMULUS	STI	[o] × [si]	STI (*symphony*, *heard*)
Association	ASSOCIATION	ASO	[ent] × [ent]	ASO (*salt*, *pepper*)
	KINSHIP	KIN	[aco] × [aco]	KIN (*John*, *his wife*)
None	IS-A	ISA	[o] × [o]	ISA (*sedan*, *car*)
	PART-WHOLE	PW	[o] × [o] ∪ [l] × [l] ∪ [t] × [t]	PW (*handlebar*, *bicycle*)
	MAKE	MAK	[co ∪ ntao] × [co ∪ ntao]	MAK (*cars*, *BMW*)
	POSSESSION	POS	[co] × [co]	POS (*Ford F-150*, *John*)
	MANNER	MNR	[ql ∪ st ∪ ntao] × [si]	MNR (*quick*, *delivery*)
	RECIPIENT	RCP	[co] × [ev]	RCP (*Mary*, *gave*)
	SYNONYMY	SYN	[ent] × [ent]	SYN (*a dozen*, *twelve*)
	AT-LOCATION	AT-L	[o ∪ si] × [loc]	AT-L (*party*, *John's house*)
	AT-TIME	AT-T	[o ∪ si] × [tmp]	AT-L (*party*, *last Saturday*)
	PROPERTY	PRO	[ntao] × [o ∪ si]	PRO (*height*, *John*)
	QUANTIFICATION	QNT	[qn] × [o ∪ si]	QNT (*a dozen*, *eggs*)

Table 1. The 26 Binary Relations

3.1 State

We augment the usage of the State entity by allowing it to be associated with a semantic relation to indicate the State where a particular relation holds. For example, a state s may be associated with the binary relation POS(x, y) to indicate the State constituted by x being possessed by y.

3.2 Non-temporal Abstract Objects

Like the State entity, we also allow Non-temporal Abstract Objects to be associated with a semantic relation. This indicates a concept that is the object of a cognitive process (i.e. *thought*, *idea*, *belief*). That is, such a Non-temporal Abstract Object is the cognitive *concept* of the associated semantic relation.

We have noted that non-temporal abstract objects may also be entities that are sensually perceived (pain, odor, fear). We put a finer point on this by further identifying cognitive perceptions (*idea*, *belief*, *thought*). To represent these, we allow non-temporal abstract objects to be associated with a semantic relation to indicate the cognitive *concept* of that relation as an entity.

3.3 Events

We define an Event as either an ongoing, continuous change in status of an entity (e.g. *growing*, *rotating*), or the transition of one discrete State to another State. In the latter case, we note additional properties of a State transition related to lexical aspect: Durativity and Telicity. A State changing Event may be durative and take place over a time period (e.g. *drive*, *eat*) or non-durative (e.g. *sneeze*, *hit*). Telicity indicates whether an Event has a defined goal or completion. “Repaired cars for a week” is atelic, while “repaired *a* car last week” is telic. These event properties and the terms associated with their intersection are summarized in Table 2.

	Durative	***Non-durative***
Telic	Accomplishment	Achievement
Atelic	Activity / Process	Semelfactive

Table 2. Event lexical aspect properties

3.4 Semantic Widgets

To more easily illustrate the event semantics of our framework we introduce the notion of *semantic widgets*, an informal and convenient way of graphically viewing semantic relationships. Our use of the term *widget* is suggestive of how it may be combined with other compatible widgets in the representation of larger blocks of text. Figure 2 shows an illustration of the semantic relationships entailed by the verb “find”, whose WordNet synset is “find, regain”. This example demonstrates the major components of the framework.

Entities are represented by encircled letter variables. The negation character “!” is a shorthand way of indicating a distinct entity that is of the same type, but whose value is disjoint with its counterpart. In Figure 2, k and $!k$ indicate two different States, where if $k = s_1$ and $!k = s_2$, then $s_1 \neq s_2$.

Relations are symbolized by a directional shape that indicates the order of the arguments.

Events that illustrate the transition of one State to another have an equivalence double line connecting it to a triangle. In this example, the Event is z. The triangle is in turn connected to the two States, $!k$ and k, and shows the direction of the transition.

Finally, the non-temporal abstract object c, is the *concept* of y being at location p.

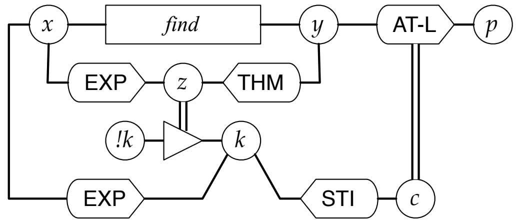


Figure 2. Semantic widget for the verb “find”

To get a better sense for the distinction between semantic frames and our semantic representation of archetypal events, we present a specific example. The FrameNet analog for the representation in Figure 2 is the verb entry `find.v`, a member of the LOCATING frame whose definition states “A [PERCEIVER] determines the [LOCATION] of a [SOUGHT_ENTITY] within a [GROUND]”. An example sentence from the LOCATING frame is shown in Figure 3.

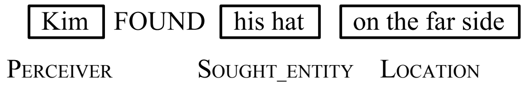


Figure 3. FrameNet example for LOCATING frame.

Figure 4 then shows the same sentence annotated with typical binary semantic relations.

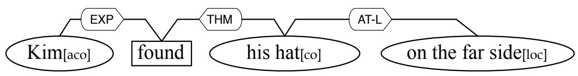


Figure 4. FrameNet LOCATING frame example sentence with typical semantic relations.

Figure 5 shows the same FrameNet example from the LOCATING frame represented using our framework. We do not account for the FrameNet concept of GROUND. However, we do account for more detail in the semantic nature of the event itself. The “found” event is an *achievement*, both telic (a realized goal) and non-durative (the transition from not found to found is instantaneous).

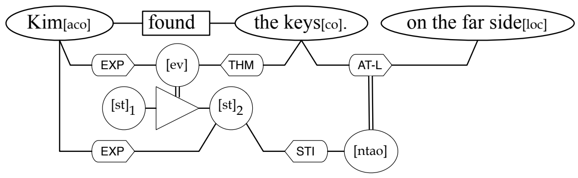


Figure 5. FrameNet LOCATING frame example sentence in our framework.

We show another widget in Figure 6 to illustrate how States may be complex, being associated with multiple relations simultaneously.

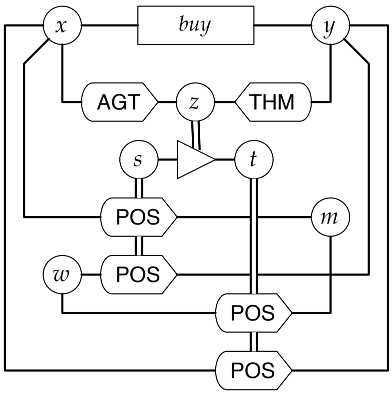


Figure 6. Example of an event with complex states.

Here we see that the possession of a good *y* and of monies *m* are exchanged by a buyer *x* and seller *w*.

3.5 Representation Encoding

The example widget diagram from Figure 2 is more formally encoded as below. Each element has been placed on a separate line for clarity.

```
find(x,y) ->
        EXP(x,z) &
        THM(y,z) &
        EXP(x,k) &
        STI(c,k) &
        x[aco] &
        y[co] &
        k[st] &
        z[ev]=(!k,k) &
        p[loc] &
        c[ntao]=AT-L(y,p)
```

It indicates the valid entity type for each variable, binary relations on those variables, and Event and Non-temporal Abstract Object associations. This is the form used in our resource that is mapped to WordNet verb synsets.

We indicate the entity typing of a particular variable with a bracketed suffix, e.g. *x*[aco]. This indicates that the entity variable *x* is an Abstract Concrete Object.

3.6 Taxonomy

For our taxonomy, we categorize by pairs of binary relations based upon verb arguments, one each from the Subject and Direct Object clusters of semantic relations. These relations share a common semantic Situation denoted by the variable *z* in Figures 2 and 6.

We write these categories in shorthand by the abbreviation for the Subject and Direct Object relations combined with a plus sign. Thus, EXP+THM and AGT+THM would symbolize the categories for the event structures in Figures 2 and 6 respectively.

Those relations that are compatible are shown in Table 3 with a check mark. Those combinations without a check are syntactic combinations that do not hold semantically. We recognize seven high-level categories.

		Direct Object		
		THM	***TPC***	***STI***
Subject	***AGT***	√	√	√
	EXP	√		√
	INS	√		√

Table 3. Event categories

4 Acquisition and Evaluation

Since event attributes may be inherited through associated synset hypernymy, acquisition of event semantics can be automatically propagated down the WordNet hierarchy. We therefore used semantic coverage created manually at the top of a synset tree to automatically seed constituent hyponyms.

Additionally, we note that some languages have properties that better allow for the automatic detection of particular semantic features than others. For instance, there are linguistic features of Modern Persian that allow for the automatic detection of lexical aspect (Folli, et al., 2003).

Farreres, et al. (2010) provide a theoretical foundation for mapping WordNet synsets between languages. Using standard bilingual alignment techniques (Och, 1999) and (Och and Ney, 2000), we mapped synsets between English and Persian, propagating (a) lexical aspect features back to English that were automatically acquired from Persian verbs, and (b) semantic entity information from Persian nouns.

	Precision	**Recall**	**F-measure**
Verb	69.23%	52.94%	60.0%
Noun	84.0%	53.85%	65.63%

Table 4. English-Persian synset mapping accuracy

Table 4 shows an evaluation of our English-Persian WordNet synset mapping of verbs associated with semantic event representations and nouns associated with semantic entities.

References

Baker, Collin F., Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet Project. In *Proceedings of the 17th international conference on Computational Linguistics*, Montreal, Canada.

Blanco, Eduardo, Hakki C. Cankaya and Dan Moldovan. 2010. Composition of Semantic Relations: Model and Applications. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, Beijing, China.

Farreres, Javier, Karina Gibert, Horacio Rodríguez, and Charnyote Pluempitiwiriyawej. 2010. Inference of lexical ontologies. The LeOnI methodology. *Artificial Intelligence*, Vol. 174, No. 1. (25 January 2010), pages 1-19.

Fellbaum, Christiane. 1998. *WordNet: An Electronic Lexical Database*. MIT Press

Fillmore, Charles J. 1968. The case for case. *Universals in Linguistic Theory*, pages 1–88, April.

Fillmore, Charles J. 1971. Some problems for case grammar. *Monograph Series on Languages and Linguistics*, 24:35–36.

Folli, Raffaella, Heidi Harley, and Simin Karimi. 2003. Determinants of event type in persian complex predicates. In *Cambridge Working Papers in Linguistics*, pages 101–126.

Helbig, Hermann. 2005. *Knowledge Representation and the Semantics of Natural Language*. Springer, 1st edition.

Moldovan, Dan and Adrian Novischi. 2002. Lexical Chains for Question Answering. In *Proceedings of the 19th International Conference on Computational Linguistics (Coling 2002)*. Taipei, Taiwan.

Moldovan, Dan I. and Vasile Rus. 2001. Logic Form Transformation of WordNet and its Applicability to Question Answering. In *Proceedings of ACL 2001*. Pages 394-401

Och, Franz Josef. 1999. An Efficient Method for Determining Bilingual Word Classes. *Ninth Conference of the European Chapter of the Association for Computational Linguistics*, pages 71-76, EACL'99.

Och, Franz Josef and Hermann Ney. 2000. Improved Statistical Alignment Models. In *Proceedings of ACL 2000*, pages 440-447.

Palmer, Martha, Daniel Gildea, and Paul Kingsbury. 2005. The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106

The interaction of semantic-syntactic context with "mental activity" interpretations of Italian verbs of visual perception

Martina Johnson
University of Pisa
martinaljohnson@gmail.com

Abstract

This paper draws on a frame semantic analysis of Italian verbs of visual perception in order to discuss the distributional features of these verbs (e.g. the syntactic and semantic characteristics of their arguments, but also their lexical collocates) that cause them to have an interpretation in the domain of mental activity.

1 Introduction

This paper is a corpus-based study on the "mental activity" senses of three Italian verbs – *vedere* (see), *intravedere* (make out or glimpse), and *scorgere* (a near-synonym of *intravedere*). Much like the English verb *see*, these verbs may be used to describe experiences that are predominantly mental as well as the experience of visual perception, as examples (1 a) and (1 b) show.

(1) (a) *Vedo un cane.*
I see a dog.
(b) *Vedo una difficoltà.*
I see a difficulty.

The data are based on a frame semantic analysis of the verbs. This was carried out in the context of the Italian FrameNet project (Lenci et al., 2010), whose goal is to create a frame-based electronic lexicon similar to the original Berkeley FrameNet.[1] The main tenet of Frame Semantics (Fillmore, 1985; Fillmore and Atkins, 1992; Fillmore et al., 2003) is that each sense of a word evokes a semantic frame – a schematic representation of a situation or an event – in the mind of language users. Each frame is constituted by a group of participants in the situation, or Frame Elements (FEs); these are instantiated syntactically by the frame-bearing word's arguments (if it is a verb) or complements (if it is a noun or adjective). The information for individuating a language's semantic frames is obtained by annotating corpus sentences with FEs (similar to semantic roles) and syntactic information.

Frame Semantics focuses prevalently on a static description of the syntax-semantics interface: a frame is devised to appropriately capture the meaning of a word in context, and then the semantic roles of its arguments (or complements) are described, along with any possible syntactic alternations. What I would like to discuss in this paper, however, are the dynamics of semantic interpretation: what are the distributional features of a word (in particular, of a verb) that cause it to have a certain meaning?

This subject has traditionally been the object of James Pustejovsky's line of inquiry. In the Generative Lexicon (Pustejovsky, 1998), the mechanisms of coercion and co-composition show that a verb can influence the semantics of its arguments, but that the reverse is true, too: in some cases a verb can force the appropriate semantic type on its object, but in other cases it is the argument that picks out certain semantic features of its head verb (and excludes others).

A similar argument is developed in Hanks' "Corpus Pattern Analysis" (CPA) approach (Hanks and Pustejovsky, 2005). According to CPA, words out of context have no specific meanings, but a «multifaceted *potential* to contribute to the meaning of an utterance» (Hanks and Pustejovsky, 2005, 64). The meaning of a word is influenced not only by the syntactic pattern it occurs in, but also by the semantic type of the words in that pattern. In fact, the combination of different semantic types in the same syntactic pattern often gives rise to different word senses: for example, *shoot* in the sentence *shoot a person* could conceivably be ambiguous, depending on

[1] See e.g. (Ruppenhofer et al., 2006). FrameNet may be consulted online at http://framenet.icsi.berkeley.edu

whether the subject of the sentence is an armed attacker or a film director (Hanks and Pustejovsky, 2005, 68). However, CPA does not just take the characteristics of a verb's arguments into account, but also any additional and recurrent collocates of the verb which act as «clues» to its interpretation, such as *dead* in *shoot a person dead*.

In this paper, I will put these positions to the test, using the distributional data afforded by our frame semantic analysis as a base. I will examine the semantic and syntactic characteristics of the FEs occurring with *vedere, intravedere,* and *scorgere,* as well as other significant elements in the linguistic context, in order to determine their effect on the interpretation of these verbs. As a result, I will present an overview of the contextual features that cause a "mental activity" interpretation, and how this comes about.

2 "Mental activity" interpretations of *vedere, intravedere,* and *scorgere*

2.1 The epistemic interpretation

Many linguistic and philosophical studies on verbs of perception (both in Italian and English) focus on these verbs' so-called "epistemic" interpretation.[2] A verb of perception is said to be used epistemically when it does not express an experience of perception, but an act of deduction or reasoning, possibly based on perceivable objects. For example, in the sentence "I **see** John playing tennis", the speaker is relating a direct perceptual experience: s/he is in fact seeing John in the act of playing tennis at the moment of the utterance. If s/he says "I **see** that John is playing tennis", on the other hand, this does not necessarily mean that s/he can see him playing (although this interpretation is also possible). S/he might have simply noticed that his racket and tennis shoes are missing from the usual place where he keeps them, and made a deduction based on that perceptual data. There are also cases where the verb loses its perceptual meaning entirely: in a sentence like "I can **see** that the economical situation is difficult", it is not implied that the speaker has reached this knowledge through direct observation (in fact, it is fairly unlikely). While important, these studies miss the wider scope of the polysemy of verbs of perception, which does not just include a "perceptual" and an "epistemic" interpretation. In addition, most of these studies focus on the difference between *that*-clauses and perception verb-specific complements, such as NP followed by a naked infinitive or an *-ing* form in English, and do not take the wider variety of syntactic constructions that these verbs can occur with into consideration. An exception to this is (Baker, 1999), a study on the polysemy of the verb *see* based on Frame Semantics, which includes a detailed description of the various syntactic patterns that occur with this verb.

2.2 A frame semantic analysis

As mentioned above, a frame semantic analysis of a word begins with the study of corpus data. A sample of sentences that is deemed representative of the word's most typical FE combinations and their syntactic realizations is extracted from the corpus. Each sentence is assigned an appropriate frame, representing the meaning of the frame-bearing word; then, it is annotated with information on the FEs.[3] The data for this analysis were extracted from *La Repubblica* (Baroni et al., 2004), one of the largest corpora for Italian (ca. 390 million tokens), composed of newspaper texts.

The representative sample of sentences for each verb featured about 15 syntactic patterns, and included instances both of perception-related senses and mental activity ones. I found that the interpretations related to mental activity accounted for about half of the instances in each sample. Since the selection was not random, but based on syntactic patterns, the figure is not statistically significant, but it is still interesting.

Assigning the appropriate frame to the mental activity senses was not a simple task. For a first approximation, I paraphrased each instance with a verb of mental activity (e.g. *think, believe, consider*) and selected from the FrameNet inventory the frame that was evoked by that verb. I then checked whether the meaning of the whole sentence was compatible with that frame and its FE structure. At the end of this process, I identified three main senses, expressed by the frames AWARENESS, EXPECTATION, and CATEGORIZATION.

Since standard FrameNet frames are fairly general, this procedure had the effect of downplaying the differences in meaning between *vedere* on the one hand and *intravedere* and *scorgere* on the

[2] See for example (Barwise, 1981; Declerck, 1981; Guasti, 1993; Higginbotham, 1983; Kirsner and Thompson, 1976).

[3] For more information on the methodology of our analysis, see (Lenci et al., 2010).

other. While *vedere* expresses an experience of perception without specifying anything about its circumstances, *intravedere* and *scorgere* describe an experience which is either fleeting and transitory in nature, or made difficult and uncertain by obstacles to vision. This element of meaning is generally carried over in mental activity interpretations as well, with the result that the subject of the verb is less certain of the validity of the cognitive content s/he is entertaining. This element is lost in the analysis that I am presenting here. On the other hand, the similarities in behavior among these three verbs are quite striking: they all occur with the same syntactic patterns and nearly with the same senses (apart from the distinction that I just made).

One more thing that must be noted is that each sense had realizations that were metaphorical and ones that were not. I rely here on the definition of metaphor proposed by Lakoff and Johnson (Lakoff and Johnson, 1980): a mapping between two conceptual domains that enables us to intepret one in terms of the other. There were sentences where two domains seemed to be activated at the same time – one related to perception and one to mental activity – and sentences where the verb the perceptual meaning of the verb seemed entirely absent. I will discuss typical syntactic patterns both for metaphorical and non-metaphorical realizations.

Here are the frames I assigned to the mental activity readings of *vedere, intravedere* and *scorgere*.

2.3 AWARENESS

AWARENESS is the frame representing the verbs' epistemic interpretation. This frame refers to a situation where «a Cognizer has a piece of Content in their model of the world. The Content is not necessarily present due to immediate perception, but usually, rather, due to deduction from perceivables»,[4], which is very close to the definition I gave for epistemicity above. Other verbs that evoke this frame are *know, understand, be aware, believe,* and *think*.

The typical syntactic complement for epistemic uses of verbs of visual perception (and therefore for their AWARENESS sense) is the declarative *che* (that)-clause, which expresses the conceptual Content, as in sentence (2).

(2) *[Con la tomografia $_{\text{MEANS}}$] abbiamo potuto* ***intravedere*** *[che c'è una sedimentazione tra i due cervelli $_{\text{CONTENT}}$].*
Thanks to the CAT scan, we could glimpse that there is some sedimentation between the two brains.

In sentence (2), *con la tomografia* 'thanks to the CAT scan' expresses the Means by which awareness of the Content was achieved. Since a CAT scan provides perceptual data, which is the basis for the deduction expressed by the *che*-clause, *intravedere* retains some perceptual meaning, although the Content is actually a conclusion that must be believed or thought of. However, there are other cases that make no reference at all to physical perception (as shown in Section 2.2 above).

The AWARENESS sense also often emerges when the verb's direct object is an abstract noun, as in (3). The fact that the object of the verb is a non-perceivable entity reinforces the "mental activity" interpretation.

(3) *[Elena $_{\text{COGNIZER}}$] ha certo le conoscenze sufficienti per* ***vedere*** *[il senso della sua posizione $_{\text{CONTENT}}$].*
Helen certainly has enough knowledge to understand the meaning of her position.

Finally, another typical argument for verbs used in this sense is a complement headed by the preposition *da* (from), as in example (4).

(4) *[Discendiamo dagli egiziani $_{\text{CONTENT}}$], si* ***vede*** *[dal nostro viso, dal taglio degli occhi e dei capelli $_{\text{EVIDENCE}}$].*
We are descended from the Egyptians, you can see it from our faces, from the shape of our eyes and of our hair.

The *da*-complement represents the Evidence on which the awareness is based. This syntactic pattern is typical of other verbs that evoke this frame but do not have any readings associated with perception, such as *capire* (understand) and *dedurre* (deduce). This could be termed a case of exploitation, in Patrick Hanks' terms: a syntactic pattern that is the norm for other verbs is exploited in order to assign *vedere* and the others these verbs' meaning. The *da*-complement that expresses Evidence should not be confused with the locative complement introduced by the same preposition, which expresses the Location of Perceiver, as in *Ho visto i fuochi d'artificio [dal*

[4] The frame definitions are taken from the FrameNet website, `http://framenet.icsi.berkeley.edu`

tetto$_{LOCATION\ OF\ PERCEIVER}$*]* (I saw the fireworks from the roof).

The main difference between the two seems to be that in the second case, the noun instantiating the Location of Perceiver must be a possible location. Otherwise, the argument is interpreted as Evidence. In this case, therefore, it is a combination between syntactic form and semantic type of the argument that triggers the AWARENESS interpretation.

The metaphorical patterns associated with AWARENESS are mostly constituted by a direct object and a locative expression introduced by *in* (in)*, dietro* (behind)*, oltre* (beyond)*, attraverso* (through), and so on. These expressions create a spatial scene which reinforces the perceptual sense of the verb.

(5) *[Surin* $_{COGNIZER}$*]* ***intravede*** *[in Jeanne* $_{EVIDENCE}$*] [le stesse passioni, gli stessi desideri dai quali è torturato lui* $_{CONTENT}$*].*
Surin believes Jeanne has the very same passions and desires that he has always been tortured by.

(6) *[Tanti* $_{COGNIZER}$*]* ***scorgono*** *[dietro la sollevazione* $_{EVIDENCE}$*] [una ricerca di dignità e autonomia* $_{CONTENT}$*].*
Many believe there is a search for dignity and autonomy at the root of the rebellion.

In (5) and (6), the locative expression also represents the Evidence. It is by looking at Jeanne (or, in a broader sense, by talking to her, observing her behavior, coming into contact with her) that Surin comes to the (subjective) conclusion that she has the same passions and desires as him. Similarly, it is by studying the rebellion and investigating its context that many come to the conclusion that a search for dignity is at its root.

Once again, complements introduced by locative prepositions may be used to express FEs relative to the Perception experience frame, too. *In*-complements usually express the Ground and *dietro*-complements express the Direction of perception, as in ***Ho visto*** *un cane [in giardino* $_{GROUND}$*]/[dietro il cancello* $_{DIRECTION}$*]* (I saw a dog in the garden/behind the fence).

What causes the AWARENESS interpretation in (5) and (6)? On the one hand, the nouns in object position refer to abstract, non-perceivable entities. On the other, the locative complements must express an actual location if the literal interpretation of the verb is to make any sense. Here, too, then, the combination between syntactic pattern and semantic features of the arguments motivates the semantic interpretation of the head verb.

The level of metaphoricity of these sentences is not always the same. Hanks maintains in (Hanks, 2007) that metaphoricity is gradable, depending on how different the two conceptual domains that are involved in the metaphor are. The fewer semantic features they share, the more metaphorical the expression. However, in these cases the "force" of the metaphor seems to derive from the complexity and definition of the spatial scene that is created in the sentence. If the only lexical element in the sentence referring to space is a preposition, the sentence reads as somewhat less metaphorical than cases where the context is more richly built up by other elements. See, for example, sentence (7).

(7) *C'è un'altra ipotesi che si* ***intravede*** *nel nebbione.*
There is another hypothesis that we can glimpse among the fog.

In this sentence, *nel nebbione* (among the fog) absolutely cannot be interpreted as Evidence; it is, effectively, the Ground of perception. The visual "scene" is given more substance here, but the mental activity interpretation is maintained because *ipotesi* (hypothesis) is an entity that cannot be seen, but must be grasped conceptually.

2.4 EXPECTATION

The second frame that I selected is EXPECTATION. This refers to a situation where «a Cognizer believes that some Phenomenon will take place in the future». It is typically evoked by verbs such as *expect, foresee* and *predict*; I assigned it to *vedere, intravedere,* and *scorgere* in cases where they assume a "foreseeing" interpretation.

The EXPECTATION sense is not associated with any particular syntactic constructions. It may occur with a direct object or with a *che*-clause, but these are very widespread syntactic patterns. Furthermore, in some cases the sentences featuring the EXPECTATION sense are identical to AWARENESS sentences both for syntactic patterns and noun semantic types, except for one element which expresses a reference to the future. This can be a noun whose meaning has to do with the future, such as *futuro* (future), *prospettive* (possibilities), *rischio* (risk), *obiettivo* (goal). In sentences (8)-(10), it is the direct object of the verb, but it can be any other element of the sentence. In (10), for instance, the EXPECTATION sense emerges because *previsioni economiche* (financial estimates) are mentioned. It is most plausible that an economist writing a financial estimate will not write merely about being aware of the economic recovery; rather, he or she will try to predict if it will happen. Finally, a reference to the future may also be made through the choice of verb tenses, as in sentence (11).

(8) *[Che futuro$_{PHENOMENON}$] lei **vede** per la Nato?*
What kind of future do you foresee for Nato?

(9) *[Che prospettive nuove$_{PHENOMENON}$] **intravede** alla guida di uno stabile?*
What new possibilities do you see, now that you are in charge of a theatre?

(10) *Nelle previsioni economiche dell'anno scorso [la ripresa economica$_{PHENOMENON}$] si **intravedeva** fra mille segnali contraddittori.*
In last year's financial estimates, the economic recovery could only be glimpsed amongst a myriad of contradictory signals.

(11) *Si può **vedere** con una certa sicurezza [quale sarà l'evoluzione futura del commercio estero americano$_{PHENOMENON}$].*
We can foresee with some certainty what the future evolution of American foreign trade will be.

These data show a different pattern from the one we have seen until now. In the case of the EXPECTATION sense, the syntactic form and semantic type of verb arguments alone cannot be used to predict verb meaning. What we must do here in order to understand the interaction of the verb with its context is to study its collocates in a more general sense, and how they affect its meaning. We may therefore conclude that sometimes, the fine semantics of lexical collocates is crucial for interpreting a verb's meaning.

There are also metaphorical collocations for the EXPECTATION sense, such as *vedere nel futuro* (to see in the future) and *vedere all'orizzonte* (to see on the horizon). As with AWARENESS above, these sentences construct a spatial "scene" which recalls the perceptual interpretation of the verb, thus causing a metaphorical interpretation.

(12) *Dottor Falcone, cosa è possibile **intravedere** nel futuro della mafia?*
Doctor Falcone, what can we foresee for the mafia's future?

(13) *Non **scorgiamo** all'orizzonte alcun referendum.*
We do not foresee any referendum (in the near future).

2.5 CATEGORIZATION

Finally, the third frame I selected is CATEGORIZATION. In this frame, «a Cognizer construes an Item as belonging to a certain Category». It is typically evoked by verbs like *classify, consider,* and *regard* (as in, "I regard him as a brother"). This sense occurs only in connection with two specific syntactic patterns, in contrast to the other two frames which show such a wide variety of realizations. Collin Baker calls this sense a "semi-collocation": «semi-collocations are separate senses that tend to co-occur with a small number of lexical forms or syntactic patterns, but are not as fixed as real collocations; the range of words they require as part of their context is usually best described intensionally than extensionally» (Baker, 1999, 45). The patterns are exemplified in sentences (14) and (15).

(14) *Un famoso critico scrisse una volta che [mi $_{ITEM}$] **vedeva** [come un lanciatore di giavellotto che si volta indietro per fare arrivare il più lontano possibile la sua asta $_{CATEGORY}$].*
A famous critic once wrote that he saw me as a javelin thrower, who turns backwards in order to throw his pole as far as possible.

(15) *Nessuno, onestamente, può **intravedere** [nei vari segretari di partito messicani $_{ITEM}$] [un Pancho Villa $_{CATEGORY}$].*
No one, honestly, can see in the various Mexican party secretaries a new Pancho Villa.

In one case, the categorized Item is expressed as the direct object of the verb and the Category is expressed as a complement introduced by *come* (as); in the second case, the Item is expressed by a complement introduced by *in* and the Category is expressed by a direct object. The first pattern is unambiguous; the second can be confused with the metaphorical pattern for AWARENESS exemplified in 5 above. The differences in semantic types are not particularly helpful here: the main criterion is that the direct object must be a category that the *in*-complement can fit into. In the case of CATEGORIZATION, this syntactic pattern is not metaphorical. The reason is probably that it is very highly conventionalized, so that even the idea of a spatial scene has been bleached out completely.

3 Conclusions

In this paper, I used data from a frame semantic analysis of the verbs *vedere, intravedere,* and *scorgere* to answer the question: what are the distributional features of these verbs that cause them to have a mental activity reading? The underlying goal was to represent semantic interpretation from a dynamic point of view, through the study of these verbs' interaction with linguistic context.

I found that the interaction of these verbs' meaning with their linguistic context is multifaceted and complex, involving three different factors: the syntactic patterns that the verbs occurred with, the semantic type of their arguments, and the semantic features of recurrent lexical collocates. In some cases, a typical syntactic pattern is enough to force a certain meaning, such as *Lo vedo come un amico* (I see him as a friend) in relation to the CATEGORIZATION sense. Often, however, a combination between syntactic pattern and semantic type is required for a specific sense to be triggered: see the difference between *Ho visto un cane in giardino* (I saw a dog in the garden) and *Vedo in te un grande coraggio* (I see great courage in you). As the discussion on the EXPECTATION sense shows, though, sometimes a certain meaning emerges through the interpretation of more subtle cues, related to the fine semantics of the verb's lexical collocates. These data are in keeping with Pustejovsky and Hanks' approaches to semantic interpretation, the Generative Lexicon and Corpus Pattern Analysis.

References

Collin Freeman Baker. 1999. *Seeing clearly: Frame Semantic, Psycholinguistic, and Cross-linguistic Approaches to the Semantics of the English Verb See*. Ph.D. thesis, University of California, Berkeley.

Marco Baroni, Silvia Bernardini, Federica Comastri, Lorenzo Piccioni, Alessandra Volpi, Guy Aston, and Marco Mazzoleni. 2004. Introducing the "la Repubblica" corpus: A large, annotated, TEI (XML)-compliant corpus of newspaper Italian. In *Proceedings of LREC 2004*, pages 1771–1774, Lisboa, Portugal.

Jon Barwise. 1981. Scenes and Other Situations. *The Journal of Philosophy*, 78(7):369–397.

Renaat Declerck. 1981. On the role of progressive aspect in nonfinite perception verb complements. *Glossa*, 15:83–114.

Charles J. Fillmore and Beryl T. (Sue) Atkins. 1992. Towards a frame-based lexicon: The semantics of RISK and its neighbors. In A Lehrer and E.F. Kittay, editors, *Frames, fields and contrasts*, pages 75–102. Lawrence Erlbaum Associates, Hillsdale, NJ.

Charles J. Fillmore, Christopher R. Johnson, and Miriam R. L. Petruck. 2003. Background to Framenet. *International Journal of Lexicography*, 16:235–250.

Charles J. Fillmore. 1985. Frames and the semantics of understanding. *Quaderni di semantica*, 6:222–254.

Maria Teresa Guasti. 1993. *Causative and perception verbs: A comparative study*. Rosenberg & Sellier, Turin, Italy.

Patrick Hanks and James Pustejovsky. 2005. A pattern dictionary for natural language processing. *Revue Française de linguistique appliquée*, 10(2):63–82.

Patrick Hanks. 2007. Metaphoricity is gradable. In Anatol Stefanowitsch and Stefan Thomas Gries, editors, *Corpus-based approaches to metaphor and metonymy*, pages 17–35. Walter de Gruyter, Berlin.

James Higginbotham. 1983. The Logic of Perceptual Reports: An Extensional Alternative to Situation Semantics. *The Journal of Philosophy*, 80(2):100–127.

Robert S. Kirsner and Sandra A. Thompson. 1976. The role of pragmatic inference in semantics: a study of sensory verb complements in English. *Glossa*, 10:200–240.

George Lakoff and Mark Johnson. 1980. *Metaphors we live by*. The University of Chicago Press, Chicago.

Alessandro Lenci, Martina Johnson, and Gabriella Lapesa. 2010. Building an Italian FrameNet through Semi-automatic Corpus Analysis. In *Proceedings of LREC 2010*, Valletta, Malta.

James Pustejovsky. 1998. *The generative lexicon*. MIT Press.

Josef Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. 2006. *Framenet II: Extended theory and practice*. http://framenet.icsi.berkeley.edu/book/book.pdf.

Representation of Changes in Valency Structure of Verbs in the Valency Lexicon of Czech Verbs*

Václava Kettnerová and **Markéta Lopatková**
Charles University in Prague
Institute of Formal and Applied Linguistics
{kettnerova,lopatkova}@ufal.mff.cuni.cz

1 Introduction

Valency behavior of verbs is so multifarious that it cannot be described by formal rules; instead, it must be listed in the form of a lexical entry separately for each verb. Prototypically, a single verbal meaning corresponds to a single valency frame. However, in many cases, semantically close uses of verbs can be syntactically structured in different ways. See the following examples:

(1) a. *Petr namazal máslo na chleba.*
b. *Petr namazal chleba máslem.*
Eng. a. Peter smeared butter on bread.
b. Peter smeared bread with butter.

(2) a. *Turisté vylezli na kopec.*
b. *Turisté vylezli kopec.*
Eng. a. Tourists climbed up the hill.
b. Tourists climbed the hill.

(3) a. *Petr řekl, že je Marie chytrá.*
b. *Petr řekl o Marii, že je chytrá.*
Eng. a. Peter said that Mary was clever.
b. 'Peter - said - about - Mary - that - (she) is - clever.'

The uses of the verb *namazat* 'to smear', *vylézt* 'to climb' and *říci* 'to say' in (1a)-(1b), (2a)-(2b) and (3a)-(3b), respectively, refer to the same situations. However, they differ in their respective valency frames. We discuss three typologically different changes in verbal valency structure. We propose their adequate representation in the valency lexicon of Czech verbs, VALLEX.[1]

The VALLEX lexicon uses as its theoretical background the *Functional Generative Description* (henceforth FGD). In FGD, valency is related to the tectogrammatical layer, i.e., a layer of linguistically structured meaning, (Sgall et al., 1986; Panevová, 1994). Valency characteristics of a verb are encoded in a form of valency frames. The frames are modeled as a sequence of valency slots, each valency slot standing for a single valency complementation. The slots consist of (i) a functor (rather coarse-grained tectogrammatical role labeling the relation of a complementation to a verb), (ii) information on obligatoriness, and (iii) a list of possible morhemic form(s), (Žabokrtský, Lopatková, 2007).

2 Situational vs. Structural Meaning

Let us turn back to examples (1), (2) and (3). The pairs of the uses *namazat* 'to smear' in (1a)-(1b), *vylézt* 'to climb' in (2a)-(2b) and *říci* 'to say' (3a)-(3b), respectively, denote the same situations, i.e., the individual situations portrayed by these uses are characterized by the same set of participants related by the same relations. We refer to this part of a verbal meaning as a *situational meaning* and to its components as *situational participants*, (Mel'čuk, 2004). The situational meaning represents such part of the verbal meaning which has not been syntactically structured yet. The part of the verbal meaning in which the components of the situational meaning is syntactically structured is referred here as a *structural meaning*. The structural meaning corresponds to the tectogrammatical layer and its components to the *valency complementations*, (Panevová, 1994).

Each lexical unit of a verb is characterized by both situational and structural meaning in a unique way: any change in the situational or structural part of its meaning leads to a change of lexical unit. We can observe that the pairs of the uses of the verbs *namazat* 'to smear' in (1a)-(1b), *vylézt* 'to climb' in (2a)-(2b) and *říci* 'to say' in (3a)-(3b), respectively, share the same situational meaning; however, they differ from each other in the structural part of meaning: The same set of sit-

*The research reported in this paper is carried out under the project MŠMT ČR No. MSM002162083. It is supported by the grant No. LC536 and partially by the grant No. GA P406/2010/0875.

[1] http://ufal.mff.cuni.cz/vallex/2.5/

uational participants are mapped onto the valency complementations in a different way, respectively, i.e., these uses of the verbs are characterized by different valency frames. It follows that the uses of the verbs *namazat* 'to smear' in (1a)-(1b), *vylézt* 'to climb' in (2a)-(2b) and *říci* 'to say' in (3a)-(3b), respectively, represent separate lexical units.

However, we observe that these separate lexical units are characterized by different types of the asymmetry in the correspondence between situational participants and valency complementations. We assume that the types of the asymmetry are closely related to characteristics of changes in a valency frame. As a consequence, these asymmetries determine a lexicographic representation of changes in a valency frame. For these reasons, the asymmetries represent a starting point in our analysis of changes in valency structure of verbs, in contrast to other approaches, e.g., (Levin, 1993).

Now let us focus on the verbs *namazat* 'to smear', *vylézt* 'to climb' and *říci* 'to say' again. On the basis of three types of the asymmetry, we determine three typologically different changes in the valency structure of these verbs: (i) semantic diatheses illustrated by the lexical units of the verb *namazat* 'to smear' in (1a)-(1b) (Section 3) (ii) multiple structural expression of a situational participant represented by the lexical units of the verb *vylézt* 'to climb' in (2a)-(2b) (Section 4) and (iii) structural splitting of a situational participant illustrated by the lexical units of the verb *říci* 'to say' in (3a)-(3b) (Section 5). We emphasize that all these changes are connected with separate lexical units, i.e., they are embedded in the lexical structure of the language.

In contrast, there are changes in valency frames which belong to the grammar structure of a language, (Apresjan, 1974). These changes are characterized by differences in the mapping of valency complementations onto surface syntactic positions while the correspondence of situational participants and valency complementations is preserved, (Kettnerová, Lopatková, 2010). With respect to the same situational and structural meaning, these changes are connected with different uses of a single lexical unit of a verb. As a consequence, different types of changes – the changes embedded in the grammar and lexical structure of a language – can be combined. For example, passive grammatical diathesis can be combined with the constructions in the relation of locative semantic diathesis, as in (4a)-(4b):

(4) a. *Máslo bylo namazáno na chléb (od Petra).*
b. *Chléb byl namazán máslem (od Petra).*
Eng. a. Butter was smeared on bread (by Peter).
b. Bread was smeared with butter (by Peter).

In this paper, we focus primarily on the changes in valency structure of verbs representing a part of the lexical structure of the language. The changes arising from the grammar structure of the language are left aside here. As the discussed changes in valency structure are based on the asymmetries in the correspondence between situational participants and valency complementations, an adequate representation of situational as well as structural meaning is necessary for the purpose of their description. However, whereas the representation of the structural meaning of verbs has been elaborated in detail in FGD, an adequate description of the situational meaning has not been worked up within this framework so far. Hence, we propose to enhance FGD (i) with lexical-semantic representation of the situational part of verbal meaning based on the lexical-conceptual structures, and (ii) with an open set of labels of situational participants.

3 Semantic Diatheses

The first type of the asymmetry in the mapping of situational participants and valency complementations can be illustrated by the uses of the verb *namazat* 'to smear' in (1a)-(1b). The situation denoted by this verb consists of three situational participants: 'Agent', 'Cover' and 'Surface'. The participants 'Cover' and 'Surface' can be mapped onto the valency complementations in two ways: 'Cover' onto PAT(ient) (1a) or EFF(ect) (1b) and 'Surface' onto DIR(ection) (1a) or PAT(ient) (1b). Thus either the participant 'Cover' or the participant 'Surface' are structured as the valency complementation PAT, which occupies the prominent syntactic position of object, see Figure 1. This fact results in a slight semantic difference between the uses in (1a) and (1b). This difference is associated with a holistic interpretation of the participant which is expressed as PAT, (Anderson, 1971; Dowty, 1991).

The asymmetry in the mapping is connected with the change of lexical unit of the verb, i.e., it is

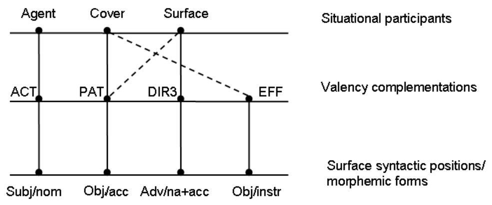


Figure 1: The asymmetry in the mapping of the situational participants and the valency complementations of the verb *namazat* 'to smear' associated with a semantic diathesis.

based on lexical-semantic means. We refer to the relations between such lexical units as semantic diatheses.[2] The changes in the valency structure of verbs associated with semantic diatheses may affect the number of valency complementations, their type and their morphemic form(s); moreover, these changes may vary even within one type of the diathesis. Thus we propose to represent separate lexical units related to a semantic diathesis by separate valency frames stored in the data component of the lexicon and to interlink them by a relevant type of the diathesis. E.g., two lexical units of the verb *namazat* 'to smear' are interlinked by the *locative semantic diathesis*.

In the grammar component, the changes in the mapping between situational participants and valency complementations are described by lexical rules based on an adequate lexical-semantic representation of the members of the semantic diathesis. For this purpose, we adopt the lexical-conceptual structure (henceforth LCS) proposed in (Rappaport Hovav, Levin, 1998). E.g., the uses of the verb *namazat* 'to smear' in (1a) and (1b) can be described by the following LCSs:

(a) [[x ACT$<_{SMEAR}>$] CAUSE [BECOME [y ON z]]]

(b) [x CAUSE [BECOME [z <*SMEARED*>]] BY MEANS OF [[x ACT$<_{SMEAR}>$] CAUSE [BECOME [y ON z]]]]

Commentary on the LCSs. LCS (a) corresponding to variant (1a) represents a change of location consisting of two subevents: (i) the first one represented as [x ACT$<_{SMEAR}>$] identifies the action of the 'Agent' x. The verb <*SMEAR*> in the subscript serves as a modifier of the action. (ii) The second part of the LCS [BECOME [y ON z]] represents the change of location of the 'Cover' y resulted from the first subevent, see the predicate CAUSE. In comparison with the LCS (a), the LCS (b) is more complex. In addition, it contains the component [BECOME [z <*SMEARED*>]] specifying the change of state of the 'Surface' z indicated as <*SMEARED*>. Relating the component [BECOME [z <*SMEARED*>]] with the whole LCS (a) indicates that this event arises as a consequence of the event identified by the LCS (a). The labels of the situational participants are associated with the position of the variables in the LCSs as follows: x $\sim$ 'Agent', y $\sim$ 'Cover', and z $\sim$ 'Surface'.

With respect to their complexity, we consider the LCS (a) as unmarked and the LCS (b) as marked. We formulate the following lexical rule Loc.r1 determining the change in the mapping between the situational participants and the valency complementations:

	LCS(a)		LCS(b)
y $\sim$ 'Cover'	PAT	$\Rightarrow_{Loc.r1}$	EFF
z $\sim$ 'Surface'	DIR	$\Rightarrow_{Loc.r1}$	PAT

Let us mention some other types of Czech semantic diatheses which can be represented in the lexicon in a similar way, i.e., by means of lexical rules determining the differences in the correspondence between situational participants and valency complementations:

Material-Product diathesis

a. *Nařezal kládu*.PAT-Material *na tři polena*.EFF-Product.

b. *Nařezal tři polena*.PAT-Product *z klády*.ORIG-Product

Eng. a. He cut the log.PAT-Material into three pieces.EFF-Product

b. He cut three pieces.PAT-Product from the log.ORIG-Material

Source-Substance diathesis

a. *Slunce*.ACT-Source *vyzařuje teplo*.PAT-Substance

b. *Teplo*.ACT-Substance *vyzařuje ze Slunce*.DIR-Source

Eng. a. The sun.ACT-Source radiates heat.PAT-Substance

b. Heat.ACT-Substance radiates from the sun.DIR-Source

Agent-Location diathesis

a. *Včely*.ACT-Agent *se hemží na zahradě*.LOC-Location

b. *Zahrada*.ACT-Location *se hemží včelami*.EFF-Agent

[2]The term diathesis generally refers to the uses of verbs characterized by shifts of some of situational participants from the prominent syntactic positions of subject or object to a less prominent syntactic position, (Kettnerová, Lopatková, 2010).

Eng. a. Bees.ACT-Agent are swarming in the garden.LOC-Location
b. The garden.ACT-Location is swarming with bees.EFF-Agent

4 Multiple Structural Expression of Situational Participant

The second type of the asymmetry in the correspondence between situational participants and valency complementations is illustrated by the uses of the verb *vylézt* 'to climb' in (2a) and (2b). The situation portrayed by this verb is composed by two situational participants: 'Agent' and 'Location'. In contrast to the semantic diatheses, this type of the asymmetry is not associated with any changes of situational participants in the prominent surface syntactic position of subject or object. Contrary, it results from two possible mappings of a single situational participant onto different valency complementations. In case of the verb *vylézt* 'to climb', the participant 'Location' is mapped either onto the valency complementation DIR(ection) (2a) or PAT(ient) (2b), see Figure 2.

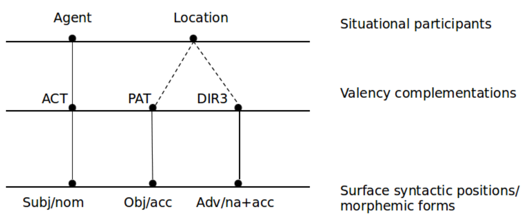


Figure 2: The multiple mapping of the situational participants 'Location' onto the valency complementations of the verb *vylézt* 'to climb'.

As in case of semantic diatheses, this type of the asymmetry, based in the lexical structure of the language, is connected with a change of lexical units of verbs. The changes in the valency structure of verbs can be described in a similar way as in case of semantic diatheses. E.g., in case of the verb *vylézt* 'to climb', separate lexical units characterized by different valency frames are stored in the data component of the lexicon. These frames are interlinked by a relevant type of the relation: In the grammar component, the lexical rule Dir.r1, based on the LCSs (c) and (d) (that correspond to (2a) and (2b), respectively), describes the change in the mapping of the situational participant 'Location'.

(c) [BECOME [x <*PLACE*>]]

(d) [BECOME [x $< PLACE_{TOP-OF} >$]]

Commentary on the LCSs. LCS (c) representing variant (2a) describes the change of location of the participant 'Agent' x. The location is identified with the constant <*PLACE*>. In contrast to variant (2b), an exact endpoint of the change of location is not implied here: 'Agent' can occur anywhere in the hill, including on the top of the hill. In the LCS (d) describing variant (2b), the constant is modified by the subscript specifying an endpoint of the change of location of 'Agent' – the top of the hill. With respect to the markedness of the constant in the LCS (d), we consider the LCS (c) as unmarked and the LCS (d) as marked. The label of the situational participant 'Agent' is attributed to the positions of the variables x in the LCSs. The situational participant 'Location' is implied by the constant <*PLACE*>.

The lexical rule Dir.r1 describing the change in the mapping of the situational participant 'Location' follows:

	LCS(c)		LCS(d)
'Location'	DIR	$\Rightarrow_{Dir.r1}$	PAT

Let us mention another type of multiple structural expression of a situational participant frequent in Czech, illustrated by the uses of the verb *umístit* 'to place'. In these uses of the verb, the situational participant 'Location' is mapped either onto the valency complementation DIR(irection) (5a) or onto the valency complementation LOC(ation) (5b). Two alternative mappings of this situational participant results in slightly different meanings: the event denoted by the first use of the verb *umístit* 'to place' in (5a) is conceived dynamically whereas the second use in (5b) is rather of static character:

(5) a. *Jana.*ACT-Agent *umístila dítě.*PAT-Patient *do jeslí.*DIR-Location
b. *Jana.*ACT-Agent *umístila dítě.*PAT-Patient *v jeslích.*LOC-Location
Eng. a. Jane.ACT-Agent placed her child.PAT-Patient into the nursery.DIR-Location
b. Jane.ACT-Agent placed her child.PAT-Patient in the nursery.LOC-Location

We assume that the above-mentioned example of multiple structural expression of the situational participant 'Location' may be described in the lexicon in a similar way, i.e., on the basis of a lexical rule determining two alternative mappings of the participant.

5 Structural Splitting of Situational Participant

The third type of the asymmetry in the correspondence between situational participants and valency complementations is illustrated by the verb *říci* 'to say' in (3a) and (3b). The situational participant 'Information' is mapped either onto one valency complementation EFF(ect) (3a) or onto two valency complementations PAT(ient) and EFF(ect) (3b), see Figure 3. We refer to these cases as a structural splitting of a situational participant.

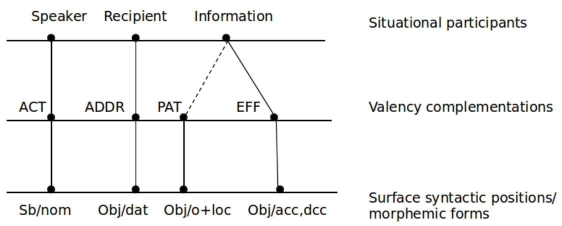


Figure 3: The structural splitting of the situational participant 'Information' of the verb *říci* 'to say'.

In case of this type of the asymmetry, the formulation of lexical rules describing changes in verbal valency structure is precluded as the coreferential relations between the split parts of a situational participant may be too complicated. See the following corpus example:

(6) *A Šaron o Arafátovi řekl, že tomuto "králi vrahů a teroristů" nikdy nepodá ruku.*
Eng. 'Sharon - about - Arafat - said - that - this - "king of murderers and terrorists" - never - shakes - hand.'
(i.e., Sharon said [about Arafat] that he (=Sharon) would never shake hands with this "king of murderers and terrorist" (=Arafat's hands).)

For this reason, we propose to capture lexical units characterized by the structural splitting of a situational participant within a single valency frame in the data component of the lexicon. The split parts of a situational participant are represented by two co-indexed valency complementations. E.g., PAT and EFF corresponding to 'Information' are co-indexed by *S* in the valency frame of the verb *říci* 'to say':

$$\boxed{\mathrm{ACT}_{obl}\ \mathrm{ADDR}_{obl}\ \mathrm{PAT}^{S}{}_{opt}\ \mathrm{EFF}^{S}{}_{obl}}$$

This valency frame explicitly describes the use of the verb *říci* 'to say' with split 'Information', as in (3b). In case of 'univalent' expression of 'Information' (3a), the situational participant corresponds to a more prominent valency complementation from the pairs of co-indexed valency complementations (e.g., in (3a) Information' is mapped onto EFF due to its obligatoriness).

In addition to the verbs of communication, the verbs denoting perception allow structural splitting of a situational participant:

(7) a. *Jana*.ACT-Perceiver *viděla, (jak Petr vchází do dveří)*.PAT-Phenomenon
b. *Jana*.ACT-Perceiver *viděla Petra*.PAT-Phenomenon, *(jak vchází do dveří)*.EFF-Phenomenon
Eng. a. 'Jane.ACT-Perceiver - saw - (as - Peter - is entering - in - the doors)'.PAT-Phenomenon
b. 'Jane.ACT-Perceiver - saw - Peter.PAT-Phenomenon - (as - (he) is entering - in - the doors)'.EFF-Phenomenon
(i.e., Jane saw Peter entering the doors.)

Then the structural splitting of the situational participant 'Phenomenon' can be described in the lexicon in a similar way; i.e., both lexical units of the verb *vidět* 'to see' in (7a)-(7b) are represented by a single valency frame with co-indexed valency complementations corresponding to the split situational participant.

6 Conclusion

We distinguished three types of changes in valency structure of Czech verbs on the basis of three types of asymmetry in the correspondence between situational participants and valency complementations. We demonstrated that these changes, embedded in the lexical structure of the language, are connected with separate lexical units. In case of semantic diathesis and multiple structural expression of a situational participant, we proposed to represent these separate units by separate valency frames interlinked by a relevant type of the relation stored in the data component of the lexicon. Then in the grammar component, lexical rules determining the changes in the mapping between situational participants and valency complementations are formulated. In case of the structural splitting of a situational participant, possible complicated coreferential relations obstruct formulating lexical rules. Hence, we propose to represent both lexical units within a single valency frame with an abbreviated notation.

References

Stephen R. Anderson. 1971. On the Role of Deep Structure in Semantic Interpretation. *Foundations of Language*, 7:387–396.

Jurij D. Apresjan. 1974. *Leksicheskaja semantika. Sinonimicheskie sredstva jazyka.* Nauka, Moskva.

David Dowty. 1991. Thematic Proto-Roles and Argument Selection. *Language*, 67(3):547–619

Václava Kettnerová and Markéta Lopatková. 2010. The Representation of Diatheses in the Valency Lexicon of Czech Verbs. In H. Loftsson, E. Rögnvaldsson, S. Helgadóttir, eds.: *Proceedings of IceTAL 2010.* LNAI 6233, Berlin Heidelberg, Springer Verlag. 185–196.

Beth C. Levin. 1993. *English Verb Classes and Alternations: A Preliminary Investigation.* University of Chicago Press, Chicago, IL.

Igor A. Mel'čuk. 2004. Actants in Syntax and Semantics. *Language*, 12:1–66.

Jarmila Panevová. 1994. Valency Frames and the Meaning of the Sentence. In P. A. Luelsdorff, ed.: *The Prague School of Structural and Functional Linguistics.* John Benjamins Publishing Company, Amsterdam, Philadelphia. 223–243.

Malka Rappaport Hovav and Beth Levin. 1998. Building Verb Meanings. In M. Butt and W. Geuder, eds.: *The Projection of Arguments. Lexical and Compositional Factors.* SLI Publications, Stanford, California. 97–134.

Petr Sgall and Eva Hajičová and Jarmila Panevová. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Reidel, Dordrecht.

Zdeněk Žabokrtský and Markéta Lopatková. 2007. Valency Information in VALLEX 2.0: Logical Structure of the Lexicon. *The Prague Bulletin of Mathematical Linguistics*, 87:41–60

Verbal Valency in the MT Between Related Languages

Natalia Klyueva
Inst. of Formal and Applied Linguistics
Charles University in Prague
kljueva@ufal.mff.cuni.cz

Vladislav Kuboň
Inst. of Formal and Applied Linguistics
Charles University in Prague
vk@ufal.mff.cuni.cz

Abstract

The paper analyzes the differences in verbal valency frames between two related Slavic languages, Czech and Russian, with regard to their role in a machine translation system. The valency differences are a frequent source of translation errors. The results presented in the paper show that the number of substantially different valency frames is relatively low and that a bilingual valency dictionary containing only the differing valency frames can be used in an MT system in order to achieve a high precision of the translation of verbal valency.

1 Introduction

Numerous experiments, such as Česílko (Hajič et al., 2000) and Apertium (Sánchez et al., 2007), with the machine translation (MT) between related languages support the claim that direct (word for word or phrase for phrase) methods guarantee better translation quality than complicated MT architectures. The more related the source and target languages are, the better the results provided by simple direct methods. Very closely related languages have similar morphological and syntactic properties, their lexicon usually also demonstrates a great number of similarities not only with regard to the lexical values, but also to important phenomena as e.g. the valency. For the translation of those languages it is therefore possible to ignore valency completely, because the system can rely on the similarity (or even identity) of valency frames of corresponding words and thus it is possible to translate expressions from individual valency slots directly, as e.g. in the Czech-to-Slovak MT system Česílko.

The languages which belong to the same language group, but which are not as closely related constitute a greater challenge, they require a different treatment of the verbal valency. In subsequent sections of this paper we present an examination of differences between Czech (a western Slavic language) and its Eastern Slavic counterpart, Russian.

Experiments in automatic extraction of verbal valency frames from different resources were carried out by many researchers. One of the first attempts was made in early 90's by (Rosen et al., 1992) where the process of English verb frame derivation from a learner's dictionary is described. The similar goal for extracting verb frames for both Czech and English was set in a research by (Bojar et al., 1984). Valency frames were extracted automatically from a parallel treebank PCEDT, resulting in a list of verbs and their modifications.
To the best of our knowledge such experiments were not carried out for related languages.

2 Existing resources

Manually built and handchecked dictionaries of verbal valency frames exist both for Czech and Russian. Vallex (Žabokrtský et al., 2007) is a lexicon of Czech valency frames having its roots in FGD (Functional Generative Description) theory. For Russian language, verbal valency frames can be found in the TKS (Tolkovo-Kombinatornyj Slovar - Explanatory combinatorial dictionary) – cf. (Mel'čuk, 1984). The lexicon of TKS is based on a Meaning-Text theory, it contains rich syntactic and semantic information for lexical entries of all parts of speech.
The formalisms on which Vallex and TKS are based are different in many ways, therefore it is almost impossible to map the entries from both dictionaries directly.

The first attempt to achieve a high quality MT between Czech and Russian, the transfer-based system Ruslan, was carried out in 80's (Oliva, 1992). This project left a valuable resource in a form of a bilingual dictionary that includes various kinds of information necessary for lexical, morphological, syntactic and semantic transfer. In our current work we use only morphological and syntactic information from this dictionary.

Another system we work with is an MT system between closely related languages Česílko (Hajič et al., 2000), which uses a direct word-for-word (and tag-for-tag) translation. Initially the system translated between Czech and Slovak languages reaching rather high quality, as the two languages are very closely related. When other languages from Slavic group - Polish, Lithuanian and Russian - were included into the system, it became evident that some additional shallow syntactic rules must be used.

3 Valency

Valency frame of a verb contains syntactic and semantic information crucial for proper analysis and synthesis of a sentence. In our work we will use a notion of a valency frame at the level of shallow syntax, we will not take into consideration deep syntactic structure. So we avoid such terms as Actor, Patient, Recipient etc., and we use rather surface forms of the verbal actants - cases: Nom, Gen, Dat, Acc, Ins, Loc for which we use shortcuts n, g, d, a, i, l respectively. Our work is carried out on the two Slavic languages, Czech and Russian, and for the sake of simplicity we partly follow the representation of verb structure used in the MT system Ruslan. In addition we use the following terms added for the present experiments (Czech case is always listed first, followed by a Russian one enclosed by brackets):

Simple frame constituents:

n(n) means that Czech nominative case corresponds to the same case in Russian.

a(d) means that whereas accusative form is used in Czech, Russian uses dative case.

Frame constituents including prepositions:

s(i,s(i)) means that the Czech preposition *s* (with) requires an instrumental case in Czech and the same situation holds for Russian language.

Other constituents:

(inf(inf)) means that both languages use infinite form of an additional verb as a valency constituent.

A translation valency frame therefore consists of a set of simple and/or prepositional or other constituents for both Czech and Russian. Example:

trvat|**(n(n),na(l,na(l)))**|nastaivat' - to insist

3.1 Dictionary of Ruslan

Dictionary entries in Ruslan contain morphological, syntactic and semantic information. In the first stage of our study we do not make use of semantic features, leaving it for future experiments.

The dictionary has 10023 entries, 2080 of which are verbs. Let us now present two examples of original dictionary entries from Ruslan, one for a noun and one for a verb:

NA2PAD==H(@(*A),FI1023, IDEJA) - *idea.*

H represents a nominal declension class(hrad).

DOBE3H==R(5,TI,?(N(N)),D2,KONC2IT6SJA)

- *to finish running*

R represents a verb, 5,TI - conjugation type (tisknout), (N(N)) - the valency frame of an intransitive verb with a single slot for a subject in nominative case in both languages.

,D2,KONC2IT6SJA - conjugation class + Russian lexical equivalent of the verb.

4 Classification of valency frames

Out of the 2080 verbal dictionary entries from Ruslan we have analyzed 1856 unique verbs. The reason of this difference is the fact that the original dictionary contains a number of verbal pairs with identical valency frames, usually two variants of a Czech lemma in the present and past tense. We made a classification of how the Czech valency frames correspond to the Russian ones. We have sorted verbs on the basis whether the verb requires the prepositional case or the simple one. The most important categories of verbs are those showing differences between both languages - these verbs will serve as a basis of a list of verbs with different valency frames which will be used for an improvement of our experimental MT system. The subsequent subsections describe examples for all analyzed categories of words.

4.1 Equal simple frame constituents

Cases when a verb have an actant structure without a preposition and Czech and Russian frames correspond to one another:

vyzývat|(n(n),a(a) or n(n))|vyzyvat' - to call

The most typical sequence of frame patterns is **n(n),a(a)** , which represents simple transitive verbs. 1317 (70 % of all verbs) have this structure. The fact that Czech and Russian have practically the same number of cases that are meaningful [1]

[1] Vocative case is not used in modern Russian unlike in Czech

Table 1: Case correspondences

Cs/Ru	Nom	Gen	Dat	Acc	Ins
Nom	3070	8	10	6	3
Gen	0	25	0	4	0
Dat	0	3	178	7	0
Acc	3	19	12	1388	7
Ins	5	0	0	3	1355

when speaking of verb valency makes the comparison easier and it apparently also influences the number of identical frames.

4.2 Different simple frame constituents

The first group of verbs that will form our list of verbs having different valency frames in both languages are those translation pairs in which Czech and Russian verbs govern different simple cases:
vyžadovat|(n(n),**a(g)** or n(g),i(i))|trebovat'
- to demand, Acc in Czech, Gen in Russian:
povšimnout|(n(n),refl(si),g(a))|zametit'
- to notice, Gen in Czech, Acc in Russian
rušit|(n(n),**a(d)** or n(d),i(n))|mešat'
- to disturb, Acc in Czech, Dat in Russian
hýbat|(n(n),a(a),**i(a)**)|dvigat'
- to move, Ins in Czech, Acc in Russian

Table 1 presents the statistics of simple frame patterns giving a picture of how simple cases in Czech and Russian mutually correspond. [2]

As we can see from the table, Czech and Russian non-prepositional valency slots have usually identical cases, the list of verbs exhibiting differences is very short.

4.3 Equal prepositional frame constituents

Verbs in this class have the valency slots containing prepositions. We have considered the translation frames to be equal in a case when prepositions are translated straightforwardly or typically from Czech into Russian. The problem is to set a border between typically translated prepositions and those translated differently. This issue lies outside of the scope of our study. We have used the data from (Nadykta, 2007), in which the author addresses in detail many aspects of Czech and Russian prepositions. Following are verbs and frames that constitute a typical translation of each other according to our criteria:
do(g,v(a)):ponořit|(n(n),do(g,v(a)))|pogruzit' - to sink into
z(g,iz(g)):vycházet|(n(n),z(g,iz(g)))|vychodit' - to go out from

[2] Locative case is not included as it is governed by a preposition in both languages.

4.4 Different prepositional frame constituents

To select verbs that have different prepositional frames we just excluded verbs with similar frame patterns described in the previous section. 104 (5.6 %) of verbs belong to this group. Below are some examples of such verbs:
záležet|(n(n),**na(l,ot(g))**)|zaviset' - to depend on
narazit|(n(n),**na(a,s(i))**)|stolknut'sja - to face
We also define some special cases which are irrelevant from computational point of view, as they will be processed as the common cases. They may still be of some interest to theoretical study of verb valency differences.
Those special cases form a rather small group of verbs that:
1. they are followed by an infinitive:
přestat|(n(n) or inf(inf) or **v(l,inf)**)|perestat' - to stop + inf
2. they govern identical prepositions that have different case:
klást|...**před(a,pered(i))** or na(a,na(a))| klast' - to put behind
3. they govern a preposition in one language, while in the other a simple case is used:
vystačit|(n(d),**s(i,g))**|chvatit' - to be enough

5 Statistics of Valency Difference List

The main output of our work is a list of verbs that have different valency structure in Czech and Russian. Table 2 shows the statistics of those verbs with regard to our classification on simple and prepositional case frames.

Table 2: Types of valency frames incorrespondences

Type of difference	N of verbs	Percentage
Simple case	68	3.6%
Prepositional case	104	5.6%
Total	1856	100%

6 Evaluation

In this section we present a semi-manual evaluation of our list of verbs carried out on sentences translated by the Česílko MT System. In the process of MT evaluation we have evaluated only parts of sentences that include a verb and its arguments and we have determined whether our data might improve the translation. The test did not evaluated overall translation quality due to the observation that because of the overall imperfection of the system there are many other errors that have greater influence on the translation quality and which would bias the evaluation of our experiment. We aim primarily at an estimation to which extent the knowledge of differing valency frames ultimately might improve the translation quality by its own, not in combination with other phenomena. We are actually aiming at a kind of upper boundary of possible improvement.

The evaluation was carried out on a relatively small sample of 100 sentences translated from Czech into Russian.

As mentioned above, we have evaluated not the whole sentences, but smaller units. In accordance with (Lopatková et al., 2009), we took linguistically motivated units (segments) containing only one finite verb. This made it easier to analyze valency issues of concrete verbs. This approach was motivated by the fact that in complex sentences it might be difficult to define a verb and its arguments when a clause is divided into two or more parts by an embedded segment, and a verb is situated in another part of a sentence than its dependent arguments:

Mnozí provozovatelé ***považovali*** *naši shůzku, k níž došlo bezprostředně po konferenci v Anapolisu, kde se sešli představitelé všech arabských států včetně Sýrie a Izraele,* ***za projev*** *nevůle...*

(Many observers considered our meeting which took place immediatelly after the conference in Anapolis, where the deputies of all Arabic states including Syria and Israel met, to be a manifestation of ill will...)

In the evaluated phrase the verb *považovat* and its dependent prepositional construction *za projev* stands more than 20 tokens from one another, and could not be analyzed properly without breaking a sentence into several less complex segments.

The evaluation process was performed in several steps:

Table 3: Errors in verbal valency

mistakes	34	12,45 %
improvements	16	5,86 %
Total No. of verbs	273	100 %

1. Detect segments of sentences with Czech verbs with different valency structure

2. Determine whether the verbs and their arguments have been translated into Russian by the MT system in a correct way

2b. ...and whether or not adding our Valency DATA can improve the translation quality (Sometimes even this will not help because of the totally different structure)

The table 3 describes the results of the evaluation: the **mistakes** column presents a number of incorrectly translated verbal valency constructions, the **improvements** column shows the number of cases where our valency list could have helped to achieve better results.

The table shows that errors in verbal valency occur in slightly more than 10 % of all verbs. Almost half of those mistakes can be captured by our list of valency differencies that contains most frequent verbs. Here comes an example of an error in MT, that can be improved:

*pokračovat v diplomatických snahách.*LOC(cz) (continue diplomatic attempts)
**prodolzhat' v diplomatičeskich popytkach.*LOC(ru - Česílko MT)(v + loc)
*prodolzhat' diplomatičeskie popytki.*ACC(ru - improved)

The verb *pokračovat* - to continue - in Czech has as its arguments the preposition *v* + noun in locative case, the entry from our data (pokračovat (v(l,a)) prodolzhat') will make sure that a noun in accusative case will follow the verb in Russian.

7 Conclusion

In this article we have shown that the number of really different verbal valency frames between Czech and Russian is relatively low and that instead of using a complete bilingual valency dictionary it is reasonable to create only a list of differences and to translate the remaining verbs and their constituents in a default manner. We have also evaluated the expected impact our data will have on translation of verbs and their arguments. This evaluation shows that although the valency dictionary will definitely improve the translation quality, it's influence is relatively small and it will be necessary to investigate also other phenomena in order to achieve a more substantial improvement.

Nevertheless, this experiment has also brought interesting results from the linguistic point of view. It shows that in the future it might be possible to translate both existing valency dictionaries for Czech and Russian and compare them. This might bring an enrichment of the frames contained in both dictionaries. The extension of our list of differences will then come as a side effect of this process.

Acknowledgments

This research is supported by grants MSM0021620838 and GA405/08/0681.

References

Hajič, J., Hric, J., and Kuboň, V. 2000 *Machine translation of very close languages*. In Proceedings of the 6th Applied Natural Language Processing Conference

Sánchez-Martínez, Felipe and Forcada, Mikel L. 2007 *Automatic induction of shallow-transfer rules for open-source machine translation*. In Proceedings of the 11th Conference on Theoretical and Methodological Issues in Machine Translation 181-190

Alexandr Rosen, Eva Hajicová, Jan Hajic. 1992 *Derivation Of Underlying Valency Frames From A Learner's Dictionary.*, COLING 1992, 553-559

Lopatková, M., Klyueva, N., Homola, P. 2009 *Annotation of Sentence Structure; Capturing the Relationship among Clauses in Czech Sentences.* In Proceedings of the Third Linguistic Annotation Workshop, Law III, ACL-IJCNLP 2009 Conference 74-81

Karel Oliva. 1989 *A Parser for Czech Implemented in Systems Q.* in Explizite Beschreibung der Sprache und automatische Textbearbeitung, MFF UK Prague, 1989.

Igor Mel'čuk and Alexander Zholkovsky. 1984. *Explanatory Combinatorial Dictionary of Modern Russian.* Vienna: Wiener Slawistischer Almanach

Ondřej Bojar and Jan Hajič. 2005. *Extracting Translation Verb Frames.* Proceedings of Modern Approaches in Translation Technologies, workshop in conjunction with Recent Advances in 2-6.

Zdeněk Žabokrtský and Markéta Lopatková. 2007. *Valency Information in VALLEX 2.0: Logical Structure of the Lexicon.* Prague Bulletin of Mathematical Linguistics, (87):41-60, 2007.

Nadzeya Nadykta. 2007. *The Use of Prepositions in Czech and Russian according to Parallel Corpus Data.* Diploma, Praha

Representation of verbal event structure in sign languages

Evie Malaia
Indiana University
Bloomington, IN, USA
emalaia@indiana.edu

Ronnie B. Wilbur
Purdue University
West Lafayette, IN, USA
wilbur@purdue.edu

Abstract

Sign languages recruit physical properties of visual motion to convey linguistic information. The present experiment investigated the effect of sign position and grammatical aspect on kinematic parameters of predicates in American Sign Language (ASL) and Croatian Sign Language (Hrvatski Znakovni Jezik, HZJ) using motion capture data. Kinematic features of signs recorded on the dominant hand were affected by both grammatical aspect of the predicate, and its position within the sentence. The study demonstrates independent, but interactive effects of grammar and prosody on kinematic parameters of signs, and provides cross-linguistic confirmation that physical properties of articulator motion are recruited in sign languages to express linguistic features.

1 Introduction

Humans perceive and conceptualize reality in terms of discrete events, and use linguistic labels – verbs, or predicates – to denote these events. Event boundaries represented by the predicate have long been of interest to linguistic theory as possible semantic primitives (Dowty, 1979; Jackendoff, 1991; Pustejovsky, 1991; Ramchand, 2008; van Hout, 2001; Van Valin, 2007; Vendler, 1967; Verkuyl, 1972). Predicates denoting events with an inherent boundary representing a change of state (break, appear) are considered semantically telic, as opposed to predicates describing homogenous – atelic – events, such as swim or sew. These predicate properties are also known as event structure template, or aktionsart. Predicate telicity, or linguistic representation of event boundary, has been shown to affect syntactic structure of the sentence in spoken languages (Ramchand, 2008; Tenny, 1994), and thematic role assignment in online sentence processing (Malaia, Wilbur, & Weber-Fox, 2009).

Sign languages recruit physical properties of visual space and motion to convey linguistic information. Prior research has demonstrated that kinematic (motion-related) parameters are utilized for expression of linguistic features in a regular manner, both within linguistic modules and at their interfaces (Brentari, Gonzalez, Seidl, & Wilbur, in press). A growing body of research in psychology also indicates that perceptual segmentation of reality into discrete events is determined by kinematic properties of the scene, namely speed and acceleration in the motion of actors (Zacks, Kumar, Abrams, & Mehta, in press; Zacks, Swallow, Vettel, & McAvoy, 2006).

Interestingly, despite their mutual unintelligibility, sign languages (SLs) appear to be more similar to each other than spoken languages are (Newport & Supalla 2000). Sign components, especially for predicate signs, are grammaticalized from universally available physics of motion and geometry of space, which are therefore fundamentals on which more advanced meanings can be constructed (Wilbur 2003, 2005, 2008.) Cross-linguistic research on SLs can provide an explanation for their apparent visual similarity. At the same time, SL signs are grammaticalized units of meaning, which have to be learned as part of the linguistic system, and that distinguishes them from conventional gestures or pantomime.

Wilbur (2003) made the linguistic observation that ASL lexical verbs could be analyzed as telic (denoting a change of state, such as throw, fall) or atelic (denoting homogenous activities, such as swim, walk) based on their kinematic parameters: telic verbs appeared to have a sharper ending movement to a stop, reflecting the semantic end-state of the affected argument. The observation that semantic verb classes are characterized by certain movement profiles was formulated as the Event Visibility Hypothesis (EVH).

Crosslinguistic quantitative research into event structure expression in SLs then became necessary to provide an insight into the interface between (possibly) language-independent perceptual cues of event structure used in SLs, and linguistic systems of different unrelated SLs.

HZJ presented an especially interesting case for investigation of event structure expression in sign kinematics, by virtue of being unrelated to ASL (on the basis of which EVH had been formulated), and having a member of Slavic language family as its spoken substrate. Slavic languages are characterized by a conflation of internal (event structure) and external (viewpoint) aspect within their lexicon, leading to fusion of temporal and aspectual domains in verbal predicates, equivalent to overt morphological specification of verbal event structure (Bertinetto, 2001; Borik, 2006; Filip, 1999). Hence, two unrelated sign languages (ASL and HZJ) were chosen in order to investigate kinematic parameters of both lexical (ASL) and grammatical (HZJ) expression of event structure.

2 Grammatical and prosodic markers in sign languages

In comparison on spoken languages, SLs are more likely to use simultaneous means of expressing grammatical markers ("layering"; Wilbur 2000). This strategy compensates for the longer time needed to articulate a sign compared to a spoken word. For example, adverbs can be made using lower face configurations while a verb is being signed on the hands.

Brentari (1998) demonstrated that the Prosodic Hierarchy, which is based on increasing breaks in rhythmic structure, is valid for sign languages: syllables contribute to prosodic words which combine into prosodic phrases which combine into intonational phrases. There is increasingly obvious Phrase Final Lengthening at these domains. Furthermore, the sign and pause durations are affected by signing rate (Wilbur 2009a).

In contrast to such rhythmic marking, components that are held in position from the beginning to the end of a domain generally mark the scope of syntactic and semantic operators. In ASL. Two such markers are lowered brow for wh-questions, and headshake for negation.

Beyond measuring sign and pause durations, previous investigations of sign kinematics have been lacking in quantitative measures, as motion capture equipment has only recently become more available.

3 Data collection and analysis

Various tests have been used in the literature to demonstrate that telicity is a relevant linguistic notion reflected in the grammatical system. The most widely used tests for spoken languages include the temporal adverbial modification test (Dowty, 1979; Verkuyl, 1972), and the conjunction test. Additionally, 'almost' modification has been used as a test in sign language research to identify telic predicates (Smith, 2007).

For the purposes of our study, a group of 50 ASL signs were tested in an interview with a native ASL signer/ linguistic consultant. The native signer's intuitions were elicited in the adverbial modification test, the conjunction test, the 'almost' modification test, and STOP/FINISH combinability test. Telicity of the predicates was established based on results of elicitation. For the adverbial modification test, ASL predicate signs were considered telic if they combined with 'IT TOOK AN HOUR'[1], and atelic if they combined with 'FOR AN HOUR'. Additionally, if the predicate combined with the adverbial meaning 'ALMOST (implemented as an adverbial, or as a modification of the formation of the sign's movement) yielding the meaning of "one did not complete doing X", we interpreted this as presence of end-point (which was not reached) in the event structure of the predicate. If the predicate combined with 'ALMOST' meant *only* "one did not start doing X", the predicate was considered atelic; as expected, some of the telic predicates allowed both interpretations.

For the conjunction test, we tested the predicates' meaning in the sentence 'she did V(erb) on Sunday and on Monday'. If the sentence was interpreted as denoting two discrete events, the predicate was considered telic; if the sentence referred to one long event, the predicate was considered atelic.

Finally, the predicates were examined for combinability with the signs FINISH and STOP. In cases where the predicate combined with FINISH with the 'completive' meaning (Fischer & Gough, 1999), it was interpreted as

[1] ASL expression best transcribed as 'IT TOOK AN HOUR' is equivalent to "in an hour"-type adverbials in spoken English, the temporal modifiers specifying the time elapsed to a referenced time-point.

having an inherent end-point (i.e. telic); if the predicate did not combine with FINISH meaning 'completed', but only with FINISH meaning 'already, in the past', and/or instead could only be combined with STOP, it was considered an atelic predicate. Results of these linguistic tests were then combined in order to classify the predicate as either telic or atelic. When telicity interpretations differed between the four tests for one predicate, signaling possibility of frame structure alternation (Levin, 1993) the predicate was eliminated from the final set of 40 stimuli, which included 24 telic and 16 atelic signs.

The following ASL predicates, which were identified as belonging to telic or atelic classes based on the results of all four linguistic tests, were selected for investigation:
Telic predicates (N=24): STING, THROW, HIT, PLUG-IN, APPEAR, CATCH-UP, OPEN-DOOR, RUIN, EAT-UP, CHECK, TAKE-FROM, ZIP, CLOSE-DOOR, SEIZE, DISAPPEAR, ARREST, BECOME, LOOK-AT, ARRIVE, DIE, RELAX, STEAL, SUGGEST, SHUT-DOWN-COMPUTER
Atelic predicates (N=16): TRAVEL, RIDE-IN, COLLECT, LIVE, PROCEED, SHAVE, FOLLOW, WRITE, STAY, INTERRUPT, DRAW, SEW-WITH-MACHINE, SEND, HAVE, INVESTIGATE, SWIM.

For the study of predicate production in HZJ, 120 imperfective-atelic Croatian verbs and 120 of their perfective counterparts were translated into HZJ in order to identify the mechanisms of temporal-aspectual category expression (Milkovič & Malaia, 2010). 3 major groups of temporal-aspectual verb pairs were identified. The largest group (104 signs) formed temporal-aspectual verb pairs based on the properties of sign kinematics: telic (perfective) signs in this group were formed by using shorter, sharper movement, as compared to atelic-imperfective roots. The second group did not allow formation of telic (perfective) signs from atelic-imperfective roots; the third group allowed formation of telic-perfective signs by suppletive means, including quantification of the internal argument, and use of verbal complements. A subset of 30 temporal-aspectual sign pairs from the first group was selected for further investigation using motion capture recording (see Table 1).

For the motion capture study of ASL predicates, 24 telic and 16 atelic signs were randomized, and elicited from 6 participants in the following linguistic conditions: in isolation, in the carrier phrase 'SIGN X AGAIN', sentence-medially 'SHE X TODAY', and in sentence-final position 'TODAY SHE X'. The conditions were the same for all participants: after completing a practice trial, they saw the stimuli in the same order, and signed to the camera while standing. One production per condition was collected for each signer (thus, we recorded 160 productions per signer for six signers). For motion capture study of HZJ, one participant followed the same protocol on 5 separate days of recording. A simultaneous video recording at 30fps rate was made with a NTSC video camera on a tripod outside the motion capture recording field. The positional data from the marker on the right wrist, tracking the movement of the dominant signing hand, was used for the analysis. Both the video and the 3-D positional data were imported into ELAN annotation software, and aligned using the audio marker and T-pose (the signer standing with hands extended to the sides at shoulder level) at the beginning and end of each recording. The video was annotated in ELAN by a native ASL signer, who marked the beginning and end of each target sign following procedures established by (Green, 1984), assuming the first frame of recognition of the sign-initial handshape as the beginning of each predicate, and either the point of contact, or maximal distance traveled by the hand, as the end of the sign. Thus, the onset and the ending of each sign were defined linguistically based solely on the video cues, without access to kinematic variables. The time points for the beginning and end of each sign were extracted from ELAN annotation of the video data, and processed in MATLAB to extract speed and acceleration profiles for each predicate from the recorded kinematic files.

The kinematic metrics for analysis were selected based on previous investigations in linguistics and psychology. Prior research in event perception has suggested that movement speed and acceleration/deceleration are the markers which enable humans to segment meaningful event from continuous reality (Zacks, Kumar, Abrams, & Mehta, in press). Event Visibility Hypothesis (Wilbur, 2003) proposed that sign languages denote event structure by the slope of deceleration from peak velocity to the end of the sign, which leads to concomitant changes in other kinematic properties of the sign – namely, sign duration, peak velocity, and timing of peak velocity within the predicate.

Table 1. Croatian verbs used as stimuli, and their English translations.

Imperfective form	English Translation	Perfective form	English Translation
buditi	to be waking up	probuditi	to wake up
putovati	to be travelling	otputovati	to take off
putovati	to be travelling	doputovati	to arrive
gledati	to be looking at	ugledati	to spot, to notice
gurati	to be pushing	gurnuti	to give a push
brisati	to be wiping	obrisati	to wipe off
crtati	to be drawing	nacrtati	to draw up
češljati	to be combing	počešljati	to comb through
čistiti	to be cleaning	očistiti	to clean up
čitati	to be reading	pročitati	to read through
dijeliti	to be dividing	podijeliti	to split
brijati	to be shaving	obrijati	to shave
bježati	to be fleeing	pobjeći	to run away
disati	to be breathing	udahnuti	to breathe in
dizati	to be lifting	dignuti	to pick up
dolaziti	to be coming	doći	to show up
donositi	to be carrying	donijeti	to bring
dopuštati	to tolerate	dopustiti	to permit (once)
dovoditi	to be bringing (someone)	dovesti	to bring (to someplace)
dovoziti	to be driving	dovesti	to drive up
govoriti	to be speaking	reći	to tell
gristi	to be biting	ugristi	to bite (someone)
gubiti	to be losing	izgubiti	to have lost
iskorištati	to be exploiting	iskoristiti	to take advantage of
oblačiti-se	to be dressing	obući-se	to put clothes on
odgovarati	to be responding	odgovoriti	to answer
prodavati	to be selling	prodati	to sell
propadati	to be decaying	propasti	to fail
birati	to be choosing	izabrati	to pick
grmjeti	to be thundering	zagrmjeti	to thunder

Based on these proposals, the following metrics were calculated for each verb sign:

a) the duration of the sign in milliseconds (duration);
b) peak instantaneous speed achieved within each sign (maxV);
c) the percent of sign movement elapsed to the moment where peak speed occurred (% elapsed), which is also the point at which deceleration starts,
d) minimum instantaneous negative acceleration (i.e. maximal deceleration) within each sign (minA);
e) the slope of deceleration, calculated as the difference between maxV and the following local minimum, divided by the number of milliseconds over which it occurred. The slope measured the overall steepness of the deceleration from maxV to the following minimum velocity, whereas minA measured the maximum instantaneous negative acceleration (deceleration).

Multivariate analysis of variance (MANOVA GLM) was conducted to determine the effect of each independent factor (Predicate, Position) and their interaction (Predicate x Position) on each of the dependent kinematic variables; the results for ASL and HZJ are presented in Tables 1 and 2, respectively.

4 Results

Kinematic features of verb signs were affected both by the verb type, and by its position within the sentence in a regular manner. Statistical analysis demonstrated regular kinematic distinctions between verb classes.

Table 2. Significant effects of Predicate Type and Position on ASL signs

Kinematic variable	Predicate Type			Position			Predicate Type x Position		
	F (1,916)	$p<$	η_p^2	F (1,916)	$p<$	η_p^2	F (1,916)	$p<$	η_p^2
duration	11.036	.001	.012	29.573	.001	.031			
maxV	78.301	.001	.079	13.092	.001	.014			
% elapsed	4.393	.036	.005	4.323	.038	.005	4.099	.043	.004
slope	29.645	.001	.031						
minA	52.614	.001	.054						

Table 3. Significant effects of Predicate Type and Position on HZJ signs

Kinematic variable	Predicate Type			Position			Predicate Type x Position		
	F (1,1170)	$p<$	η_p^2	F (1,1170)	$p<$	η_p^2	F (1,1170)	$p<$	η_p^2
duration	68.375	.001	.055	31.292	.001	.026			
maxV	641.448	.001	.354						
% elapsed	28.925	.001	.024	22.288	.001	.019			
minA	356.863	.001	.234	6.522	.011	.006			
slope	306.2	.001	.207	8.886	.003	.008	4.58	.033	.004

In ASL, measures of deceleration (slope, minA), and in HZJ - peak velocity, were robust to the prosodic effect of Phrase Final Lengthening. The findings showed that Event Visibility in kinematic parameters, demonstrated at the lexical level in ASL verbs, can be grammaticalized in sign languages, such as HZJ. The latter allows formation of temporal-aspectual verb classes from the same sign root, such that rapid deceleration following peak velocity constitutes a morphemic affix similar to those observed for various aspectual purposes, e.g. different types of reduplication (Wilbur, 2005, 2009b).

5 Conclusion

The motion capture data on sign production in two unrelated sign languages demonstrates that the final part of syllables in predicate signs denoting bounded (telic) events is marked by a rapid deceleration at the end of the sign, made even more prominent by higher peak velocity, as compared to verb signs denoting unbounded (atelic) events.

The two experiments show independent and interactive effects of grammar and prosody on kinematic parameters of verb sign, providing cross-linguistic confirmation that physical properties of articulator motion are recruited in sign languages to express linguistic features.

From the standpoint of linguistic theory, the significance of the finding that kinematics of sign production map onto event structure representation has implications for modeling the syntax-semantics interface in both signed and spoken languages. From the standpoint of computational linguistics, the evidence that minimal semantic feature (such as telicity) can affect multiple parameters of the sign's kinematic pattern, which merge the semantic and syntactic levels of a sign with its phonological level, can be utilized for machine translation of signed languages (cf. Malaia, Borneman & Wilbur, 2008).

Reference

Bertinetto, P. M. (2001). On a Frequent Misunderstanding in the Temporal-Aspectual Domain: the 'Perfective = Telic' Confusion. In C. Cecchetto, G. Chierchia & M. T. Guasti (Eds.), Semantic Interfaces [Reference, Anaphora and Aspect] (pp. 9-42). Stanford: CSLI Publications.

Borik, O. (2006). Aspect and Reference Time: Oxford University Press.

Brentari, D. (1998). A prosodic model of sign language phonology. Cambridge, MA: The MIT Press.

Brentari, D., Gonzalez, C., Seidl, A., & Wilbur, R. B. (in press). Sensitivity to visual prosodic cues in signers and nonsigners. Language and Speech.

Dowty, D. (1979). Word Meaning and Montague Grammar. Dordrecht: D. Reidel.

Filip, H. (1999). Aspect, Situation Types, and Noun Phrase Semantics. New York/London: Garland Publishing, Inc.

Green, K. (1984). Sign boundaries in American Sign Language. Sign Language Studies, 42, 65-91.

Jackendoff, R. (1991). Parts and boundaries. Cognition, 41(1-3), 9-45.

Malaia, E., Borneman, J., & Wilbur, R. B. (2008). Analysis of ASL Motion Capture Data towards Identification of Verb Type. In J. Bos & R. Delmonte (Eds.), Semantics in Text Processing. STEP 2008 Conference Proceedings (pp. 155-164): College Publications.

Malaia, E., Wilbur, R. B., & Weber-Fox, C. (2009). ERP evidence for telicity effects on syntactic processing in garden-path sentences. Brain and Language, 108(3), 145-158.

Milkovic, M., Malaia, E. (2010) Event visibility in Croatian Sign Language: separating Aspect and Aktionsart. *Theoretical Issues in Sign Language Research-10,* Purdue University, IN, USA

Newport, E. & T. Supalla (2000). Sign language research at the millennium. In Lane, H. and Emmorey, K. (eds.), The signs of language revisited: Festschrift for Ursula Bellugi and Edward Klima, pp. 103-114. Hillsdale, NJ: Lawrence Erlbaum.

Pustejovsky, J. (1991). The Syntax of Event Structure. Cognition, 41(1-3), 47-81.

Ramchand, G. (2008). Verb Meaning and the Lexicon - The First-Phase Syntax: Cambridge University Press.

Tenny, C. (1994). Aspectual roles and the syntax-semantics interface. Dordrecht: Kluwer Academic Press.

van Hout, A. (2001). Event Semantics in the Lexicon-Syntax Interface. In C. Tenny, Pustejovsky, J. (Ed.), Events as Grammatical Objects (pp. 239-282): CSLI Publications.

Van Valin, R. (2007). Some universals of verb semantics In R. Mairal & J. Gil (Eds.), Linguistic Universals): Cambridge University Press.

Vendler, Z. (1967). Linguistics in Philosophy. Ithaca, NY: Cornell University Press.

Verkuyl, H. (1972). On the Compositional Nature of Aspects. Dordrecht: Reidel.

Wilbur, R. B. (2000). Phonological and prosodic layering of non-manuals in American Sign Language. In Lane, H. and Emmorey, K. (eds.), The signs of language revisited: Festschrift for Ursula Bellugi and Edward Klima, pp. 213-241. Hillsdale, NJ: Lawrence Erlbaum.

Wilbur, R. B. (2003). Representation of telicity in ASL. In Chicago Linguistic Society 39 (pp. 354-368). Chicago, IL.

Wilbur, R. B. (2005). A reanalysis of reduplication in American Sign Language. In B. Hurch (Ed.), Studies in Reduplication (pp. 593-620). Berlin/New York: de Gruyter.

Wilbur, R. B. (2009a). Effects of varying rate of signing on ASL manual signs and nonmanual markers. Language and Speech 52(2/3): 245-285.

Wilbur, R. B. (2009b). Productive reduplication in ASL, a fundamentally monosyllabic language. Language Sciences, 31, 325-342.

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (in press). Using movement and intentions to understand human activity. Cognition.

Zacks, J. M., Swallow, K. M., Vettel, J. M., & McAvoy, M. P. (2006). Visual motion and the neural correlates of event perception. Brain Res, 1076(1), 150-162.

Acknowledgements

This work was supported by National Institutes of Health grant DC005241 and NSF BCS-0345314 grant to Ronnie Wilbur.

Verb Meaning as Verb Argument Realization: evidence from brain-damaged and non brain-damaged populations

Christina Manouilidou
University of Patras, Patras, Greece
Institute of Neurophysiology, University Medical Center, Ljubljana, Slovenia
chmanouilidou@upatras.gr

Abstract

The aim of this paper is to show the importance of Argument Realization as a crucial factor of understanding and processing verb meaning. We discuss experimental data from brain-damaged and non brain-damaged populations regarding verb processing demonstrating the significance of correct thematic role assignment.

1 Introduction

Previous research on verbs has mainly focused on the following principal areas of verb representation and processing: 1. Argument Structure, i.e. the specification of structural relations between the predicate and its arguments. 2. Thematic Roles, i.e. the interpretation of these arguments in terms of the roles they play in the event or state denoted by the predicate. 3. Representation of a verb's meaning in terms of its *internal structure*, traditionally studied in terms of conceptual primitives within a semantic template representing a verb or verb class. While previous theoretical and experimental work has shown the importance of these three domains in our understanding of verbs, what is less clear is the process by which the abstract thematic roles match with the specific structurally represented arguments, in other words the way thematic roles are realized as arguments at the sentence level. In this paper we will show the importance of argument realization (AR) in understanding and processing verb meaning, and by extension, sentence meaning.

2 Theoretical Background

There are various theories that attempted to shed light on the relationship between the abstract thematic roles and the concrete arguments of a verb by targeting the interpretation of sentence constituents according to their syntactic position (i.e. thematic hierarchy, e.g. Fillmore, 1968; Grimshaw, 1990;), their general semantic content (i.e. proto-roles, e.g. Dowty, 1991), and their properties of animacy and definiteness (i.e. animacy hierarchy, e.g. Croft, 2003). Moreover, various specific hypotheses about the linking between thematic roles and a verb's arguments have been proposed (e.g., Perlmutter & Postal's 1984 Universal Alignment Hypothesis—UAH—and Baker's 1988 Universality of Theta Assignment Hypothesis – UTAH). While these theories provide the theoretical framework highlighting the importance of AR, the main purpose of the research presented here is to explore how argument realization affects language processing. We discuss experimental data from brain-damaged and non brain-damaged populations by looking at sentences which require *non-canonical* AR, such as in (1).

(1) a. *The thunder frightened the boy* (*Theme* before *Experiencer*)

b. *The boy feared the thunder* (no *Agent*)

We focus on two experiments which shed light, from different directions, on the same phenomenon, i.e. the role of AR in accessing verb meaning, and consequently sentence meaning. We predict that sentences with non-canonical AR will increase processing load, compared to other features of verb representation that could increase complexity, such as complex internal structure, thus highlighting the distinctive way verb AR contributes to language interpretation (Exp. 1). Similarly, we anticipate brain-damaged populations (i.e. Alzheimer's patients) to demonstrate difficulties dealing with sentences that require non-canonical AR (Exp. 2). Based on the results of these two independent experiments we will support the idea that AR has a unique contribution to verb meaning interpretation and, consequently, to sentence processing.

3 Exp. 1: Evidence from on-line sentence processing in non brain-damaged populations

Sentences with non-canonical AR impose processing difficulties resulting in longer Reading Times (RTs) (Frazier & Clifton, 1996; Bornkessel et al. 2002, 2003; Bornkessel & Schlesewsky, 2006). In the experiment reported here (Manouilidou & de Almeida, under review), we attempted to contrast the roles of internal structure, argument structure and AR and investigate whether we can establish primacy relationships between them. We explored the performance of native speakers of English in four groups of verbs that differ with respect to their internal structure (change-of-state [+CS] vs. non-change-of-state verbs [-CS]) and their thematic roles which might result in non-canonical AR (Agent [+AG] vs. NonAgent [-AG]).

Participants. Thirty-eight undergraduate students participated in the study for course credit. They were all native speakers of English and had normal or corrected-to-normal vision.

Materials. 128 experimental sentences were included, divided into four conditions, according to the variables of change of state (+/-CS) and agentivity (+/-AG). These sentences formed 32 sets such as the one presented in (2). All sentences were normed for plausibility and naturalness and had the same basic syntactic structure, with a NP+Adv+VP (V+NP). We employed manner and degree adverbs in an attempt to affect the volition of the NP occupying the canonical subject position of the sentence. This manipulation was particularly important in conditions such as 2b ([+CS, -AG]), which could denote an intentional act on the part of the Causer of the fright state. Adverbs were also used to further enforce an agentive reading in conditions such as 2a ([+CS, +AG]), and 2c ([-CS, +AG], as well as to keep constant structure and length for all sentence types.

(2)
a. The hunter maliciously *killed* the bear
(+CS,+AG, lexical causatives)
b. The hunter unintentionally *frightened* the bear
(+CS, -AG, object-experiencer)
c. The hunter persistently *followed* the bear
(-CS, +AG, agentive transitive)
d. The hunter barely *sensed* the bear
(-CS, -AG, subject-experiencer)

Procedure. We employed a self-paced reading moving window paradigm. Participants were first presented with a row of dashes on the screen. Each dash represented a letter in the to-appear sentence (such as "--- ------ ---------- ------ --- ---" for sentence (2a)). They were told that each time they pressed the space bar on the computer keyboard, a word would appear in place of the dashes and, as each new word appeared, the previously presented word would turn back to a set of dashes.

Results. RTs for each of the three words of the VP (Verb, Determiner, and Noun) for the four sentence types (lexical causatives, object-experiencer, agentive transitive, and subject-experiencer) constituted the raw data for analyses. Figure 1 depicts the reading times at the **verb** position for the four conditions. A 2 (verb type: +CS vs. -CS) x 2 (agency: +AG vs. -AG) x 3 (VP segment: Verb, Det, Noun) repeated-measures ANOVA showed **no effect of verb type**, $F1\ (1, 37) = .25$, $p = .62$, a marginal **effect of agency**, $F1\ (1, 37) = 3.52$, $p = .069$ and a **significant effect of segment**, $F1\ (2, 74) = 20.73$, $p < .0001$. There was also a significant **interaction between verb type and segment**, $F1\ (2, 74) = 3.35$, $p = .041$. In order to understand how different verb types behaved with regards to different agency manipulations, we performed planned comparisons between the four conditions.. For the **analysis of [-CS] verbs** (*love*, *follow*), there was no effect of agency, $F1\ (1, 37) = .36$, $p = .55$, while the analysis of [+CS] verbs (*kill*, *frighten*), showed a significant effect of agency, $F1\ (1, 37) = 6.17$, $p = .02$. In the analysis of the two agentive sentence types [+AG], a repeated-measures ANOVA showed again no main effect of verb type, $F1\ (1, 37) = .113$, $p = .74$, suggesting that [+CS] and [-CS] structures behave similarly when they are both agentive. Finally, in the analysis of the **two non-canonical structures, [-AG],** the object-experiencer (+CS, -AG) (*frighten*) and subject-experiencer (-CS, -AG) (*love*), a 2 (verb type: [+CS, -AG] vs. [-CS, -AG]) x 3 (VP segment) repeated-measures ANOVA showed no main effect of verb type, $F1\ (1, 37) = .56$, $p = .46$, but a significant effect of segment, $F1\ (2, 74) = 19.55$, $p < .0001$, and a significant interaction between verb and segment, $F1\ (2, 74) = 4.05$, $p = .014$. In pairwise analyses, we found a significant difference between the two constructions at the verb position, t1 (37) = 2.86, p = .007, with [+CS] taking longer than [-CS]. There was no difference between the two [+AG] sentences at the verb locus.[1] In sum, Exp. 1 showed an effect of [CS] only for the [-AG] structures, but not in the [+AG] structures. This suggests that CS *per se* does not account for increased RTs. With respect to non-canonical AR, results showed that there is an effect of agency in [+CS] structures which is only marginal in the [-CS] structures, with [-AG] structures being more difficult to process. Thus, the results suggest that atypical AR, in terms of absence of typical Agent, increases sentence complexity and yields longer RTs in verbal position. In contrast, internal structure does not seem to have an effect in sentence processing. The current findings seem to be in accordance with previous studies indicating a *thematic reanalysis* in sentence processing, which appears to have a processing cost, when the processor's expectations of a thematic role in a particular structural position are not met (see Bornkessel and colleagues). Looking at this outcome in the bigger picture, the present study has shown how *structural and thematic* properties of a verb play the primary roles in sentence comprehension, thus reflecting their prominent role in verb representation.

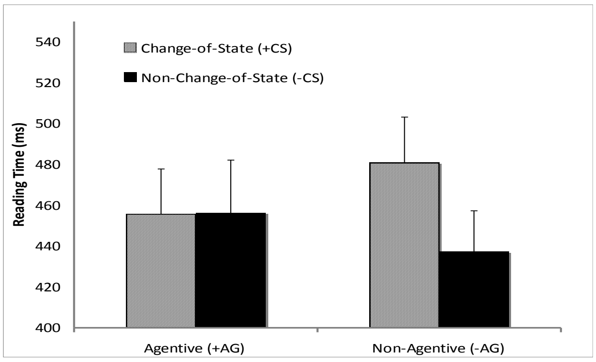


Figure 1: Reading times (in milliseconds) at the verb position in the four sentence conditions.

[1] Paired t-tests showed no significant difference between the mean length of the adverbs used in each condition ([+CS vs -CS]: p=.20; [+AG vs -AG]: p = .12) neither between their mean frequency ([+CS vs -CS]: p=.15; +AG vs -AG]: p = .20). Besides, a post-hoc analysis on adverb RTs showed no significant difference among them. Thus, a possible influence of the adverb length and adverb frequency on the verb RTs should be ruled out.

4 Exp. 2: Evidence from off-line sentence processing in brain-damaged populations[2]

Language impairment studies examining the correspondence between thematic roles and syntactic properties highlight the importance of AR (e.g.; Pinango, 2006). The notion of canonicity in verb-argument relations has been reported to influence sentence processing in aphasic patients (Caplan & Hildebrandt, 1988), and in dementia (Kemper, 1997; Kemper et al., 1993; Lyons et al., 1994; Small, Andersen, et al., 1997; Small, Kemper, et al., 1997; Small et al., 2000). However, in these studies, canonicity in verb-argument relations is usually described in terms of non-canonical thematic role assignment as a result of syntactic manipulations, such as in passive sentences. Hence, non-canonical thematic role assignment and its processing cost cannot be dissociated from other factors that affect sentence processing, such as syntactic movement. For this reason, we cannot be certain that the observed difficulties arise exclusively from non-canonical AR. It could be the case that patients' difficulties are associated with memory and cognitive resources something that Small et al. (2000) also point out. In the sentence completion task described below (Manouilidou et al., 2009) we examined the question of non-canonical thematic role assignment in the performance of populations suffering from dementia in terms of verb-specific requirements and not as a side effect of syntactic manipulations.

Participants. 10 individuals with the diagnosis of probable Alzheimer's Disease (pAD) (mean age: 75.8; s.d. 5.99), 11 elderly controls (mean age 87.25; s.d. 2.5) and 49 young controls (age range: 18-25). The pAD patients' Mini Mental State Examination (MMSE; Folstein, Folstein, and McHugh, 1975) scores ranged from 19 to 27 indicating mild to moderate cognitive impairment. They were all native speakers of English with a minimum education level of sixth grade.

Materials. Patients were required to complete 72 active and passive written sentences by choosing the correct verb. Materials were divided into 6

[2] This section is based on Manouilidou et al. (2009). Thematic Hierarchy violations in Alzheimer's disease: the case of psychological verbs. *Journal of Neurolinguistics* Vol. 22, pp.167-186.

conditions, with 12 sentences in each of them: (1) subject-experiencer verbs (e.g., *fear*); (2) object-experiencer verbs (e.g., *frighten*); (3) subject-agent verbs (e.g., *kick*); and (4), (5) and (6) were the passive equivalent of (1) (2) and (3), respectively (e.g., *was feared, was frightened* and *was killed*).

Design and Procedure. Participants were presented with the sentences with the verb missing marked by a blank line (e.g., *The boy_____the thunder*). They had to choose the correct verb from a list of four verbs, which included the target (e.g., *fear*) its "thematic" distractor[3] (e.g., *frighten*), a syntactically anomalous distractor (e.g., *sleep*) and a semantically unrelated distractor (e.g., *cook*). Materials were divided into four blocks. For sentences corresponding to the *fear-frighten* minimal pairs, four versions were created (e.g., *The boy feared the thunderThe thunder frightened the boy, The boy was frightened by the thunder,* and *The thunder was feared by the boy*), with one version in each block. Active and passive versions of the agentive verbs complemented the blocks (e.g., *the hunter killed the deer, the deer was killed*). Patients and elderly controls saw all four blocks, with two blocks in each of the two sessions, one week apart. Sentences were presented on a computer screen and participants had to choose the correct verb by pressing a key on the keyboard. Each verb on the screen lead to a specific key by an arrow to facilitate the choice by the patients. Testing was completed in two sessions one week apart.

Results. Percentages of correct responses were calculated for each condition (Figure 2). A 3 (group: patients vs. elderly controls vs. young controls) x 3 (voice: active vs. passive) x 2 (subject thematic role: subject-experiencer, object-experiencer, subject-agent) repeated-measures ANOVA[4] showed that patients' data differed significantly from those of the elderly (p<.0001) and the young controls (p<.0001); also, a main effect of verb type was obtained (p=.013), but not of voice (active vs. passive) (p=.13). Repeated measures ANOVAs on the patient data showed a main effect of verb type (p<.001) and a tendency for a main effect of voice (p=.067). Error analysis showed that patients chose the reverse distractor more times when confronted with a psych verb than when confronted with an agent verb. They seldom chose the unrelated distractors. Thus, for example, when confronted with a sentence frame such as *The thunder___the boy* patients selected the correct response *frightened* only 58% of the time—confusing it with *fear* the other times. Most interestingly, there was also a difference between (1) (subject-*experiencer*) and (2) (object-*experiencer*) in the active voice (p= .02) but no difference between their passive equivalents (p=.46).

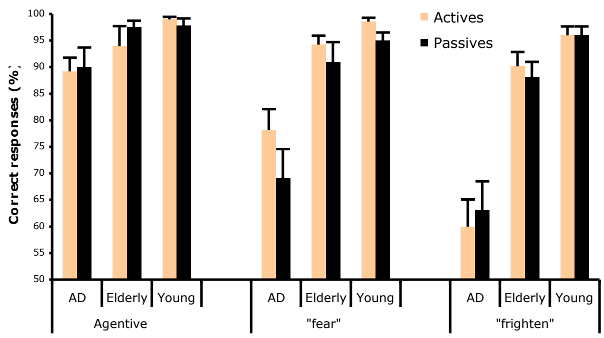


Figure 2: Mean percentage correct responses for the three groups in all conditions. Error bars indicate standard errors.

In sum, the results of Exp. 2 are consistent with our predictions, showing that patients had difficulties completing sentences that required non-canonical argument realization. More importantly, the present study allowed us to identify as source of this difficulty patients' inability to assign the correct thematic roles to the NPs (patients did have access to the correct core meaning of the verb since they almost never chose the unrelated distractors). We take this result to consist additional evidence for the importance of AR in accessing verb and consequently, sentence meaning.

5 Conclusion

The experiments described above come to add to the body of previous research showing the processing costs of non-canonical AR and by extension the general role of AR in sentence processing and verb meaning. Both studies highlight the

[3] In case of agentive verbs, the reverse distractor was a verb in the same semantic field but with different thematic roles. For instance, the distractor for *kill* was *die.*

[4] In all cases, arcsine transformation was employed. However, since we obtained the same effects as with raw data, we choose to report the analyses on the untransformed data.

importance of the [+agentive] feature in verb representation, which projects a canonical argument realization, as a decisive factor in thematic role assignment. The absence of the Agent argument creates difficulties in thematic role assignment, either resulting in thematic reanalysis, and thus, increased processing times (Exp. 1), or in the creation of implausible sentences in neurologically damaged populations (Exp. 2).

REFERENCES

Baker, Mark. 1988. *Incorporation: a theory of grammatical function changing*. University of Chicago Press.

Bornkessel, Ina, Matthias Schlesewsky & Angela D. Friederici. 2002. Beyond syntax: Language-related positivities reflect the revision of hierarchies. *Neuroreport* 13. 361–364.

Bornkessel, Ina, Matthias Schlesewsky & Angela D. Friederici. 2003. Eliciting thematic reanalysis effects: The role of syntax-independent information during parsing. *Language and Cognitive Processes* 18. 268–298.

Bornkessel, Ina & Matthias Schlesewsky. 2006. The extended argument dependency model: a neurocognitive approach to sentence comprehension. *Psychological Review* 113(4): 787-821.

Caplan, David & Nancy Hildebrandt. 1988. *Disorders of syntactic comprehension.* Cambridge, MA: MIT Press.

Croft, William. 2003. *Typology and universals, second edition.* (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67. 547–619.

Fillmore, Charles J. 1968. Lexical entries for verbs. *Foundations of Language* 4(4). 373–393.

Frazier, Lynn & Charles Clifton. 1996. *Construal.* Cambridge, MA: MIT Press.

Grimshaw, Jane. 1990. Argument Structure. [Special issue]. *Linguistic Inquiry* 18.

Kemper, Susan. 1997. Metalinguistic judgments in normal aging and Alzheimer's disease. *Journal of Gerontology: Psychological Sciences*, 52B(3), 147-155.

Kemper, Susan, Emily LaBarge, Richard F. Ferraro, Hintat Cheung, Him Cheung & Martha Storandt. 1993. On the preservation of syntax in Alzheimer's disease: evidence from written sentences. *Archives of Neurology*, 50, 81-86.

Lyons, Kelly, Susan Kemper, Emily LaBarge, Richard F. Ferraro, David Balota & Martha Storandt. 1994. Oral language and Alzheimer's disease: a reduction in syntactic complexity. *Aging and Cognition*, 1, 271-281.

Manouilidou, Christina. & Roberto G. de Almeida, (under review). On the Psychological Reality of Verb Typologies: contrasting the roles of internal structure and argument realization.

Manouilidou, Christina, Roberto G. de Almeida, George Schwartz & N.P.Vasavan Nair. 2009. Thematic Hierarchy violations in Alzheimer's disease: the case of psychological verbs. *Journal of Neurolinguistics* Vol. 22, pp.167-186.

Pinango, Maria Mercedes. 2006. Thematic roles as event structure relations. In I. Bornkessel, M. Schlesewsky, & A. Friederici (Eds.), *Semantic role universals and argument linking: Theoretical, typological, and psycholinguistic perspectives.* Berlin: Mouton.

Perlmutter, M. David & Paul M. Postal. 1984. The 1-Advancement Exclusiveness Law? in D.M. Perlmutter and C. Rosen, eds., (1984) *Studies in Relational Grammar 2*, University of Chicago Press, Chicago, IL, 81-125.

Small, A. Jeff, Elaine S. Andersen, & Daniel Kempler. 1997. Effects of working memory capacity on understanding rate-altered speech. *Aging, Neuropsychology, and Cognition*, 4(2), 126-139.

Small, A. Jeff, Susan Kemper & Kelly Lyons. 1997. Sentence comprehension in Alzheimer's disease: effects of grammatical complexity, speech rate and repetition. *Psychology and Aging*, 12(1), 3-11.

Small, A. Jeff, Susan Kemper & Kelly Lyons. 2000. Sentence repetition and processing resources in Alzheimer's disease. *Brain and Language*, 75, 232-258.

Attribute verbs within and across languages

Rainer Osswald and **Thomas Gamerschlag**
Institute for Language and Information
Department of General Linguistics
University of Düsseldorf, Germany
{osswald|gamer}@phil.uni-duesseldorf.de

Abstract

We describe a project that aims at a cross-linguistic resource of *attribute verbs*, which are stative verbs that encode attributes such as weight (*to weigh*) and price (*to cost*). To this end, we compiled a comprehensive lexical database of German attribute verbs, which have been classified with respect to the attributes they encode and a number of other relevant features. Based on these analyses, attribute verbs from other languages are added to the database in order to reveal typological differences in the encoding of attributes in the verbal domain.

1 Introduction

Language has different ways to encode attribute-value descriptions. If, as in English and many other languages, an attribute like weight or price can be expressed by a noun then the ascription of an attribute and its value can usually be expressed by a copula construction (1-a) or a *have*-possession construction (1-b).

(1) a. The weight of the laptop is two kilos.
  b. The laptop has a weight of two kilos.
  c. The laptop weighs two kilos.

English also has the verb *weigh* for expressing such a description (1-c). An adjectival encoding such as *be two kilos heavy* is not possible in English, in contrast, for instance, to German:

(2) Der Laptop ist zwei Kilo schwer.

The objective of the work reported in this paper is to explore the domain of verbs that encode attributes in the described way, with an eye on cross-linguistic variation. A related goal is to investigate the space of attributes encoded by verbs. As a first step, we compiled a comprehensive list of such verbs in German by systematically exploring existing dictionaries. The verbs are manually classified with respect to the attribute they encode and certain other morpho-syntactic and semantic criteria, and the annotated entries are stored in a lexical database. After a careful revision and adjustment of the attribute space revealed during this process, data from other languages are added to the database, partly with the aim of full coverage, partly for contrastive purposes.

2 Attribute verbs

By an *attribute verb* (or *dimensional verb*) we mean a verb that, in one of its senses, characterizes an entity by specifying the value of an attribute of that entity. Attribute verbs are by definition stative. Standard examples are measurement verbs such as *weigh* (1-c), *cost* (3-a), and *last* (3-b):

(3) a. The ticket costs two dollars. [PRICE]
  b. The lesson lasts two hours. [DURATION]

Each of the sentences in (3) describes the entity denoted by the subject with respect to the attribute expressed by the verb by specifying the value of that attribute through a "value phrase" (*two dollars*, *two hours*). The notion of attribute is taken here in a broad sense that subsumes also location, meaning, function, etc.

2.1 Empirical basis and coverage

We used the German standard dictionary *Duden Deutsches Universalwörterbuch* (DDUW) as a primary source for compiling a comprehensive list of German attribute verbs. Roughly 800 of the more than 13,000 verbs in the dictionary have been identified as attribute verbs. All of them are classified along the scheme described in Section 2.2 and are stored in a database (cf. Section 5). A representative set of French, Spanish, and English attribute verbs is currently classified along the same lines and added to the database. In addition, a questionnaire has been used to gather data from native speakers and experts of further languages, which are mainly considered for contrastive purposes and with less ambition of lexicographic coverage.

2.2 Classification

The encoded attributes are by definition the most important features of the classification. Their formal properties will be discussed in Section 2.3. While the analyses and statistics presented in this section are based on the German sample introduced above, we often give corresponding English examples for ease of exposition (but see Figure 1). In addition to the encoded attribute, the verbs in the sample are characterized with respect to the following features:

Scalarity. Attributes like those in (3) are *scalar* in the sense that their value range is linearly ordered.[1] The attributes expressed by the verbs in (4), by comparison, are non-scalar.

(4) a. The yoghurt tastes of bananas. [TASTE]
b. The path leads to a lake. [GOAL]

Although scalarity is a property of the attribute itself, and not of the verb, we used it as a verb feature in the first round of classification for reasons of consistency control.

Value incorporation. Some attribute verbs inherently restrict or specify the value of the attribute they encode. For example, the English verb *bulge* describes the shape of an object (5).

(5) The bag bulges (with papers). [SHAPE]

Attribute verbs with inherent values can be seen as *descriptive verbs* in the sense of Snell-Hornby (1983), with the inherent value as the "modificant" of the "nucleus" given by the attribute.

Absolute use. Various attribute verbs that usually require an external value phrase can also be used absolutely, that is, without such a phrase. In this case, the implied value can depend on the typical properties of the entity denoted by the subject; viz. *The socks smell* vs. *The roses smell.*

[1] Scalar structure has been recognized as an important factor for the analysis of gradation; see e.g. Kennedy and McNally (2005).

One vs. two attributes. Some attribute verbs are "two-dimensional" in that they encode two attributes. In our German sample, about 16% of the attribute verbs have been identified as two-dimensional. More than 70% of them encode LOCATION as an attribute with external value phrase and an additional attribute with inherent value. The two most frequent of these additional attributes are POSTURE and MANNER OF CONTACT; see (6) for English examples:

(6) a. They squatted near the fire.
[LOCATION × POSTURE]
b. His wet shirt clung to his body.
[LOCATION × CONTACT]

There is moreover a small number of verbs such as English *hover* which combine LOCATION with an inherently specified "supporting medium" attribute:[2]

(7) The hummingbird hovered over the flowers.
[LOCATION × MEDIUM]

Instead of LOCATION, two-dimensional attribute verbs can also encode "path-related" attributes such as SOURCE, GOAL, and COURSE. However, as illustrated by the examples in (8), the specific path aspect is usually provided by the preposition and not by the verb itself (Jackendoff, 1990; Kaufmann, 1995; Eschenbach et al., 2000).

(8) a. The tree arched over the road.
[SHAPE × COURSE]
b. A male voice droned from the TV.
[SOUND × SOURCE]

It is nevertheless an inherent semantic property of these verbs to have an associated path. Moreover, word formation in German can give rise to attribute verbs with incorporated prepositions (see below).

Simplex vs. complex. A central concern in describing the lexical repertoire of attribute verbs in a language is to identify the *simplex* verbs, i.e., the monomorphemic lexemes within this class. Languages differ considerably in their morphological potential to form complex words. German has a rich system of verb prefixes comprising particles, prepositions, and adverbs, among others. The morphologically complex verbs in our sample are

[2] This analysis is based on Kaufmann (1995, Chap. 6.1).

subclassified into semantically transparent and intransparent formations. The first class consists mainly of regular incorporations of locative or directional prepositions, in which case the meaning of the resulting complex verb is compositionally derived from the meaning of the base verb and that of the preposition. For example, combining the directional preposition *über* with the attribute verb *sich wölben* results in the transitive complex verb *überwölben* (*arch over*), that is, *etwas überwölben* means *sich über etwas wölben*. In particular, the incorporated preposition reduces or eliminates the flexibility in the choice of the preposition usually available for locative or directional PP arguments. A special case of transparent formation is provided by verbs like *zurechnen* that are derived from proper prepositional verbs, here *rechnen zu* (*count among*).[3] Despite the fact that the semantic contribution of the preposition is subregular at best in these cases, preposition incorporation only affects the syntactic valency and is hence transparent in the above sense. All in all, about half of the entries of the German sample are morphologically simple while approximately 25% have been classified as transparent and complex.

Primary vs. secondary. A considerable number of verbs used for attribute-value descriptions are "secondary" attribute verbs in that they are non-stative in their primary sense. The stative sense of *lead* in (4-b), for example, counts as secondary. The relation between this sense and the basic sense of *lead* can be described as a lexicalized *metonymic shift*. Other examples of this type are provided by change of direction verbs such as *abdrehen* (*turn*) as in *Der Fluss dreht nach Westen ab*.[4] *Argument alternations* are a further source for deriving secondary attribute verbs from non-stative verbs.[5] Examples are the "characteristic property alternation" (e.g., *This knife cuts well*) and the middle alternation.

Approximately one out of three attribute verbs in the German sample have been analyzed as secondary. These verbs are furthermore classified with respect to the mechanism by which they are derived from the corresponding basic verb sense.

[3]Cf. Osswald et al. (2006) for more information on prepositional verbs in German.

[4]The stative sense of *turn* belongs to a class called "meander" verbs in Levin (1993) and "pseudo-motional locative" verbs in Dowty (1979).

[5]See Levin (1993) and Frense and Bennett (1996) for an overview of argument alternations in English and German.

Sense distinctions. We distinguish different senses of an attribute verb to the extent that they encode different attributes or attribute values. Under this regime, the attribute verbs in the German sample show an average polysemy of 1.2. About 13% of the entries are polysemous and the average polysemy within that set is 2.6.

Each sense is linked to the corresponding DDUW section or sections, if existent. For nearly 10% of the German entries in the sample, we added uses as attribute verbs that have no corresponding section in the DDUW. As to be expected, a good part of the missing readings are secondary in the sense introduced above, and thus related to non-stative senses in a more or less systematic way (cf. the discussion in Section 3).

Nominal and adjectival equivalents. We record if the attribute-value description expressed by an attribute verb can be expressed by a nominal (1) or an adjectival construction (2). Investigating the nominal and adjectival equivalents of attribute verbs is relevant for cross-linguistic comparison and also for questions concerning the diachronic development of attribute verbs.

A preliminary investigation of the simplex attribute verbs in the German sample has revealed that adjectival equivalents are rare. Nominal equivalents are fairly frequent, with less than half of them derivationally related to the verb. Within the latter class, deverbal nouns (e.g., *stinken* > *Gestank*) are considerably more frequent than denominal verbs (e.g., *Duft* > *duften*).

Valency. All attribute verbs in the sample are characterized with respect to their syntactic valency. In particular, the valency position of the value phrase has been explicitly marked. The optionality of this argument position corresponds to a possible absolute use of the attribute verb (see above).

2.3 The space of attributes

It is part of the project to explore the space of attributes encoded by attribute verbs on an empirical basis. The set of attributes used in the classification was not set up *a priori*, but is developed during the classification process and subsequent revision cycles.

Since attribute verbs can encode more than one attribute, there is no straightforward assignment of attributes to verbs. To put it the other way around, one cannot expect to devise a simple taxonomy of attributes with attribute verbs uniquely attached to the nodes of the hierarchy. In fact, there seem to be at least three different ways of combining attributes that have to be taken into account: pairing, conjunction, and composition. *Pairing* ($\times$) is meant to refer to the two-dimensional case discussed before. *Conjunction* ($+$) is needed, for instance, to distribute the features TEMPORAL and SPACIAL over POSITION and DIRECTION. Conjunction thus realizes multiple inheritance. The attribute encoded by the German verb *datieren*, which is used to locate events in time, can then be written as POSITION + TEMPORAL. *Composition* ($\circ$), finally, is to be understood in the formal sense of functional composition, with attributes regarded as functions. Composition can be applied to characterize the attributes encoded by verbs like *begin* and *end*, when used to locate the begin or end of a trail, road and the like. For *The trail starts at the chapel*, the encoded attribute would then be expressed as BEGIN $\circ$ LOCATION, or BEGIN $\circ$ (POSITION + SPATIAL).

Figure 1 shows part of the system of classes of attribute verbs currently under development, with corresponding German examples. (Attribute composition is neglected in the figure.)

2.4 Cross-linguistic variation

The availability of verbs for encoding attribute-value descriptions differs considerably between languages. For instance, French has less verbs than English or German for encoding attributes of the SENSATION class. While the English attribute verbs *taste* (German: *schmecken*) and *feel* (German: *anfühlen*) can express TASTE and TOUCH, French uses constructions such as *avoir un goût* and *être ... au toucher* instead.

Another difference between Romance and Germanic languages shows up with two-dimensional verbs that involve LOCATION or PATH. In German and English, there are attribute verbs that combine LOCATION and POSTURE, whereas in French, a copula construction with past participle is required instead (Schwarze, 1993); compare *sitzen* (*sit*) and *liegen* (*lie*) vs. *être assis* and *être allongé*. The same is the case for the pairing LOCATION $\times$ CONTACT, viz. *kleben* (*stick*) vs. *être collé*. This typological difference seems to be related to the distinction between verb-framed and satellite-framed languages proposed by Talmy (1985), according to which verb-framed languages such as French tend to express manner by an adjunct in conflated constructions.

In order to investigate typological differences on a broad empirical basis, we developed a questionnaire to collect data about the verbal encoding of attributes in the languages Spanish, Korean, Russian, and Lakhota (Siouan). In addition, we are currently expanding the database to include French and English entries following the classification scheme described in Section 2.2.

3 The lexicographic perspective

A resource of attribute verbs as described in this paper can contribute to lexicography in various ways. For instance, it can be employed to improve the coverage of monolingual dictionaries with respect to stative uses of non-stative verbs. An overview of secondary attribute verbs and the underlying mechanisms of meaning shift can be helpful in this respect. This includes cases of valency alternation, which are often not systematically covered in dictionaries; see, e.g., Schwarze (2008) on the transitive and intransitive uses of the attribute verb *medir* (English: *measure*) in Spanish dictionaries.

The existence of sense gaps in the DDUW has already been mentioned in the discussion of sense distinctions in Section 2.2. Even the entries of frequent verbs such a *drehen* and *wenden* (*turn*) do not make clear that they can be used to describe the change of direction of a road, river, etc. By comparison, the 5th edition of the 'Longman Dictionary of Contemporary English', is quite explicit about the corresponding sense of *turn*, which is listed in a section headed by 'Direction' and has the definition 'if a road, river etc turns, it curves and starts to go in a new direction'.

Concerning the question of how to account for secondary attribute verbs in the dictionary, we therefore agree with Apresjan (2002), who requires that "all salient lexical classes should be fully taken into account and uniformly described in a dictionary in all of their linguistically relevant properties." If applied to the classes of attribute verbs and the systematic relations between secondary attribute verbs and their non-stative base verbs, Apresjan's "principle of systematic lexicography" would surely help to improve treatment of attribute verbs in the dictionary.

4 Comparison with existing verb classifications and resources

While some of the subclasses of attribute verbs have been studied before in work on stative verbs (Gerling and Orthen, 1979; Rothmayr, 2009), there has been no systematic investigation of such verbs in lexical semantics or lexicology up to now.

Existing lexical-semantic resources such as WordNet, FrameNet, or VerbNet do not pay attention to the systematic analysis and classification of attribute verbs (and stative verbs in general), with consequential gaps and inconsistencies in this domain. It is worth mentioning that the Brandeis Semantic Ontology of Pustejovsky et al. (2006) contains a number of subtypes of the type *Value Relation* such as *Amount*, *Cost*, *Height*, *Size*, *Temperature*, *Velocity*, and *Weight*, which are all to be located under QUANTITY/MEASURE in Figure 1. However, verbs are rather rare compared to nouns and adjectives in the set of lexical items associated with these ontological types.

5 Representation and implementation

The design of the lexical database and its implementation was driven by the requirement of a lean architecture that is easily modifiable and extensible and, furthermore, supports collaborative and platform-independent access and modification of the lexical data. We chose an XML database and implemented a web interface that allows editing and flexible browsing (including XQuery support).[6] The chosen XML schema (specified via RELAX NG) has been kept as simple as possible for the moment. At a later point, we plan to migrate to a representational format more in line with existing standards for lexical data such as the Lexical Markup Framework (Francopoulo et al., 2006). It is planned to make the web interface accessible to the general public at the project end in summer 2011.

6 Ongoing work and prospects

The main focus of the project is currently on extending the lexical database with attribute verbs from other languages than German. Concerning the German sample, we plan to add more corpus-based examples.

[6]As to technical details, the implementation uses Berkley DB XML, an embedded XML database, and Pylons, a Python-based web framework.

SENSATION/APPEARANCE	
SENSATION	
SMELL	*riechen*, *duften*
TASTE	*schmecken*, *munden*
TOUCH	*sich anfühlen*, *kratzen*
SOUND	*klingen*, *dröhnen*
LOOK	*aussehen*, *glänzen*
APPEARANCE	*anmuten*, *wirken*
POSITION	
+ SPATIAL	*sich befinden*, *wohnen*
× POSTURE	*knien*, *hocken*
× CONTACT	*kleben*, *stecken*
× MEDIUM	*schweben*, *schwimmen*
+ TEMPORAL	*datieren*
STRUCTURE	
SHAPE	*sich wölben*
COMPOSITION	*bestehen aus*
EXTENT	
QUANTITY/MEASURE	
WEIGHT	*wiegen*
DURATION	*dauern*
SPEED	*fahren*, *draufhaben*
CAPACITY	*fassen*
PATH	*führen*, *gehen*
COURSE	*verlaufen*
DIRECTION	
SOURCE	*wegführen*
GOAL	*hinführen*, *zeigen*
EXTREMAL	
BEGIN	*anfangen*, *entspringen*
END	*enden*, *münden*
SOCIOCULTURAL	
POSSESSOR	*gehören*
ROLE	*darstellen*, *verkörpern*
FUNCTION	*dienen*, *fungieren*
SEMIOTIC	
NAME	*heißen*
MEANING	*bedeuten*
REFERENCE	*sich beziehen auf*

Figure 1: Sketch of the system of attribute verb classes under development, with German examples.

Moreover, the formal characterization of the attribute space needs further investigation and might benefit from taking into account existing work on formal ontologies such as DOLCE (Borgo and Masolo, 2009).

It is furthermore planned to extend the coverage of the resource to non-stative attribute verbs, that is, to verbs which encode the change of attribute values.

Acknowledgments

The work presented in this paper has been financially supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) as part of the project 'Dimensional Verbs' of the DFG research group FOR600 'Functional Concepts and Frames'.

References

Juri Apresjan. 2002. Principles of systematic lexicography. In Marie-Hélène Corréard, editor, *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins*, pages 91–104. Euralex, Grenoble.

Stefano Borgo and Claudio Masolo. 2009. Foundational choices in DOLCE. In Steffen Staab and Rudi Studer, editors, *Handbook on Ontologies*. Springer, Berlin.

David Dowty. 1979. *Word Meaning and Montague Grammar*. D. Reidel, Dordrecht.

Dudenredaktion, editor. 2003. *Duden. Deutsches Universalwörterbuch*. Duden Verlag, Mannheim, 5th edition.

Carola Eschenbach, Ladina Tschander, Christopher Habel, and Lars Kulik. 2000. Lexical specifications of paths. In Christian Freksa, Wilfried Brauer, Christopher Habel, and Karl Friedrich Wender, editors, *Spatial Cognition II*, Lecture Notes in Computer Science 1849, pages 127–144, Berlin. Springer.

Gil Francopoulo, Monte George, Nicoletta Calzolari, Monica Monachini, Nuria Bel, Mandy Pet, and Claudia Soria. 2006. Lexical Markup Framework (LMF). In *Proceedings of LREC 2006*, pages 233–236, Genoa, Italy.

J. Frense and P. Bennett. 1996. Verb alternations and semantic classes in English and German. *Language Sciences*, 18(1–2):305–317.

Martin Gerling and Norbert Orthen. 1979. *Deutsche Zustands- und Bewegungsverben*. Studien zur deutschen Grammatik 11. Gunter Narr Verlag, Tübingen.

Ray Jackendoff. 1990. *Semantic Structures*. MIT Press, Cambridge, MA.

Ingrid Kaufmann. 1995. *Konzeptuelle Grundlagen semantischer Dekompositionsstrukturen. Die Kombinatorik lokaler Verben und prädikativer Argumente*. Niemeyer, Tübingen.

Christopher Kennedy and Louise McNally. 2005. Scale structure, degree modification and the semantics of gradable predicates. *Language*, 81:345–381.

Beth Levin. 1993. *English Verb Classes and Alternations*. University of Chicago Press, Chicago.

Rainer Osswald, Hermann Helbig, and Sven Hartrumpf. 2006. The representation of German prepositional verbs in a semantically based computer lexicon. In *Proceedings of LREC 2006*, pages 461–464, Genoa, Italy.

James Pustejovsky, Catherine Havasi, Jessica Littman, Anna Rumshisky, and Marc Verhagen. 2006. Towards a generative lexical resource: The Brandeis Semantic Ontology. In *Proceedings of LREC 2006*, pages 1702–1705.

Antonia Rothmayr. 2009. *The Structure of Stative Verbs*. John Benjamins, Amsterdam.

Christoph Schwarze. 1993. Primäre und sekundäre Lokalverben im Französischen. In Frank Beckmann and Gerhard Heyer, editors, *Theorie und Praxis des Lexikons*, pages 103–122. de Gruyter, Berlin.

Brigitte Schwarze. 2008. Verbos dimensionales en el diccionario. In *Actas del II Congreso Internacional de Lexicografía Hispánica*.

Mary Snell-Hornby. 1983. *Verb-Descriptivity in German and English. A Contrastive Study in Semantic Fields*. Carl Winter, Heidelberg.

Leonard Talmy. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen, editor, *Language Typology and Syntactic Description. Volume III: Grammatical Categories and the Lexicon*, pages 57–149. Cambridge University Press, Cambridge.

Evidence from neuropsychology on verb features: The case of a patient with Semantic Dementia

Domenica Romagno
University of Pisa
Department of Linguistics
d.romagno@ling.unipi.it

Costanza Papagno
University of Milano-Bicocca
Department of Psychology

Alfonso Caramazza
Harvard University
Cognitive Neuropsychology Laboratory
Center for Mind/Brain Sciences–
University of Trento

Abstract

The aim of this study is to investigate whether Referential Semantics (RS) and Lexical Semantics (LS) are independently represented in the brain. We report the case of a Semantic Dementia patient who, despite a severe impairment in accessing verb RS, was perfectly able to access verb lexical semantic features governing specific morphosyntactic operations. This pattern of performance constitutes the first clear evidence that the distinction between RS and LS, that operates in language, does have a neuropsychological correlate.

1 Introduction

Verb meaning includes two different components at least: Referential Semantics (RS) and Lexical Semantics (LS) (in Levin & Rappaport Hovav's terms: Levin & Rappaport Hovav, 1995, among others).

RS determines verb reference: dying, for instance, refers to stopping living or existing; walking refers to moving along by lifting and setting down each foot in turn. LS, instead, corresponds to more abstract features, like agentivity and telicity, that are shared by verbs, independently of differences or similarities between their RS. Dying is telic (it entails a specified endpoint corresponding to the change of state of the subject) and unagentive (the subject does not have the control of the event). And so is collapsing, even though dying and collapsing do not have the same RS: "the man died" vs. "the roof collapsed". Walking, instead, is atelic (it denotes an event unfolding over time with no final state or specific delimitation) and agentive (the subject does have the control of the event). And so is talking: even though walking and talking do not have the same RS: "the boy walked along the river" vs. "the professor talked about Higgs boson".

The question that we address here – and that has not yet been addressed – is whether these two components of verb meaning are independently represented in the brain.

Since LS governs morphosyntax in a specified way independently of RS, we are able to investigate the neural dissociation between the two semantic components by testing morphosyntactic processing.

Features like telicity and agentivity are morphosyntactically relevant, indeed: there appear to be striking lexical semantic regularities in the composition of classes of verbs sharing the same morphosyntactic patterns, "regularities that are manifested across languages in impressive similarities" (Levin & Rappaport Hovav, 1995:2).

We can fairly claim that these morphosyntactic patterns are semantically determined (Levin & Rappaport Hovav, 1995, 2001; VanValin, 1990, Van Valin & LaPolla, 1997; Tenny, 1994; Croft, 1990; Dowty, 1979, 1991; Chomsky, 1981, 1986; Perlmutter, 1978).

Let us consider, for instance, the distribution of the temporal adverbials "in X time" and "for X time" in sentences like "the man died <u>in</u>/*for an hour" vs. "the man walked <u>for</u>/*in an hour". Dying selects "in X time", as opposed to walking.

Does the behavior of the two verbs depend on syntactic differences? Or rather, are these syntactic representations semantically driven?
The syntactic frame is identical in both sentences; in addition, both dying and walking are compatible with either "for" or "in": "the man died for his country", "the man walked in an unusual way". However, the temporal adverbial "in X time", that has a delimiting value, occurs only with telic verbs like "to die" which denotes a delimited event, as opposed to atelic verbs like "to walk" which denotes an event with no specific delimitation. Thus, syntactic features being equal, it is possible to identify the components of verb meaning that give rise to a given pattern as opposed to the other.
The list of verbal phenomena that are morphosyntactically represented but semantically determined is significantly long and well-known in the literature (Levin & Rappaport Hovav, 1995, 2001; Zaenen, 1993; Haspelmath, 1993, 2001; Alexiadou et al. 2004; Aikhenvald-Dixon-Onishi, 2001; Sorace, 2000; Centineo, 1996).
We tested our Semantic Dementia (SD) patient through a series of morphosyntactic tasks that specifically required access to the lexical semantic features determining the morphosyntactic representations involved in each task.
Patients with SD offer a unique opportunity to investigate the dissociation between RS and LS: they typically show a severe impairment in accessing the RS of words, in the face of a good ability to produce well-formed sentences.
Previous studies on SD patients – and, more generally, on Fronto Temporal Dementia (FTD) patients (Breedin & Saffran, 1999; Cotelli et al. 2007; Tyler et al. 1997; Rochon et al. 2004; Schwartz, Marin & Saffran, 1979; Patterson et al. 2001; Tyler et al. 2004; Benedet et al. 2006; Neary et al. 1998; Hodges & Patterson 1996; Hodges et al. 1992; Meteyard & Patterson, 2009; Patterson & MacDonald, 2006, Lambon Ralph & Patterson, 2008; Visser et al. 2010) – almost exclusively focused on either RS or morphosyntactic patterns that are independent of LS, such as, for instance, the so-called 'wh-movement', that perfectly applies to either "John died" ("Who died?") or "John talked" ("Who talked?"), even though dying and talking belong to different lexical semantic classes.
We investigated the patient access to LS, in order to see whether her severe impairment at RS was necessarily accompanied by an impairment at LS, or rather the two components of verb meaning were neurally distinguishable.

Here we provide the first evidence that the distinction between referential semantic and lexical semantic verb features, that operates in language, can give rise to a neuropsychological dissociation. Our patient presented with a severely damaged RS, but intact LS. Significantly, when tested on morphosyntactic operations which are not sensitive to LS, she did not perform well. Her percentage of errors was up to 40%, in the face of 100% correct responses in the morphosyntactic tasks requiring access to LS.

2 Case Presentation

Patient MC was previously reported by Papagno, Capasso & Miceli (2009), who found a reversal of the concreteness effect restricted to nouns. MC is an Italian 75-year-old, right-handed woman with 17 years of education. She worked as a teacher until 1995. MC suffers of Semantic Dementia, a neurodegenerative disease that belongs to the Fronto-Temporal Dementia (FTD) spectrum (Warrington, 1975; Neary, Snowden et al. 1998). She showed the typical pattern of an SD patient with regard to both neuropsychological behavior and neurological profile (Hodges and Patterson, 2007). At the time of our research, MC's semantic deficit affected both the grammatical categories of Noun and Verb almost to the same extent and without any distinction between either concrete and abstract terms or animate and inanimate entities.
MC's neuroimaging revealed a bilateral degeneration of the Anterior Temporal Lobes, with a greater atrophy on the left side at the earlier stages of the disease; the atrophy also progressively involved the insula and the frontal lobes bilaterally (Figure 1-2).

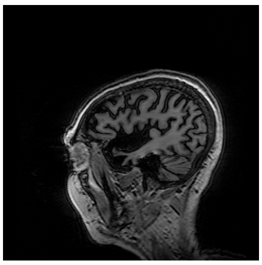


Figure 1

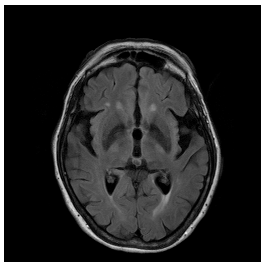


Figure 2

Figure 1 and 2. Marked atrophy involving bilaterally the temporal lobe, the insula, and the frontal operculum, more pronounced in the left hemisphere.

3 Methods and Results

3.1 Morphosyntactic Tasks

The morphosyntactic tasks that we used to test the patient's access to LS are summarized in Table 1. Each task specifically requires access to the property or the cluster of properties determining the morphosyntactic representations which are involved.
The patient was provided with written sentences (also read aloud by the examiner) and requested to respond in either written or oral modality.

Task 1
The distribution of the temporal adverbials "in X time" and "for X time" allows to distinguish between telic verbs like "to die" and atelic verbs like "to talk": "in X time", that has a delimiting value, occurs only with verbs denoting a delimited event, like telic verbs, as opposed to atelic verbs which denote an event with no specific delimitation or final state: *Luigi è morto in/*per un giorno* "Luigi died in/*for a day" vs *Mario ha parlato per/*in un'ora* "Mario talked for/*in a hour".

Task 2
Imperative mood allows to distinguish between agentive and non-agentive verbs: since imperative requires the subject to have the control of the event, it occurs only with agentive verbs, such as "to walk": *Cammina!* "Walk!" vs **Esisti!* *"Exist!".

Task 3
Present Progressive allows to distinguish between dynamic and non-dynamic verbs. Since progressive tense is a statement of dynamic process, it does not occur with states like "to possess", "to consist of": *Maria sta mangiando un gelato* "Maria is eating an ice-cream" vs **Anna sta possedendo una grande intelligenza* *"Anna is possessing a great intelligence".

Task 4
"To be" as auxiliary verb in compound tenses and PP agreement with the subject (i.e. the PP takes an ending that agrees in gender and number with the subject) allow to distinguish between the intransitive verbs which do entail a state predicate in their logical structure – that is, intransitive verbs denoting either a change of state/location, such as "to arrive", or an inherent state/location, such as "to exist" – and the intransitive verbs which do not, such as "to work". These select "to have" and lack PP agreement with the subject (i.e. the PP takes the unmarked singular ending *–o*):

*I soldat<u>**i**</u> **sono*** ("are") *arrivat<u>**i**</u>*
"The soldiers arrived "
vs.
*Gli opera<u>**i**</u> **hanno*** ("have") *lavorat<u>**o**</u>*
"The laborers worked"

Task 5
Agent nouns with *–tore* allow to distinguish between verbs which require an UNDERGOER subject, like "to belong", "to die" (whose subject is in a given state/location or undergoes a change of state/location) and verbs which do not, like "to travel". The verbs requiring an UNDERGOER subject do not produce agent nouns via the suffix *-tore*: *viaggiatore* "traveler" vs. **appartenitore* *"belonger".

3.2 Semantic Tasks

The patient's access to the RS of the verbs and the nouns included in the five tasks described above was previously tested via a vocabulary task (oral definition), like the one included in the Wechsler Adult Intelligence Scale (WAIS). Her performance was compared with five neurologically-unimpaired control subjects, matched for age, education and sex.

3.3 Referential Semantics vs Lexical Semantics

The patient's performance of the tasks testing her access to word RS was extremely poor, as reported in Table 2.

Referential Semantics		
Task 1–5		
Word Comprehension		
	Verbs	Nouns
MC's Success Rate	30.32%	28.5%

Table 2. MC's access to the RS of the verbs and the nouns used in the morphosyntactic tasks 1-5.

On the contrary, the patients's success rate in performing the tasks that specifically required access to the LS of the words involved in the morphosyntactic representations, was remarkably high, as shown in Table 3. Hence, we can conclude that the patient's pattern of performance revealed preserved LS, in the face of severely impaired RS (Table 4).

	Stimuli	Lexical Semantic Properties	Linking Rules
Task 1: Sentence Completion Temporal Adverbials	*Luigi è morto. . . (in/per) un giorno* "Luigi died. . . (in for) a day" *Mario ha parlato. . . (in/per) un'ora* "Mario talked. . . (in for) an hour"	Telicity	+Telicity = *in* X time –Telicity = *per* X time[1]
Task 2: Acceptability Judgement Imperative	*Cammina!* "Walk!" *Esisti!* "Exist!"	Agentivity	+Agentivity = +Imperative –Agentivity = –Imperative
Task 3: Acceptability Judgement Present Progressive	*Maria sta mangiando un gelato* "Maria is eating an ice-cream" *Anna sta possedendo una grande intelligenza* "Anna is possessing a great intelligence"	Dynamicity	+Dynamicity = +Pres.Progr. –Dynamicity = –Pres.Progr.
Task 4 Sentence Completion Auxiliary Selection and Past Participle (PP) Agreement with Subject	*I soldati...(sono/hanno) arrivat…(i/o)* "The soldiers…(are/have) arrived" *Gli operai. . . (sono/hanno) lavorat. . .(o/i)* "The laborers. . .(are/have) worked"	Event Structure (ES)	+State Predicate in the ES = "to be", +PP agreement –State Predicate in the ES = "to have", –PP agreement
Task 5 Acceptability Judgement Derivational Suffix *-tore*	*Andrea è un viaggiatore curioso* "Andrea is a curious traveler" *Mario è un serio appartenitore* "Mario is a serious belonger"	Subject's Semantic Role	ACTOR Subject = +Suffix *-tore* UNDERGOER Subject = –Suffix *-tore*[2]

Table 1. Morphosyntactic tasks used to test MC's access to verb LS features.

	Morphosyntactic Behavior				
	Task 1 Temporal Adverbials	**Task 2** Imperative	**Task 3** Present Progressive	**Task 4** Aux.Sel. and PP agreement	**Task 5** Derivational Suffix *–tore*
MC's Success Rate	100%	100%	100%	100%	100%

Table 3. MC's success rate in the morphosyntactic tasks requiring access to verb LS features.

	Verbs	LS	RS
Task 1	*morire* vs. *parlare* "to die" vs. "to talk"	100%	20%
Task 2	*camminare* vs. *esistere* "to walk" vs. "to exist"	100%	33,3%
Task 3	*mangiare* vs. *possedere* "to eat" vs. "to possess"	100%	33,3%
Task 4	*arrivare* vs. *lavorare* "to arrive" vs. "to work"	100%	25%
Task 5	*viaggiare* vs. *appartenere* "to travel" vs. "to belong"	100%	40%

Table 4. MC's success rate in the tasks testing her access to verb LS vs RS features.

---

1 We refer to Bertinetto (1986) on different uses of the temporal adverbial "*per* X time" in combination with accomplishment and achievement verbs.

2 The Italian suffix *–tore* is also used to derive *nomina instrumenti* (e.g. *contenitore* "container") from verbs: in this case, verb agentivity, obviously, is not required. When *–tore*, instead, encodes an agent noun (*nomen agentis*), the ACTOR semantic role of the subject is required.

4 Discussion

Despite a severe impairment in accessing verb referential semantic features, MC was perfectly able to access the lexical semantic features determining the morphosyntactic behavior of verbs. One could again suppose that MC's pattern of performance relies on a general dissociation between semantics and morphology. However, the patient's morphological processing was not uniformly preserved. MC performed well only the tasks that specifically required access to LS, whereas her success rate in processing morphological operations that are not sensitive to LS, such as inflectional forms which are lexically controlled, was significantly lower. She was provided with written sentences (also read aloud by the examiner) including incorrect verb forms – e.g. *"ieri Mario ha **corruto** nel parco"* (***corruto*** instead of ***corso***, Past Participle of *correre* "to run") "yesterday Mario has run in the park",*"lo scorso anno **mettei** su tre chili in un mese* (***mettei*** instead of ***misi***, Past Tense of *mettere* "to put") "last year, I put on three kilos in a month", *"Filippo venirà domani* (***venirà*** instead of ***verrà***, Future Tense of *venire* "to come") "Filippo will come tomorrow" – alternating with sentences including correct verb forms, and asked to recognize the incorrect forms and to produce the correct ones. *Corso* vs *corruto*, *misi* vs *mettei* refer to inflectional forms that are independent of LS. In this task, MC's percentage of errors was up to 40%, in the face of 100% correct responses in the morphosyntactic tasks requiring access to LS.

One could again suppose that the patient's performance in the morphosyntactic tasks 1–5 (Table 1) relies on her possibly preserved episodic memory. However, when provided with sentences including unusual/impossible combinations (e.g. "the tree walked...(in/for) an hour", "the table died... (in/for) an hour"), she replicated her success rate across all the five tasks.

5 Conclusions

MC's pattern of performance revealed a clear dissociation between two components of verb meaning: RS and LS.

There is a need for further investigation into the RS/LS domain. This study provides a first answer to a question that has long baffled linguists and cognitive-neuroscientists: is the morphosyntactically relevant component of word meaning (LS) neurally distinguishable from the referential meaning of words (RS)? Here we provided the first evidence that the distinction between RS and LS does have a neuropsychological correlate.

References

A.Y.Aikhenvald, R.M.W. Dixon and M. Onishi (eds.). 2001. *Non-canonical marking of subjects and objects*, Amsterdam/Philadelphia, Benjamins.

A. Alexiadou, E. Anagnostopoulou and M. Everaert. 2004. *The Unaccusativity Puzzle. Explorations of the Syntax-Lexicon Interface*, University Press, Oxford.

M. Benedet, K. Patterson, I. Gomez-Pastor, M. L. Garcia de la Rocha. 2006. 'Non-semantic'Aspects of Language in Semantic Dementia: As Normal as They're Said to Be? *Neurocase*, 12:15-26.

S.D. Breedin and E.M. Saffran. 1999. Sentence Processing in the Face od Semantic Loss: A Case Study, *Journal of Experimental Psychology*, 128(4):547-562.

G.Centineo. 1996. A lexical theory of auxiliary selection in Italian, *Probus*:223-271.

N.Chomsky. 1981. *Lectures on Government and Binding*, Foris, Dordrecht.

N.Chomsky. 1986. *Knowledge of Language: Its Nature, Origin and Use*, Praeger, New York.

W. Croft. 1990. *Possible verbs and the structure of events*. In: S.L. Tsohatzidis (ed.), *Meanings and prototypes. Studies in linguistic categorization*, Londra/New York, Routledge.

M. Cotelli, B.Borroni, R. Manenti, V. Ginex, M. Calabria, A. Moro, A. Alberici, M. Zanetti, O. Zanetti, S.F. Cappa and A. Padovani. 2007. Universal grammar in the frontotemporal dementia spectrum. Evidence of a selective disorder in the corticobasal degeneration syndrome, *Neuropsychologia*, 45:3015–3023.

D.R. Dowty, 1979, *Word meaning and Montague grammar*, Reidel, Dordrecht.

D.R. Dowty, 1991. Thematic Proto-Roles and Argument Selection, *Language*, 67(3):547-619.

M. Haspelmath. 1993. *More on the typology of inchoative/causative verb alternations*. In: B. Comrie and M. Polinsky (eds.), *Causatives and transitivity*, Amsterdam/Philadelphia, Benjamins, pp. 87-120.

M. Haspelmath. 2001. *Non-canonical marking of core arguments in European languages*. In: A.Y. Aikhenvald, R.M.W.Dixon and M. Onishi. 2001. pp. 53-83.

J.R. Hodges, K. Patterson, S. Oxbury, E. Funnell. 1992. Semantic dementia: progressive fluent aphasia with temporal lobe atrophy. *Brain,* 115:1783–806.

J.R. Hodges and K. Patterson, 1996. Non-fluent progressive aphasia and semantic dementia: a comparative neuropsychological study, *J Int Neuropsychol Soc,* 2: 511–24.

J.R. Hodges and K. Patterson, 2007. Semantic dementia: a unique clinicopathological syndrome. *Lancet Neurol*, 6: 1004–1014.

M.A. Lambon Ralph and K. Patterson. 2008. Generalization and Differentiation in Semantic Memory Insights from Semantic Dementia, *Ann. N.Y. Acad. Sci.* 1124: 61–76.

B. Levin and M. Rappaport Hovav. 1995. *Unaccusativity. At the Syntax–Lexical Semantics Interface*, The MIT Press, Cambridge, MA.

B. Levin and M. Rappaport Hovav. 2001. An event structure account of English resultatives, *Language*, 77(4):766-797.

L. Meteyard and K. Patterson. 2009. The relation between content and structure in language production: An analysis of speech errors in semantic dementia, *Brain and Language*, 110(3):121-134.

D. Neary, J.S. Snowden, L. Gustafson, U. Passant, D. Stuss, S. Black, et al.1998. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology*, 51:1546–54.

C. Papagno, R. Capasso and G. Miceli, Reversed concreteness effect for nouns in a subject with semantic dementia. *Neuropsychologia*, 47:1138–1148.

K. Patterson, M.A. Lambon Ralph, J.R. Hodges and J.L. McClelland. 2001. Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge, *Neuropsychologia* 39: 709–724.

Patterson, K., & MacDonald, M. C. (2006). Sweet nothings: Narrative speech in semantic dementia. In S. Andrews (ed.), *From Inkmarks to Ideas: Current Issues in Lexical Processing*, Psychology Press, Hove, UK, pp. 288–317.

D.M. Perlmutter. 1978. Impersonal passives and the unaccusative hypothesis*, Proceedings of the Forth Annual Meeting of the Berkeley Linguistics Society*, 4:157-189.

E. Rochon, G. Kavé, J. Kupit, R. Jokel and G. Winokur. 2004. Sentence Comprehension in Semantic Dementia: A Longitudinal Case Study, *Cognitive Neuropsychology*, 21(2/3/4):317-330.

M. Schwartz, O.S.M. Marin and E.M. Saffran. 1979. Dissociation of Language Function in Dementia: A Case Study, *Brain and Language*, 7:277-306.

A.Sorace. 2000. Gradients in Auxiliary Selection with Intransitive Verbs, *Language*, 76(4):859-890.

L.K. Tyler, H.E. Moss, K. Patterson and J. Hodges. 1997. The gradual deterioration of syntax and semantics in a patient with Progressive Aphasia, *Brain and Language*, 56:426-476.

L.K. Tyler, E.A. Stamatakis, R.W. Jones, Peter Bright, K.Acres, and W.D. Marslen-Wilson. 2004. Deficits for Semantics and the Irregular Past Tense: A Causal Relationship? *Journal of Cognitive Neuroscience* 16(7):1159–1172.

C.Tenny. 1994. *Aspectual Roles and the Syntax-Semantics Interface*, Kluwer, Dordrecht.

R.D.Jr.Van Valin. 1990. Semantic parameters of split intransitivity, *Language*, 66: 221-260.

R.D.Jr.Van Valin and R.LaPolla. 1997. *Syntax: Structure, Meaning & Function*, Cambridge, University Press.

M. Visser, E. Jefferies and M.A. Lambon Ralph. 2010. Semantic Processing in the Anterior Temporal Lobes: A Meta-analysis of the Functional Neuroimaging Literature, *Journal of Cognitive Neuroscience* 22 (6):1083–1094.

E.K. Warrington. 1975. Selective impairment of semantic memory. *QJ Exp Psychol*, 1975: 635–657.

A. Zaenen. 1993. *Unaccusativity in Dutch: Integrating Syntax and Lexical Semantics*. In: J.Pustejovsky, *Semantics and the Lexicon*, Kluver, Dordrecht, pp. 129-161.

A Specialised Verb Lexicon as the Basis of Fact Extraction in the Biomedical Domain

C.J. Rupp, Paul Thompson, William Black, John McNaught and Sophia Ananiadou
National Centre for Text Mining, University of Manchester, UK
c.j.rupp@cs.man.ac.uk,
{paul.thompson,william.black,john.mcnaught,sophia.ananiadou}@manchester.ac.uk

Abstract

The BioLexicon is a standardised, reusable, lexical and conceptual resource suitable for advanced biomedical text mining. One of the unique features of the BioLexicon is the incorporation of rich syntactic and semantic patterns for a wide range of domain-relevant verbs, which have been acquired semi-automatically from biomedical corpora. Such types of information can be highly beneficial for information and fact extraction applications. In this paper, we describe the collection of the verb-specific information for inclusion in the BioLexicon, and explain how it is being employed in a specific scenario (the UKPMC project) to leverage fact-based information extraction on a large collection of biomedical papers.

1 Introduction

Information Extraction (IE) applications that focus on extraction of event information require sophisticated lexical resources, which include both syntactic patterns for verbs, and their corresponding semantic interpretations (in terms of semantic roles). The BioLexicon (Sasaki et al., 2008) is the first large-scale, specialised lexical resource that includes such information for a wide range of domain-relevant verbs.

The BioLexicon is being evaluated within the context of the UK PubMed Central (UKPMC) project, where it is used as a fulcrum to leverage fact-based information extraction over a large collection of approximately 1.8M research articles in the biomedical domain.

The UKPMC project is developing a specific search portal for UK researchers, which accesses the PubMed Central repository of biomedical research, and adds additional functionalities, including text mining capabilities.

Among the text mining applications under development is a semantic search engine that allows specific facts to be located within the document collection. To support this application, an extensive index of analysed facts occurring within the collection is under compilation. The detailed information provided in the BioLexicon regarding the behaviour of verbs in the biomedical domain forms the primary criterion for recognising facts and extracting their constituents.

In the first part of this paper, we describe the motivation for the construction of the BioLexicon, followed by a description of the collection of the verb-specific information contained within it. In the second part, we explain in more detail the method by which the BioLexicon is employed in the context of the UKMPC project, and provide an initial evaluation of this method.

2 Lexical Resources for IE

A number of large-scale computational lexicons containing syntactic and semantic information for verbs and other parts-of-speech have been developed for general English language, e.g., FrameNet (Rupenhoffer et al., 2006), Propbank (Palmer et al., 2005) and VerbNet (Kipper-Schuler, 2005). However, these resources are not well suited for use in IE systems that operate in specialized domains such as biomedicine, and may lead to incorrect analyses.

Descriptions of events in biomedical texts have a number of domain-specific features. Firstly, there are verbs that appear rarely in general language texts (e.g., *phosphorylate*), and hence are not accounted for in the general language resources. Secondly, verbs that occur frequently in both general language and biomedical texts often have different syntactic or semantic properties in each domain, e.g., differing numbers of arguments (Wattarujeekrit et al., 2004) or different meanings. In addition, strongly selected modifiers (such as location, manner and timing), are considered to be much more important to the correct interpretation of biomedical events than general language events (Tsai et al., 2007). This is exemplified in the following sentence, which specifies both a manner and a location for the event described by the verb *directs*:

*A promoter has been identified that **directs** relA gene transcription towards the pryG gene in a counterclockwise direction on the E. coli chromosome*

Although domain-specific extensions of general language computational lexicons have been attempted, e.g., BioFrameNet (Dolbey et al., 2006) and PASBio (Wattarujeekrit et al., 2004), their coverage is very limited. The SPECIALIST lexicon (Browne et al., 2003) is a larger resource containing biomedical vocabulary. Although syntactic complementation patterns are included for verbs, they are somewhat limited, and based on general language patterns. In addition, no semantic information for verbs is provided.

3 The BioLexicon

The BioLexicon[1] (Sasaki et al., 2008) is a standardised, reusable, lexical and conceptual resource suitable for advanced biomedical text mining, containing over 2.2M lexical entries, with a particular emphasis on gene regulation.

Whilst the vast majority of entries in the BioLexicon correspond to biomedical terms, a major design criterion was to include syntactic and semantic patterns for a wide range of domain-relevant verbs, in order to address the previous lack of a suitable resource. The BioLexicon thus incorporates 658 domain-relevant verbs, all of which are accompanied by syntactic subcategorization frames, and 168 of which include semantic event frames, as well as explicit linking between the syntactic and semantic levels.

A corpus-based approach was taken to the construction of the verbal part of the lexicon to ensure that the behaviour of the verbs recorded in the lexicon reflects the way they are used in domain-specific texts. In contrast to the manual construction of many other lexical semantic resources, the verbal information in the BioLexicon was derived semi-automatically, using different techniques and different sizes of corpora to obtain each type of information.

3.1 Syntactic Subcategorisation Frames

The extraction of subcategorisation frames (SCFs) (Venturi et al., 2009) used an unsupervised learning technique, applied to a corpus of approximately 6M tokens (both MEDLINE abstracts and full biomedical papers) on the subject of E. coli. The corpus was automatically annotated for predicate-argument structure using a version of the Enju parser tuned to biomedical texts (Hara et al, 2005). Based on the parse results, observed dependency sets (ODSs) were computed for each verbal occurrence and used as the basis of the SCFs.

Each ODS is represented as a set of dependencies described in terms of relation type (e.g. ARG1, ARG2, etc.). The order of the dependencies in each ODS is normalised and does not reflect their order of occurrence in context. According to their importance in biomedical events, the induced SCFs include strongly-selected modifiers as well as strongly selected arguments in the description of biomedical events.

For each ODS, the conditional probability given the verb was computed. Thresholding based on this probability was used to filter out noisy frames (i.e., frames containing not only arguments and strongly selected modifiers, but also adjuncts) as well as possible errors of either parsing or ODS extraction. The remaining 1760 ODSs (distributed amongst the 658 verbs) were selected as SCFs for inclusion in the BioLexicon.

3.2 Semantic Event Frames

The extraction of event frames was carried out on a subset (677 abstracts) of the corpus used for SCF extraction. Each abstract was manually annotated with gene regulation events, centred on both verbs and nominalised verbs, by a group of domain experts (Thompson et al., 2008). For each event, semantic arguments occuring within the same sentence were labelled with both semantic roles and named entity (NE) types.

Although somewhat comparable to the GENIA event annotation (Kim et al., 2008), our annotation differs in that it was geared specifically towards the acquisition of semantic frame information for verbs, using a richer set of semantic roles to capture detailed information regarding verb behaviour. Our corpus uses a total of 13 roles (compared to 6 in GENIA), which are intended to characterise all the sublanguage semantic arguments of relevant events.

The semantic roles used in the BioLexicon are event-independent, and constitute a closed set, which is advantageous in facilitating generalization over different types of events (Cohen and Hunter 2006; Merlo and Plas 2009). Although application of a closed semantic role set to general language events may be problematic (Palmer et al., 2005), the use of such a set is more viable in a restricted domain, as domain-specific definitions can be provided for each semantic role type.

[1] http://catalog.elra.info/product_info.php?products_id=1113

Role Name	Description	**Example** ([…] = semantic argument, small capitals = focussed verb)
AGENT	Drives/instigates event	[The narL gene product] ACTIVATES the nitrate reductase operon
THEME	a) Affected by/results from event b) Focus of events describing states	[recA protein] was INDUCED by UV radition [The FNR protein] RESEMBLES CRP
MANNER	Method/way in which event is carried out	cpxA gene INCREASES the levels of csgA transcription by [dephosphorylation] of CpxR
INSTRUMENT	Used to carry out event	EnvZ FUNCTIONS through [OmpR] to control NP porin gene expression in E. Coli.
LOCATION	Where *complete* event takes place	Phosphorylation of OmpR MODULATES expression of the ompF and ompC genes in [Escherichia coli]
SOURCE	Start point of event	A transducing lambda phage was ISOLATED from [a strain] harboring a glpD’’lacZ fusion
DESTINATION	End point of event	Transcription is activated by BINDING of the cyclic AMP (cAMP)-cAMP receptor protein (CRP) complex to [a CRP binding site]
TEMPORAL	Situates event in time/ w.r.t. another event	The Alp protease activity is DETECTED in cells [after introduction] of plasmids
CONDITION	Environmental conditions/changes in conditions	Strains carrying a mutation in the crp structural gene fail to REPRESS ODC and ADC activities in response to [increased cAMP]
RATE	Change of level or rate	marR mutations ELEVATED inaA expression by [10- to 20-fold] over that of the wild-type.
DESCRIPTIVE-AGENT	Descriptive information about AGENT of event	HyfR ACTS as [a formate-dependent regulator]
DESCRIPTIVE-THEME	Descriptive information about THEME of event	The FNR protein RESEMBLES [CRP].
PURPOSE	Purpose/reason for the event occurring	The fusion strains were USED [to study] the regulation of the cysB gene

Table 1: Semantic roles and definitions

Our semantic roles are based largely on the verb-independent roles used in VerbNet (Kipper-Schuler, 2005) and SIMPLE (Lenci et al, 2000). Through the examination of a large number of relevant events within MEDLINE abstracts, in consultation with biologists, it was concluded that arguments of gene regulation events may be characterised using a subset of these general language roles, with some name changes to make them more easily understandable to biologists, and with the addition of the domain-specific CONDITION role, corresponding to descriptions of environmental conditions. The full set of roles is shown in Table 1.

NE categories are organised into 5 different hierarchies, corresponding to the following 5 supercategories: *DNA, PROTEIN, EXPERIMENTAL, ORGANISMS* and *PROCESSES.* The categories are mapped to classes in the Gene Regulation Ontology (GRO) (Beisswanger et al, 2008).

A set of 856 verb-specific semantic frames was extracted from the annotated corpus for inclusion in the BioLexicon. We have chosen to create verb-specific frames, as these allow more detailed argument specifications than those resources that group verbs into classes (e.g.,. VerbNet, FrameNet). The importance within the domain of phrases that identify location, manner, timing and condition mean that individual verbs can behave idiosyncratically.

Extracted semantic frames include the semantic roles annotated, in addition to NE types, if available. These allow selectional restrictions to be applied to the fillers of each role. An example event frame is as follows:

```
activate(Agent=>Protein,
         Theme=>DNA)
```

3.3 Linking Syntactic and Semantic Frames

Syntactic arguments of predicates have been manually linked to their semantic counterparts in the event frames, in order to facilitate the automatic labelling of syntactic arguments of verbs with semantic roles. This step was carried out for the 168 verbs for which both subcategorisation and event frame information was available, taking into account the following types of information:

a) General linguistic constraints regarding the alignment of hierarchies of semantic roles and grammatical functions. Given a semantic role hierarchy (agent>theme ...) and a grammatical functions hierarchy (subject>object ...), the mapping usually proceeds from left to right;

b) A list of 'prototypic' grammatical realisations of semantic arguments;

c) General language repositories of individual semantic frames containing both syntactic and semantic information.

4 Fact-Based Information Extraction

The sheer volume of publications in biomedicine has made it a focus for text mining research. Much of this activity involves named entity recognition (NER), i.e., the identification of technical terms and designations relevant for the domain. Text mining systems may either manage the documents themselves, i.e. information retrieval, or the information contained within the documents, i.e. information extraction (IE). IE applications aim to locate relations between entities, e.g., Hoffman and Valencia (2004). These relations may be evidenced by proximity in the text, or inferred based on domain knowledge. The most specific relations are the claims explicitly made in the text detailing the research itself.

We aim to support searching over the evidence and claims presented in research papers by indexing the relations that occur at the lexical level. We refer to the combinations of lexical relations and arguments, which typically centre on verbs or deverbal nominalizations, as *facts*[2].

The detailed information encoded about verbs in the BioLexicon forms the keystone of an IE method applied to the UKPMC corpus. In order to support queries against this collection focussing on specific evidence presented in the text, we analyse the verbal relations, along with their argument structure and predicted modifiers. We extract representations of the key facts, and then index these for efficient query and retrieval. The BioLexicon provides the information that ultimately decides which constructions are recorded as facts.

There are three knowledge sources used in the fact extraction process:

The papers are syntactically analysed using the Enju parser, which is the same parser used in the development of the BioLexicon. It has been optimised for the biomedical domain by the use of a parse preference model (Hara et al., 2005), meaning that we can be confident in selecting only the highest rated parse for each sentence. The size of the collection would make the considering competing parses impractical.

The extended verb frames of the BioLexicon, which provide patterns of argument structure and systematic modification, are used as predictions of the arguments and modifier structures that can identify relevant facts within the domain.

Within the UKPMC project, a suite of standardised NER recognisers for various classes of named entities is used across all applications. The NER results play a significant role in determining which facts to extract and index.

The fact extraction process consists of three steps, each refining the set of potential facts more precisely. Firstly, we locate within the Enju parse result those verbs with corresponding entries in the BioLexicon. Next, we require that at least one of the named entities recognised by the NER components be involved in the relation centred on the verb, either as an argument or as a predicted modifier. Finally, we ensure that whole construction is consistent with a verb frame definition in the BioLexicon.

As an example, consider the two syntactic frames provided in the BioLexicon for *induce*:

```
1) induce,ARG1#ARG2#
2)induce,ARG1#ARG2#PP-in#
```

NER recognises *mutant p53* as a protein. This allows us to add to our index sentences where this protein appears with verb *induce*, either as the subject, object or a prepositional modifier headed by *in*, e.g.:

This scenario suggests that <u>mutant p53</u> could use different mechanisms to ***induce*** *malignant properties in epidermal keratinocytes.*

The overall conclusion from our work is a direct relationship between chemoresistance ***induced*** *by <u>mutant p53</u> and its transactivation ability.*

[2] However, we are aware that not all of the claims they express are factual

The primary search domain of this extraction process is the analysis tree provided by the Enju parser. The alignment of recognised named entities with analysed constituents is performed via the standoff annotations provided by each component. The relevant sections of text are retrieved and recorded, in order to present search results from the index. The information from the BioLexicon is the primary filter and determines the final choice of facts to be indexed, but the results of both parsing and NER make a significant contribution.

5 Evaluation of the BioLexicon for Fact Extraction

An initial quantitative evaluation of the method described above has been carried out on a subset of the UKPMC corpus, consisting of approximately 80,000 documents.

On the one hand, the BioLexicon is a strong filter, in that only the verbs it recognises are accepted as the basis of a fact. This is restrictive in that only a certain part of the domain covered by the collection is within the remit of the BioLexicon, i.e., gene regulation. The evaluation results confirm this filtering effect: only 62.7% of the instances of the verbs present in the document collection matched verbal entries in the BioLexicon. A still stronger filter is the requirement that a domain relevant NE should be present in one of the arguments, resulting in only 16.9% of the total verb instances present in the text collection being extracted as facts.

On the other hand, the lexicon also has a boosting effect on the fact base, since modifier phrases are explored which would not be considered without its input. Where these modifier phrases contain recognised named entities, this can provide enough evidence for the extraction of a fact that would not otherwise be recorded. Consider the following example:

The pXPC3 plasmid ***codes*** *for <u>an XPC cDNA</u> that is truncated by 160 bp from the N terminus compared with the wild-type XPC cDNA*

Although the Enju parse result treats *codes* as an intransitive verb, the information in the BioLexicon allows the THEME role to be assigned to the PP headed by *for.*

This boosting effect is demonstrated in the evaluation results: 9.7% of verb arguments are detected in prepositional modifier phrases, rather than in the arguments initially predicted by the parser output.

In addition to the argument and modification patterns predicted in the BioLexicon, the fact index also records patterns of negation and some other scoped modifications that are independent of the lexical predictions. We are thus able to distinguish between logically related facts retrieved in a query-based application.

5.1 Conclusion and Further Work

This paper has described the verbal component of the BioLexicon, which is a unique resource comprising rich linguistic information suitable for text mining applications operating within the biomedical domain. The corpus-driven nature of the acquisition of both syntactic and semantic information for verbs aims to facilitate the accurate identification of events, together with their participants and the semantic roles assigned to them. Such comprehensive information is not currently available in any comparable domain-specific resource.

The BioLexicon is at the heart of an IE method that is being employed to facilitate fact-based querying over a large collection of biomedical documents as part of the UKPMC project. The preliminary results provide compelling evidence that the BioLexicon can assist in building powerful tools for fact extraction within the biomedical domain.

We are currently in the process of developing applications based on the fact index extracted for the UKPMC corpus (see Black et al. (2010)).

The utility of the BioLexicon has also been shown to extend beyond IE applications; the recognition of multiword terms in the lexicon can help with a number of NLP tasks in the biomedical domain including POS tagging and syntactic parsing (Sasaki et al., 2009) and improving the performance of information retrieval (Sasaki et al., 2010).

Acknowledgments

This research was supported by the European Commission IST project FP6-028099 (BOOTStrep) and the UKPubMedCentral (UKPMC) project. UK PubMed Central is funded by: Arthritis Research Campaign, BBSRC, British Heart Foundation, Cancer Research UK, Chief Scientist Office, Department of Health - National Institute of Health Research, Medical Research Council, Wellcome Trust. We also thank ILC-CNR, Italy for their work on the production of the subcategorisation frames.

References

E. Beisswanger, V. Lee, J.J Kim, D. Rebholz-Schuhmann, A. Splendiani, O. Dameron, S. Schulz and U. Hahn. 2008. Gene Regulation Ontology (GRO): design principles and use cases. *Studies in Health Technology and Informatics*, 136:9-14.

W.J. Black, C.J. Rupp, C. Nobata, J. McNaught, J. Tsujii and S. Ananiadou. 2010. High-Precision Semantic Search by Generating and Testing Questions. In *Proceedings of the UK e-Science All Hands Meeting 2010.*

A.C. Browne, G. Divita, A.R. Aronson and A.T. McCray. 2003. UMSL Language and Vocabulary Tools: AMIA 2003 Open Source Expo. In *Proceedings of AMIA Annual Symposium 2003*, page 798

K.B. Cohen and L. Hunter. 2006. A critical review of PASBio's argument structures for biomedical verbs. *BMC Bioinformatics* 7 (Suppl. 3), S5.

A. Dolbey, M. Ellsworth and J. Scheffczykx. 2006. BioFrameNet: A Domain-Specific FrameNet Extension with Links to Biomedical Ontologies. In *Proceedings of KR-MED,* pages 87-94.

T. Hara, Y. Miyao and J. Tsujii. 2005. Adapting a probabilistic disambiguation model of an HPSG parser to a new domain. In *Proceedings of IJCNLP*, pages 199-210

R. Hoffmann, R. and A. Valencia. 2004. A Gene Network for Navigating the Literature. *Nature Genetics* 36:664

J.D. Kim, T. Ohta and J. Tsujii. 2008. Corpus annotation for mining biomedical events from literature *BMC Bioinformatics* 9:10.

K. Kipper-Schuler. 2005. *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD. Thesis. Computer and Information Science Dept., University of Pennsylvania, Philadelphia.

A. Lenci, F. Busa, N. Ruimy, E. Gola,, M. Monachini, N. Calzolari, A. Zampolli, E. Guimier, G. Recourcé, L. Humphreys, U. Von Rekovsky, A.. Ogonowski, C. McCauley, W. Peters, I. Peters, R. Gaizauskas and M. Villegas. 2000. *SIMPLE Linguistic Specifications LE-SIMPLE (LE4-8346), Deliverable D2.1 & D2.2*. ILC and University of Pisa

P. Merlo and L. Plas. 2009. Abstraction and Generalisation in Semantic Role Labels: PropBank, VerbNet or both? In *Proceedings of ACL-IJCNLP 2009*, pages 288-296

M. Palmer, P. Kingsbury and D. Gildea. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71-106.

J. Ruppenhofer, M. Ellsworth, M. Petruck, C. Johnson and J. Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice*, Available online at http://framenet.icsi.berkeley.edu/

Y. Sasaki, S. Montemagni, P. Pezik, D. Rebholz-Schuhmann, J. McNaught and S. Ananiadou 2008. BioLexicon: A Lexical Resource for the Biology Domain. In *Proceedings of the SMBM 2008*, pages 109-116.

Y. Sasaki, P. Thompson, J. McNaught and S. Ananiadou. 2009. Three BioNLP Tools Powered by the BioLexicon. In *Proceedings of EACL Demonstration Session*, pages 61-64

Y. Sasaki, J. McNaught and S. Ananiadou. 2010. The value of an in-domain lexicon in genomics qa. *Journal of bioinformatics and computational biology*, 8(1):147-161.

P. Thompson, P. Cotter, S. Ananiadou, J. McNaught, S. Montemagni, A. Trabucco and G. Venturi .2008. Building a Bio-Event Annotated Corpus for the Acquisition of Semantic Frames from Biomedical Corpora. In *Proceedings of LREC 2008*, pages 2159-2166.

R.T.H Tsai, W.C. Chou, Y.S. Su, Y.C. Lin, C.L. Sung,. H.J Dai, I.T.H. Yeh, W. Ku, T.Y. Sung and W.L. Hsu. 2007. BIOSMILE: A semantic role labeling system for biomedical verbs using a maximum-entropy model with automatically generated template features. *BMC Bioinformatics* 8:325

G. Venturi, S. Montemagni, S. Marchi, Y. Sasaki, P. Thompson, J. McNaught, and S. Ananiadou. 2009. Bootstrapping a Verb Lexicon for Biomedical Information Extraction. In *Proceedings of CICLing 2009*, pages 137-148.

T. Wattarujeekrit, P. Shah and N. Collier. 2004. PASBio: predicate-argument structures for event extraction in molecular biology. *BMC Bioinformatics* 5:155

Investigating characteristics of semantic networks of verbs in patients with Alzheimer's disease

Anderson Santos[1], Gustavo Garcia Valdez[1], Aline Villavicencio[1], Jerusa Salles[2]
Institute of Informatics[1]
Institute of Psychology[2]
Federal University of Rio Grande do Sul, Brazil
{arssantos, ggvaldez, avillavicencio}@inf.ufrgs.br,
jerusafs@yahoo.com.br

Abstract

Alzheimer's disease produces alterations of cognitive functions and of processes that are responsible for language and memory. In order to have a better understanding of language changes, we investigate the characteristics of the semantic networks of patients diagnosed with probable Alzheimer, focusing on verbs. The results of comparisons with networks of healthy individuals highlights some topological differences among them.

1 Introduction

It is estimated that 35.6 million people currently suffer from dementia and that in 20 years this number will reach 65.7 million of individuals[1], with an estimated overall treatment cost of 315 billion dollars per year in the world. Alzheimer's disease (AD) is responsible for more than 50% of the cases of dementia, and it is one of the pathologies that cause among other consequences, alteration of cognitive functions and of the processes that are responsible for language and memory (Mansur, Carthery, Caramelli, & Nitrini, 2005).

Although there is no consensus about the precise nature of the changes in semantic memory change (Mansur, Carthery, Caramelli, & Nitrini, 2005), based on the results of semantic memory tests such as the Hodges Battery (Hodges, Salmon, & Butters, 1992; Howard & Patterson, 1992), two main theories are proposed to explain the semantic deficits of cognitive performance on these explicit semantic tests. The first one proposes a degradation of the semantic memory itself while the second advocates for a failure to retrieve information from memory (Mansur, Carthery, Caramelli, & Nitrini, 2005; Rogers & Friedman, 2008).

In relation to the language capacity, previous studies have found a progressive deterioration of performance in phonetic-phonological, syntactic, semantic and pragmatic-discursive processes (Mac-Kay, Assêncio-Ferreira, & Ferri-Ferreira, 2003; Mansur, Carthery, Caramelli, & Nitrini, 2005; Ortiz, 2009). For instance, in the context of aphasia, which may result from a progressive neurological disease like Alzheimer's, there seems to be a preference for more general and frequent verbs to be more easily used (Barde, Schwartz, & Boronat, 2006; Breedin, Saffran, & Schwartz, 1998; Kim & Thompson, 2004; Thompson, 2003; Thompson & Shapiro, 2007), which may be due to these verbs being applicable in many distinct situations. Closely related factors such as polissemy and synonymy are also seen as an important role in the human learning process (Hills, Maouene, Maouene, Sheya, & Smith, 2009). Features like this may influence the organization of the mental lexicon arising, e.g., from the need of fast retrieval of concepts (Steyvers & Tenenbaum, 2005).

In this paper we investigate the characteristics of the semantic networks of AD patients[2], focusing on the lexical organization of verbs. For that we use psycholinguistic data from an action

[1] Figures from the Alzheimer's Disease International, 2009.

2 Due to the impossibility of detecting the presence of histological brain features in living elderly individuals, the diagnostics is of probable or possible Alzheimer Disease (McKhann et al., 1984).

naming task, comparing the output of AD patients with those from healthy individuals. We represent the data as semantic networks, which seem to play an important role in the modeling of the organization of lexical knowledge and have been used to describe access to the mental lexicon (Steyvers & Tenenbaum, 2005). We analyse the collective[3] semantic networks using statistical and topological analysis.

This paper is structured as follows: in section 2 we describe some relevant works on semantic networks. In section 3 we present the materials and methods in the experiments. In section 4 we present the results. We finish with some conclusions and future works.

2 Related Works

Semantic networks have been used in several studies of language. For instance, Steyvers and Tenenbaum (2005) analyzed the large scale structures of three kinds of semantic networks: word associations of naïve subjects (Nelson, McEvoy, & Schreiber, 1999), WordNet (Miller, Fellbaum, Gross, & Miller, 1990) and a thesaurus (Roget, 1911). All three networks have the features of small-world structure, characterized by the combination of short-average minimal path lengths (L)[4] and a high clustered neighborhood (extracted from the clustering coefficient, C, that represents the probability of two random nodes being neighbors). The results found suggest that these characteristics may be related to the cognitive need for the fast retrieval of concepts (Steyvers & Tenenbaum, 2005). Indeed, Sigman and Cecchi (2002) also found a small-world structure in a network of nouns constructed from four types of semantic relations in WordNet: hyponymy/hypernymy; antonymy; meronymy/holonymy; polysemy. However, it is only when the polysemy links are added to the network that it becomes a small world (Sigman & Cecchi, 2002). Similarly, a network constructed from synonyms from the Moby thesaurus (Motter, de Moura, Lai, & Dasgupta, 2002) also had small-world structure. In this paper we follow these works, and in particular Steyvers and Tenenbaum (2005) and Sigman and Cecchi (2002) in using topological analysis for comparing the semantic networks.

3 Collective networks are modeled using a group of individuals, rather than only one.

4 A minimal path length is the minimal distance between two nodes in the network.

Semantic networks have also been using in cross-linguistic investigations like that of Parente et al. (2011) who compared the semantic networks of Brazilian Portuguese speakers and Mandarin Chinese in a verb naming task, in the context of language acquisition (Parente et al., 2011). In this work we also use a verb naming task but this time to investigate possible changes in the semantic networks of AD patients.

3 Materials & Methods

Participants for the verb naming task consisted of 46 individuals divided into 2 groups:

•Alzheimer's Disease (AD) group: 23[5] patients diagnosed with probable Alzheimer's disease (Mild AD), with Mean age = 75.6 years; SD = 6.7 and

•Healthy Elderly (HE) group: 23 healthy individuals with Mean age = 72.4 years; SD = 8.2.

In addition, a third group of participants was also considered for evaluation purposes:

•Healthy Young Adult (HYA) group: with 75 adults (Mean age = 21.69; SD = 3.25).

The experimental materials consisted of 17 movies showing destruction or division actions which always included an agent, an instrument and an object (e.g. sawing a log and cutting paper) (Duvignau & Gaume, 2004; Tonietto et al., 2008). The participants were asked to name the action portrayed, and the answer given by each participant for each movie was recorded. These actions were selected according to criteria of easiness of understanding. All responses that contained verb were considered valid, if the verb was related to the main action (excluding e.g. "to eat" for the action of sawing a log) and if the answer was not metalinguistic (excluding e.g. "I don't know") or non-verbal.

For each of the elderly groups (AD and HE) one semantic network was created, where every distinct verb uttered by a participant of the group was represented by a node in the network. A link

5 The size of this sample is compatible with that of other works with Alzheimer's disease: some report from 5 to 11 patients, and others have from 20 to 26 patients (Bell, Chenery, & Ingram, 2001; Chan, 1997; Garrard, Lambon Ralph, Patterson, Pratt, & Hodges, 2005; Laisney et al., 2009; Peraita, Daz, & Anllo-Vento, 2008; Rogers & Friedman, 2008). This is partly due to the difficulties of finding a larger sample of participants with the same level of the disease (in this case Mild level).

between two nodes (verbs) was added to the network if the two verbs were uttered for the same action. The result was a clique formed by all verbs given for a movie, and the different cliques became connected due to the polysemy of some of the verbs, which were produced for more than one movie.

A comparison of the two groups is done in terms of their structure, through topological analysis, and also of their content. Table 1 shows some relevant topological measures, where:

•<*k*> is the mean degree of the network;
•*L* is average minimal path length;
•*D* is the diameter of the network (with a maximum of L nodes);
•*C* is the clustering coefficient and

More details about each of these measures can be found in (Barabási & Albert, 1999; Watts & Strogatz, 1998).

4 Results

The results are discussed in terms of two comparisons. In the first we compare the semantic networks of the two groups with each other. The results are further evaluated by first determining the expected differences that would arise from a variation in the participants (using the HYA group) and comparing with the observed differences between the two elderly groups (AD and HE).

4.1 Elderly Groups

Sharing the same global features of the other language networks, these show a small world structure: they have a small average of minimal path lengths and high clustering coefficients.

Apart from their diameters, the two networks considerably differ in all other measures. First of all, the AD group produced more distinct verbs for describing the actions, which suggests lower agreement for describing the actions and is reflected in a slightly larger number of nodes than the HE group. As a consequence, although a larger number of edges would be expected with more nodes in the AD group and their mean connectivity (<*k*>) of the HE, the observed increase was considerably larger than that.

Second, the mean and standard deviation presented by <*k*> indicate that AD's nodes have a consistently higher degree of connectivity (*k*) than those in HE. One possibility for a larger k is the use of more polysemic verbs by the AD group, since for every action that a verb is used to describe, it becomes connected to all other verbs also used to describe the action, forming a clique. Therefore, for each new context in which a verb is used, its degree increases by the size of the clique. If we assume that more connected verbs are also more generic, this would be consistent with the tendency of aphasic patients to use more general verbs (Barde, Schwartz, & Boronat, 2006; Breedin, Saffran, & Schwartz, 1998; Kim & Thompson, 2004; Thompson, 2003; Thompson & Shapiro, 2007).

Third, with a larger number edges between the nodes and a higher mean connectivity, the average minimal path length (*L*) would be expected to be smaller in the AD group than in the HE. However, the opposite is found, which is an indication that the differences between the two networks go beyond the use of a larger vocabulary and less agreement between in the AD group, but that they are structurally different too

Variable	Alzheimer	Controls
n (verbs)	46	40
Edges	243	140
<*k*>	10.57 (SD 6,55)	7.00 (SD 4,56)
L	1.94	1.57
D	4	4
C	.829	.789

Table 1: A summary of the semantic networks

In Figure 1, we can see the two networks in which the size of a node is shown in direct proportion to its degree (normalized). The image suggests the a larger number of highly connected nodes, or hubs, in the AD network.

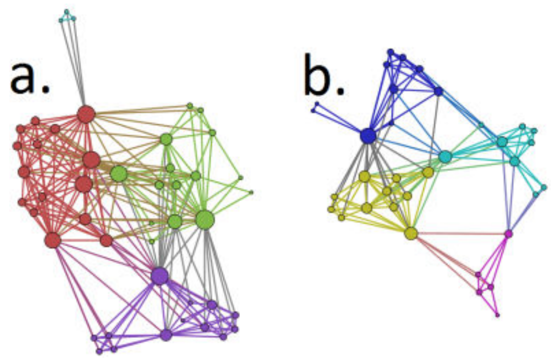


Figure 1: The Alzheimer's (a) and controls (b) semantic networks.

4.2 Adults and Elderly Groups

In order to verify the degree of variation expected from different groups of participants, and whether this variation could explain the differences found between the two elderly groups, we also created 30 subgroups of 23 participants randomly selected from the 75 in the HYA group. For each subgroup we generated a semantic network using the same method than for the elderly groups. Table 2 shows the mean and standard deviation of the topological features of the 30 groups. In addition this table also shows the module of the difference of statistics between the AD and the HE networks. All the differences are larger than the standard deviation of the adult's samples. This indicates that intra-group variations are not enough to explain the differences found between the elderly groups.

	Adults Sample		|AD-Controls|
Variable	**Mean**	**SD**	
n (verbs)	38.57	1.305	6
Edges	334	20.85	103
$<k>$	9.405	0.355	3.57
L	2.137	0.05	0.37
D	4.567	0.504	0
C	0.817	0.012	0.04

Table 2: Characterization of Sample of Adults. Including the difference between Alzheimer and controls networks.

5 Conclusions and Future Work

In this paper we presented an investigation of the lexical organization of verbs in the context of Alzheimer's disease patients. We looked at characteristics of the semantic network of verbs produced by AD patients in an action naming task, comparing with that of healthy individuals. We analysed the collective semantic networks using statistical and topological analysis, and found interesting divergences. In particular there seemed to be less agreement among the AD patients for the lexical choice to describe a given action. In addition, there were also indications of structural differences between the networks which may arise from modifications in the lexical organization caused by AD.

However, more detailed investigation of these possibilities needs to be conducted before more definite conclusions can be reached. We also plan to analyze qualitative differences among hubs between the networks. Finally we intend to inspect other statistical features of complex networks, particularly those related to network vulnerability (Criado, Flores, Hernández-Bermejo, Pello, & Romance, 2005), that are associated to network performance and helps to measure the response of complex networks subjected to attacks on vertices and edges.

Acknowledgments

This research was partly supported by CNPq (Projects 479824/2009-6 and 309569/2009-5), FINEP and SEBRAE (1194/07) and FAPERGS (08513752). We would also like to thank Maria Alice Parente, Lauren Tonietto and Bruno Menegola for providing the data.

References

Barabási, A. L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*, 509-512.

Barde, L. H., Schwartz, M. F., & Boronat, C. B. (2006). Semantic weight and verb retrieval in aphasia. *Brain and Language*, *97*(3), 266-278. doi: DOI: 10.1016/j.bandl.2005.11.002.

Bell, E. E., Chenery, H. J., & Ingram, J. C. (2001). Semantic priming in Alzheimer's dementia: evidence for dissociation of automatic and attentional processes. *Brain and language*, *76 (2)*, 130-144.

Breedin, S. D., Saffran, E. M., & Schwartz, M. F. (1998). Semantic factors in verb retrieval: an effect of complexity. *Brain and language*, *63*(1), 1-31. doi: 10.1006/brln.1997.1923.

Chan, a. (1997). The deterioration of semantic networks in patients with Alzheimer's disease: A cross-sectional study. *Neuropsychologia*, *35*(3), 241-248. doi: 10.1016/S0028-3932(96)00067-X.

Criado, R., Flores, J., Hernández-Bermejo, B., Pello, J., & Romance, M. (2005). Effective measurement of network vulnerability under random and intentional attacks. *Journal of Mathematical Modelling and Algorithms*, *4*(3), 307-316. doi: 10.1007/s10852-005-9006-1.

Duvignau, K., & Gaume, B. (2004). Linguistic, Psycholinguistic and Computational approaches to the lexicon: For early verb-learning based on analogy. *Cognitive Systems*, (Special issue on learning), 255-269.

Garrard, P., Lambon Ralph, M. a., Patterson, K., Pratt, K. H., & Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: a new approach. *Brain and language*, *93*(1), 79-94. doi: 10.1016/j.bandl.2004.08.003.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: preferential attachment or preferential acquisition? *Psychol Sci*, *20*(6), 729-739. Retrieved from http://www.biomedsearch.com/nih/Longitudinal-analysis-early-semantic-networks/19470123.html.

Hodges, J. R., Salmon, D. P., & Butters, N. (1992). Semantic memory impairment in Alzheimer's disease: Failure of access or degraded knowledge. *Neuropsychologia*, *30*, 301-314.

Howard, D., & Patterson, K. (1992). Pyramids and palm trees: A test of semantic access from pictures and words. Thames Valley Publishing. *Lambon Ralph, M. A., McClelland, J. L., Patterson, K., Galton, C. J., & Hodges, J. R*, *13*(3), 341-356.

Kim, M., & Thompson, C. K. (2004). Verb deficits in Alzheimer's disease and agrammatism: Implications for lexical organization. *Brain and Language*, *88*(1), 1-20. doi: DOI: 10.1016/S0093-934X(03)00147-0.

Laisney, M., Giffard, B., Belliard, S., de la Sayette, V., Desgranges, B., Eustache, F., et al. (2009). When the zebra loses its stripes: Semantic priming in early Alzheimer's disease and semantic dementia. *Cortex; a journal devoted to the study of the nervous system and behavior*, 1-12. doi: 10.1016/j.cortex.2009.11.001.

Mac-Kay, A., Assêncio-Ferreira, V., & Ferri-Ferreira, T. (2003). *Afasias e demências avaliação e tratamento fonoaudiológico*. (L. E. Santos). São Paulo.

Mansur, L. L., Carthery, M. T., Caramelli, P., & Nitrini, R. (2005). Linguagem e cognição na doença de Alzheimer. *Psicologia: Reflexão e Crítica*, *18*(3), 300-307. doi: 10.1590/S0102-79722005000300002.

Miller, G. A., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography*, (3), 235-244.

Motter, A., de Moura, A., Lai, Y., & Dasgupta, P. (2002). Topology of the conceptual network of language. *Physical Review E*, *65*(6), 65102. APS. Retrieved from http://link.aps.org/doi/10.1103/PhysRevE.65.065102.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). *The University of South Florida word association norms.*. Retrieved from http://www.usf.edu/FreeAssociation.

Ortiz, K. Z. (2009). *Distúrbios Neurológicos Adquiridos: Linguagem e Cognição* (2 ed., p. 510). Barueri.

Parente, M. A., Villavicencio, A., Siqueira, C., Ping, L., Tonietto, L., Duvignau, K., et al. (2011). Lexical Bootstrapping Hypothesis (LBH) and conventionality: The contribution of a crosslinguistic study in verb acquisition by Chinese Mandarin- and Brazilian Portuguese-speaking children. In D. Bittner & N. Kuehn, (eds.). Lexical Bootstrapping.

Peraita, H., Díaz, C., & Anllo-Vento, L. (2008). Processing of semantic relations in normal aging and Alzheimer's disease. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, *23*(1), 33-46. doi: 10.1016/j.acn.2007.09.001.

Rogers, S. L., & Friedman, R. B. (2008). The underlying mechanisms of semantic memory loss in Alzheimer's disease and semantic dementia. *Neuropsychologia*, *46*(1), 12-21. doi: 10.1016/j.neuropsychologia.2007.08.010.

Roget, P. M. (1911). Roget's Thesaurus of English Words and Phrases.

Sigman, M., & Cecchi, G. A. (2002). Global organization of the Wordnet lexicon. *Proc. Nat. Acad. Sci*, *99*(3), 1742-1747.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41-78.

Thompson, C. (2003). Unaccusative verb production in agrammatic aphasia: the argument structure complexity hypothesis. *Journal of Neurolinguistics*, *16*(2-3), 151-167. doi: 10.1016/S0911-6044(02)00014-3.

Thompson, C. K., & Shapiro, L. P. (2007). Complexity in treatment of syntactic deficits. *American journal of speech-language pathology / American Speech-Language-Hearing Association*, *16*(1), 30-42. doi: 10.1044/1058-0360(2007/005).

Tonietto, L., Villavicencio, A., Siqueira, M., Alice, M., Pimenta, M., Tania, P., et al. (2008). A especificidade semântica como fator determinante na aquisição de verbos. *PSICO*, *39*, 343-351.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, *393*, 440-442.

Sneezed, stretched, and vanished: Differential brain activations of different classes of intransitive verbs

Einat Shetreet
Harvard University, Tel Aviv University
einat.shetreet@childrens.harvard.edu

Naama Friedmann
Tel Aviv University
naamafr@post.tau.ac.il

Uri Hadar
Tel Aviv University
Urih1@post.tau.ac.il

Abstract

This study uses fMRI to test the neural distinctions between sub-classes of intransitive verbs: unergatives, unaccusatives, and reflexives.

1 Introduction

According to syntactic theories, the verb assigns to its arguments thematic roles that specify the mode of their participation in the event (Dowty, 1989; Gruber, 1965; Jackendoff, 1972). The *agent* role, which indicates the entity that performs the action or brings about some change, i.e., the doer or initiator, is strongly associated with the subject position, whereas the *theme* or the *patient* role, which indicates the entity that the action is performed upon, i.e., the undergoer, is strongly associated with the object position. For example, in the sentence *The girl sneezed*, *the girl* in the subject position is the *agent*. However, *the girl* in the sentence *The girl vanished* is not actively responsible for the action described by the verb, and accordingly, is assigned the *theme* role. Unaccusative verbs are assumed to be lexically derived by an operation that reduces an argument from transitive verbs (Chierchia, 1989; Reinhart & Siloni, 2005). This operation eliminates the subject of the transitive verb or, in terms of thematic roles, eliminates its *agent* (Reinhart & Siloni, 2005), leaving the direct object (or the *theme*) the sole argument. In languages such as English, the NP must move to the subject position. To create sentences of the order NP-V for sentences with unaccusative verbs, the noun moves from its original position after the verb to the subject position (Burzio, 1986; Levin & Rappaport-Hovav, 1995; Perlmutter, 1978). In Hebrew, both the order noun-verb (NP-V) and the order verb-noun (NP-V) are acceptable.

Like unaccusative verbs, reflexive verbs, which denote an action that the *agent* applies on itself (*e.g.*, *The girl stretched*), are derived from transitive verbs. Note that in Hebrew reflexives have a distinctive morphological structure that distinguishes them from the transitive verbs from which they were derived. Thus, although in English the verb *stretched* can be used for both transitive and reflexive instances, in Hebrew the two verbs are distinct. The way reflexive verbs are derived from transitives is still debated. Some accounts argue that this derivation includes a lexical operation similar to the one that derives unaccusative verbs from transitive verbs, and thus include movement from object to subject position (Kayne 1988; Marantz, 1984; Pesetsky 1995). Other accounts claim that the argument of reflexive verbs, like the argument of unergative verbs, originates in the subject position. According to these accounts, reflexive verbs are the output of a lexical operation of absorption or reduction, which applies to a transitive entry, targeting its internal argument and producing an intransitive verb (Chierchia, 1989; Grimshaw, 1982; Reinhart & Siloni, 2004). Therefore, these approaches differ with respect to whether or not the derivation of reflexive verbs includes syntactic movement (leading to word order change) or not.

Here, we report two fMRI experiments that examined the patterns of cortical activation associated with the comprehension of unaccusative and reflexive verbs. This study specifically aims to examine whether, as predicted by linguistic theory, the cortical representation of unaccusatives, reflexives, and unergatives differs. Additionally, we used patterns of cortical activation in an attempt to adduce evidence that may help to decide the nature of the derivation of reflexive verbs.

2 General Methods

The first stage of the experimental procedure included the selection of Hebrew verbs based on several distinguishing criteria including: (1) the possibility to add possessive datives, (2) the possibility to appear in V-NP order (Siloni, 2008), and (3) the possibility to occur with reflexive pronouns. In the experiments, each verb was embedded in four sentences. The sentences in each experiment were controlled for the number of phrases, phrase structure, definiteness, and duration. The verbs were controlled for verb templates and frequency. A block design paradigm was used. Each block included four sentences and each condition repeated 7 or 8 times. Eighteen (Experiment 1) or twenty-four (Experiment 2) participants were asked to listen to the sentences and to decide whether the event described in the sentence is more likely to happen at home or not (for example, for a sentence like "Dan slept in the yellow tent", participants will press a "no" button). This semantic task ensured that participants attended to the sentences and processed them fully.

3 Experiment 1: The distinction between unaccusative and unergative verbs

In this experiment (Shetreet et al., 2010), we compared unaccusative verbs (in NP-V order) to unergative verbs (with one argument) and transitive verbs (with two arguments). That is, we compare verbs that undergo lexical reduction and syntactic movement to verbs that do not. Other types of syntactic movement have shown activations in the left inferior frontal gyrus (IFG) (e.g., Ben-Shachar et al., 2003, 2004; Stromswold et al., 1996). Lexical related activation was previously revealed in several brain areas including left frontal, left temporal and left inferior parietal cortices (e.g., Fiebach et al., 2002; Keller et al., 2001; Kemeny et al., 2006; Kircher et al., 2000). A direct comparison between unaccusative and unergative verbs revealed activations in the left IFG (Brodmann area (BA) 45/46/47), left superior frontal gyrus, left middle temporal gyrus (MTG; BA 21) and right cerebellum. This indicates the brain distinguishes between unaccusative and unergative verbs, even when they appear in identical sentence structures (i.e., "*The girl vanished*" vs. "*The girl sneezed*"). These results join a growing body of findings from various methodologies and populations, suggesting evidence for the neuropsychological and psycholinguistic reality of this distinction generally, and for the analysis of unaccusative verbs specifically (Friedmann, 2007; Friedmann et al., 2008). In order to identify the activations that relate to the lexical operation and the syntactic movement of unaccusatives, and distinguish them from activations linked to a specific difference between unaccusatives and unergatives, we performed a conjunction analysis with the comparison between unaccusatives and transitive verbs. This analysis showed activations in the left IFG (BA 45/46) and the left posterior MTG. The present study cannot conclusively determine which operation is associated with each of these areas. However, previous findings regarding the functions of these areas can give us some insights for their role in the comprehension of unaccusative verbs. The left IFG has been consistently linked to syntactic processing and syntactic movement (Ben-Shachar et al., 2004; Friedmann, 2006; Grodzinsky, 2000; Shetreet et al., 2009; Stromswold et al., 1996; Zurif, 1995), and thus its activation in our experiment may be related to the movement of the object to the subject position. The left posterior temporal areas have been associated with lexical and semantic processes and with verb processing (Demonet et al., 1992; Friederici et al., 2000; Humphries et al., 2006; Palti et al., 2007; Perani et al., 1999; Price et al., 1997). An adjunct area (left fusiform), located medially to the area identified in this experiment, was also implicated in the processing of an operation that omits complements of optional complements (Shetreet et al., 2009b). This may suggest that left MTG activation with response to unaccusatives is linked to the lexical operation.

4 Experiment 2: The distinction between reflexive and unaccusative verbs

This experiment was aimed to inform the linguistic controversy regarding the derivation of reflexive verbs, and to determine whether the lexical operation that derives reflexive verbs involves the reduction of the external argument of a transitive verb or a reduction of the internal argument. In an attempt to answer this question, we compared reflexive verbs and unergative verbs. Additionally, we compared reflexives with unaccusatives, which undergo lexical reduction, as well as syntactic movement. As in Experiment 1, comparing Unaccusatives and unergatives resulted in activations in the left IFG (BA 45/46/47) and in the left posterior MTG (as well as other activations). Comparing reflexives and unergative did not reveal these areas, but instead the right MTG and the right middle frontal gyrus (MFG). Both of these areas have been linked to syntactic binding (Grodzinsky & Friederici, 2006) and was found in the binding of a noun with its reflexive pronoun (Santi & Grodzinsky, 2007a, 2007b).

To further examine the differences between the verb classes, we defined the areas identified in the unaccusative-unergative comparison, the left IFG and the left MTG, as regions of interest (ROI). For each ROI, we computed the average beta weights and compared them using ANOVA and Tukey test. In the left IFG, activation in response to unaccusative verbs was higher than for both unergatives and reflexives. By contrast, in the left MTG, unaccusative differed significantly only from unergative verbs. It seems that the activations associated with reflexive verbs are a subset of the activations associated with unaccusative verbs (NP-V order). Specifically, this subset seems to overlap with the activation attributed to the lexical operation in the derivation of unaccusative verbs. Thus, our results support the claims that the lexical operation is involved in the derivation of reflexive verbs targets the internal argument and not the external argument. This is suggested by the absence of activation in the processing of reflexives in the left IFG, which is linked to the processing of movement of the object to subject position in unaccusative verbs.

5 Conclusion

Our findings indicate that brain activation is sensitive to the difference between sub-classes of intransitive verbs: unaccusatives, reflexives, and unergatives, distinguishing between verbs that undergo lexical and syntactic operations and those that do not. Specifically, we found support for the involvement of lexical and syntactic operations in the processing of unaccusative verbs (in NP-V order) and for the involvement of a lexical operation in the processing of reflexive verbs. The latter result supports the linguistic analysis according to which reflexive verbs are derived from transitive verbs by reducing the *internal* argument.

References

Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., & Grodzinsky, Y. (2003). The neural reality of syntactic transformations: Evidence from fMRI. *Psychological Science*, *14*, 433-440.

Ben-Shachar, M., Palti, D., & Grodzinsky, Y. (2004). Neural correlates of syntactic movement: Converging evidence from two fMRI experiments. *NeuroImage*, *21*, 1320-1336.

Burzio, L. (1986). *Italian syntax. A Government-Binding approach*. Dordrecht: Reidel.

Chierchia, G. (1989). *A semantics for unaccusatives and its syntactic consequences*. Ms., Cornell University.

Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., et al. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, *115*, 1753-1768.

Dowty, D. (1989). On the semantic content of the notion of 'thematic role'. In G. Chierchia, B. Partee, & R. Turner (Eds.), *Properties, types, and meanings, vol. 2* (pp. 69-129). Dordrecht: Kluwer.

Fiebach, C., Friederici, A. D., Mueller, K., & von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience,*14, 11-23.

Friederici, A., Optiz, B., & von Cramon, D. (2000). Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types. *Cerebral Cortex*, *10*, 698-705.

Friedmann, N. (2006). Speech production in Broca's agrammatic aphasia: Syntactic tree pruning. In Y. Grodzinsky & K. Amunts (Eds.), *Broca's Region* (pp. 63-82). New York: Oxford University Press.

Friedmann, N. (2007). Young children and A-chains: The acquisition of Hebrew unaccusatives. *Language Acquisition*, *14*, 377-422.

Friedmann, N., Taranto, G., Shapiro, L. P., & Swinney, D. (2008). The vase fell (the vase): The online processing of unaccusatives. *Linguistic Inquiry*, *39*, 355-377.

Grimshaw, J. (1982). On the lexical representation of Romance reflexive clitics. In J. Bresnan (Ed.), *The mental representation of grammatical relations*. Cambridge, MA: MIT Press. Grodzinsky, Y. (2000). Syntax in the brain: Linguistic versus neuroanatomical specificity. *Behavioral and Brain Sciences*, *23*, 1-21.

Grodzinsky, Y., & Friederici, A. D. (2006). Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology*, *16*, 240-246.

Gruber, J. (1965). *Studies in Lexical Relations*. Doctoral dissertation, MIT.

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, *18*, 665-679.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Kayne, R. (1988). Romance se/si. Presented at the GLOW 11 Colloquium, Budapest, *GLOW newsletter*, *20*, 33.

Kemeny, S., Xu, J., Park, G. H., Hosey, L. A., Wettig, C. M., & Braun, A. R. (2006). Temporal dissociation of early lexical access and articulation using a delayed naming task: An FMRI study. *Cerebal Cortex,16*, 587-595.

Keller, T., Carpenter, P., & Just, M. A. (2001). The neural bases of sentence comprehension: an fMRI examination of syntactic and lexical processing. *Cerebral Cortex*, *11*, 223-237

Kircher, T. T. J., Brammer, M. J., Williams, S. C. R. & McGuire, P. K. (2000). Lexical retrieval during ¯uent speech production: An fmri study. *Neuroreport*, *11*, 4093-4096.

Levin, B. A., & Rappaport-Hovav, M. (1995). *Unaccusativity*. Cambridge, Mass: MIT Press.

Marantz, A. (1984). *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

Palti, D., Ben-Shachar, M., Hendler, T., & Hadar, U. (2007). Neural correlates of semantic and morphological processing of Hebrew nouns and verbs. *Human Brain Mapping*, *28*, 303-314.

Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., et al. (1999). The neural correlates of verb and noun processing: A PET study. *Brain*, *122*, 2337-2344.

Perlmutter, D. M. (1978). Impersonal passives and the unaccusative hypothesis. In *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistic Society* (pp. 159-189). Berkeley, CA.

Pesetsky, D. M. (1995). *Zero syntax: Experiencer and cascades*. Cambridge, MA: MIT Press.

Price, C. J., Moore, C. J., Humphreys, G. W., & Wise, R. S. J. (1997). Segregating semantic from phonological processes during reading. *Journal of Cognitive Neuroscience*, *9*, 727-733.

Reinhart, T., & Siloni, T. (2004). Against the unaccusative analysis of reflexives. In A. Alexiadou, E. Anagnostopoulou, & M. Everaert (Eds.), *The unaccusativity puzzle* (pp. 159-180). Oxford: Oxford University Press.

Reinhart, T., & Siloni T. (2005). The lexicon-syntax parameter: Reflexivization and other arity operations. *Linguistic Inquiry*, *36*, 389-436.

Santi, A., & Grodzinsky, Y. (2007a). Taxing working memory with syntax: Bi-hemispheric modulations. *Human Brain Mapping*, *28*, 1089-1097.

Santi, A., & Grodzinsky, Y. (2007b). Working memory and syntax interact in Broca's area. *NeuroImage*, *37*, 8-17.

Shetreet, E., Friedmann N., & Hadar U. (2009). An fMRI study of syntactic layers: Sentential and lexical aspects of embedding. *NeuroImage*, *48*, 707-716.

Siloni, T. (2008). Al Binyan Hitpa'el. In G. Hatav (Ed.), *Balshanut Ivrit Te'oretit* (pp. 111-138). Jerusalem: The Hebrew University Magnes Press. (In Hebrew)

Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language*, *52*, 452-473.

Zurif, E. B. (1995). Brain regions of relevance to syntactic processing. In L. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science* (pp. 381-398). Cambridge, MA: MIT Press.

Verbs, Objects, and Events: Eye-Tracking Reveals the Time-Course of Aspectual Interpretation

David J. Townsend
Montclair State University
Upper Montclair, NJ, US
townsendd@mail.montclair.edu

George Olekson
Montclair State University
Upper Montclair, NJ, US
oleksong1@mail.montclair.edu

Abstract

We evaluated specification and under-specification hypotheses of verb representation by measuring eye-fixations during reading. Participants read two sets of verb phrases that differed in object definiteness and in preference for *in*- vs. *for*-adverbial phrases. Total fixation time on the adverbial phrase depended on a predicted interaction between verb phrase preference and object definiteness. These results are consistent with claims that verbs specify boundedness information.

1 Introduction

Sentences represent events as bounded or not. For example, we interpret *John built a house* to refer to a single completed event (bounded). However, we interpret *John built houses* to refer to an indefinite series of separate house-building events (unbounded). The difference in interpretation appears when we modify these sentences with time-span vs. durative adverbial phrases (Comrie, 1976; Dowty, 1979; Smith, 1991; Vendler, 1957; and many others). A bounded interpretation occurs in *John built a house in six months* but not in *John built houses in six months.* An unbounded interpretation occurs in *John built houses for six months* but not in *John built a house for six months*. We find different patterns in *John pushed a cart* and *John pushed carts*: Both are more acceptable with *for ten minutes* rather than *in ten minutes*. These observations suggest that the mental representations of *build* and *push* differ in boundedness.

Recent discussions suggest two general approaches to the representation of boundedness. Specification maintains that the mental representation of a verb contains information about boundedness. Various theories have expressed this idea in different ways. Type theories place verbs into distinct categories according to aspectual classes such as states, processes, transitions, etc. (e.g., Jackendoff, 1997; Pustejovsky, 1991). Information that conflicts with the aspectual type of a verb introduces semantic content or operators that convert it to another type. Other examples of specification maintain that verbs are marked for boundedness in various ways (Declerck, 1979; Verkuyl, 1989). Still other approaches propose that a verb such as *build* represents boundedness through its entailment of an incremental theme whose status is related to the completeness of an event (Dowty, 1991). In contrast, an unbounded verb such as *push* entails no incremental theme. Despite substantial differences among these theories, they have one property in common: Each specifies boundedness information in verbs.

Under-specification maintains that verbs are under-specified with regard to temporal boundedness. This approach emphasizes that aspectual interpretation requires combining phrases and other elements from the entire sentence (Pickering et al., 2006; Pylkkanen & McElree, 2006).

Two observations support under-specification. First, multiple interpretations often are possible (e.g., Declerck, 1979; Dowty, 1979). For example, we may assign a habitual (unbounded) interpretation to *John built houses in six months* such that John made a practice of building a house in six months and did so on several occasions. We may interpret *John built a house for six months* by shifting the interpretation of *built* to that of *worked on*, producing an unbounded interpretation. *John pushed a cart in ten minutes* has an inchoative interpretation of the time span that elapsed before the onset of the event. *John pushed carts in ten minutes* has a habitual interpretation of this inchoative meaning. This flexibility of interpretation seems inconsistent with specification of boundedness in verbs.

Second, many factors influence interpretations of boundedness. As noted above, an adverbial phrase with *in* vs. *for* influences these interpretations, as does the definiteness of the direct object. In fact, the definiteness of nearly any noun phrase in the sentence affects aspectual interpretation (Declerck, 1979; Verkuyl, 1993). For example, *John built a house in Pisa* is bounded, *John built a house in many cities* is unbounded. *A liter of water ran out of the tap* is bounded, *Water ran out of the tap* is unbounded (Declerck, 1979). *Den Uyl gave a badge to a congress-goer* is bounded, *Den Uyl gave a badge to congress-goers* is not (Verkuyl, 1993). These facts suggest that aspectual interpretation depends greatly on the context in which a verb appears.

Questions about the nature of aspectual representation in verbs appear in the literature on sentence processing. One study supports different representations of stative vs. eventive verbs (i.e., activities, accomplishments, and achievements) (Gennari & Poeppel, 2003). Other studies suggest that event semantics has a role in comprehension, but they provide few details about the sequence of interpretive processes during uninterrupted reading or listening (Brennan & Pylkkanen, 2008; Husband et al., 2008; Pinango et al., 1999; Proctor et al., 2004; Todorova et al., 2000; Townsend & Seegmiller, 2004). A third group of studies used analyses of eye-tracking to argue for representation of individuals and events as distinct types (Pickering et al., 2006; Pylkkanen & McElree; Traxler et al., 2002). Results from this group of studies suggest that verbs are underspecified for boundedness.

We evaluated the specification and under-specification approaches in sentence comprehension. To remain theoretically neutral, we defined boundedness in terms of participants' judgments about the acceptability of sentences with *in*- vs. *for*-modification. We used a forced choice test in which we presented pairs of sentences that differed only in *in/for*:

(1) A. *The curious cat killed the grey mouse in 8 minutes*.
B. *The curious cat killed the grey mouse for 8 minutes*.

(2) A. *The black bear hunted the crimson fox in two hours.*
B. *The black bear hunted the crimson fox for two hours.*

One sentence in each pair contained an *in*-phrase that specifies a time span for completing a bounded event. The other contained a *for*-phrase that specifies the duration of an unbounded event. In order to increase naturalness, we placed each verb in a unique context, as in (1) and (2).

We asked participants to make one of four choices about the sentences within pairs such as (1) and (2): The sentence with *in* sounds better, the sentence with *for* sounds better, both sound good, or neither sounds good. We defined a bounded verb as one that participants judged more acceptable with an *in*-phrase rather than a *for*-phrase. An unbounded verb is one that they judged more acceptable with a *for*-phrase.

To examine the mechanism of aspectual interpretation, we presented the same sentences to other participants in an eye-tracking experiment. The variables were Verb (bounded vs. unbounded), Object (definite vs. indefinite), and Adverb (*in* vs. *for*). We ended the sentence after the adverbial phrase to increase the effect of sentence "wrap-up" processes. Examples appear in Table 1.

<u>Bounded Verb</u>:
Definite, *in: The curious cat/ killed/ the grey mouse/ in eight minutes*.
Definite, *for: The curious cat/ killed/ the grey mouse/ for eight minutes.*
Indefinite, *in: The curious cat/ killed/ grey mice/ in eight minutes.*
Indefinite, *for: The curious cat/ killed/ grey mice/ for eight minutes.*
<u>Unbounded Verb</u>:
Definite, *in: The black bear/ hunted/ the crimson fox/ in two hours.*
Definite, *for: The black bear/ hunted/ the crimson fox/ for two hours.*
Indefinite, *in: The black bear/ hunted/ crimson foxes/ in two hours.*
Indefinite, *for: The black bear/ hunted/ crimson foxes/for two hours.*

Table 1: Conditions and Scoring Regions for Two Items

We obtained two eye-tracking measures on the adverbial phrase (Liversedge, Paterson, & Pickering, 1998; Kennedy & Murray, 1987; Rayner, 1998). *First pass time* is the sum of fixation times from the first fixation in the adverbial phrase through the last until the eye leaves the adverbial phrase. First pass time assesses attempts to resolve processing difficulties without leaving the adverbial phrase. *Total time* is the sum of all fixation times in the adverbial phrase including fixations during the first pass through the phrase and fixations in the phrase following any regressions.

Specification theories predict an interaction between Verb and Object: An indefinite object phrase increases fixation time following a bounded verb, but not following an unbounded verb. Under-specification maintains that aspectual interpretation depends on integrating all parts of a sentence. Since aspectual interpretation depends on the meanings of phrases rather than on boundedness information in the verb, there is no reason to expect difficulty with any combination of phrases.

2 Method

2.1 Participants

Forty participants came from the Psychology Department subject pool at Montclair State University. All participants were native speakers of English with vision that was normal or corrected to normal. Participants received either course credit or payment for the 40 minute experiment.

2.2 Materials

The materials consisted of 14 bounded verbs and 16 unbounded verbs. Each verb appeared in four conditions depending on Adverb and Object (see Table 1). The direct object was either definite (singular) or indefinite (bare plural); three direct objects had irregular plural forms (e.g., *mice*). The adverbial phrase specified either a time span (*in*) or a duration (*for*). Neither the number of characters in the adverbial phrase nor its Collins Cobuild frequency differed for bounded vs. unbounded verbs, both ps > .25.

We conducted two surveys to evaluate judgments about the materials. For both surveys we adapted the forced-choice test developed by Townsend & Seegmiller (2004). Thirty-two college students who did not participate in eye-tracking took both surveys.

We first established that the two groups of verbs produce different interpretations. We presented pairs of sentences that differed only in *in/for*, as in (1) and (2). One sentence contained an *in*-phrase that specifies a time span for completing a bounded event. The other contained a *for*-phrase that specifies the duration of an unbounded event. Participants indicated whether a sentence containing a bounded or unbounded verb sounds better with *in* or *for*, whether both *in*- and *for*-sentences are acceptable, or whether neither sentence is acceptable. The results showed that participants prefer bounded verbs with *in*-phrases (75% preferred a bounded verb with *in* while 5% preferred a bounded verb with *for).* They prefer unbounded verbs with *for*-phrases (74% preferred an unbounded verb with *for* while 7% preferred an unbounded verb with *in*). The two groups of verbs differed in preference for both *in*, $F(1, 28) = 165$, $p < .001$, and *for*, $F(1, 28) = 304$, $p < .001$. The verb groups did not differ in choices of "both" (14 vs. 17% for bounded and unbounded verbs respectively) or "neither" responses (6 vs. 3% respectively), both *p*s > .25. Thus, sentences with verbs from different groups differ in interpretation but not in acceptability.

A second survey examined transitivity preferences for bounded vs. unbounded verbs. Participants received a pair of questions such as "What did the cat kill?" vs. "When did the cat kill?" They indicated whether "What…" sounds better (indicating a preference for transitive), "When…" sounds better (indicating a preference for intransitive), both are acceptable, or neither is acceptable. Verb group was unrelated to the percentage of choices of transitive questions (43 vs. 45% for bounded vs. unbounded respectively), intransitive questions (21 vs. 16%), both (33 vs. 35%), or neither (3 vs. 4%), all *F*s < 1.

2.3 Procedure

We conducted the experiment with an SR Research Eye Link 1000 desktop system, and Eye Tracking and data processing software from http://www.umass.edu/eyelab/software/. The monitor was 50 cm from the participant. Participants rested their chin and forehead on bars. The system was calibrated for right eye tracking with corneal reflection. Maximum drift error was set at 0.4 degrees and checked several times during each session. A trial began when the participant focused on a circle in the center of the screen. When the participant's gaze was stable, the Eye Track software presented a square near the left edge of the screen. When the participant looked at the square, the software displayed the sentence. The font was Arial 18. The screen width was 160 characters with a resolution of 1280 by 1024. Participants were instructed to read each sentence normally. When they reached the end of a sentence, participants looked at a sequence of XXX one line below and 5 spaces to the right of the period. They then pressed a button on the left side of a game controller. This button press either initiated another trial or presented a question. Participants answered questions by pressing a button on the right or left side of the game controller.

Each participant read 128 sentences. Thirty sentences had the form of those in Table 1. Four lists had 7-8 sentences in each of four conditions: 2 Adverb (*in* vs. *for*) x 2 Object (definite vs. indefinite). Of the remaining 98 filler sentences, 24 had clauses conjoined with *and*, 48 had clauses conjoined with a subordinating adverbial conjunction, and 26 were a mixture of one- and two-clause sentences. A question followed 48% of both filler and test sentences. Half of these questions concerned agent and patient roles (e.g., *Who did the hunting? Fox vs. Bear*); half concerned the number of events (e.g., *How many killings were there*? *Just one vs. More than one*).

Data processing software adjusted vertical displacement. The software combined fixations that were shorter than 80 ms. It excluded trials in which gaze duration exceeded 2000 ms or no fixation occurred in the region. First pass data in the adverbial phrase was missing on 1.9% of the trials.

We evaluated the statistical significance of differences in first pass time and total time for the adverbial phrase with analysis of variance by participants and by items. The variables in these analyses were Verb (bounded vs. unbounded), Object (definite vs. indefinite), and Adverb (*in* vs. *for*). We used residual reading times to factor out the effects of variability in length of the adverbial phrase. The residual reading time in a region is the difference between the actual fixation time and the fixation time that linear regression predicts from the number of characters in the region (Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). The basis for these predictions was fixation time in the subject phrase; the verb; the direct object phrase; the adverbial phrase. Table 1 marks these regions with a /.

The results of interest are the three-way interaction between Verb, Object, and Adverb, and the two-way interaction between Verb and Object. Specification theories predict that fixation time is longer for indefinite objects than definite objects only following bounded verbs.

3 Results

Table 2 shows mean first pass time and total time in the adverbial phrase region.

First pass time showed a small effect of object definiteness when the adverb was *in* and the verb was bounded: Fixation time was longer for indefinite objects than for definite objects (655 vs. 631 ms). The opposite occurred when the adverb was *in* and the verb was unbounded (633 vs. 670 ms), and when the adverb was *for* regardless of verb (bounded verbs: 668 vs. 724 ms for indefinite and definite objects respectively; unbounded verbs: 639 vs. 656 ms). However, neither the three-way interaction in residual reading time nor the interaction between Verb and Object was significant, all ps > .10.

Total time showed a similar pattern. When the adverb was *in* and the verb was bounded, fixation time was longer for indefinite objects than for definite objects (843 vs. 754 ms). The opposite occurred when the adverb was *in* and the verb was unbounded (803 vs. 854 ms), and when the adverb was *for* regardless of verb (bounded verbs: 853 vs. 887 ms for indefinite vs. definite objects respectively; unbounded verbs: 750 vs. 861 ms). Again, the three-way interaction in residual reading time was not significant, both ps > .10. However, fixation time overall was longer for indefinite than for definite objects following bounded verbs (848 vs. 821 ms) but not following unbounded verbs (777 vs. 858 ms), $F1\ (1, 39) = 6.38$, $p < .05$, $F2\ (1, 28) = 6.58$, $p < .05$.

		In				*For*			
		Bounded Verb		Unbounded Verb		Bounded Verb		Unbounded Verb	
		Definite	Indefinite	Definite	Indefinite	Definite	Indefinite	Definite	Indefinite
FPT	M	631	655	670	633	724	668	656	639
	s.e.	41.3	32.5	32.0	34.0	42.3	43.0	33.9	33.9
	RRT	28.4	51.6	35.5	-8.1	92.1	33.4	-3.1	-26.5
TT	M	754	843	854	803	887	853	861	750
	s.e.	46.6	54.7	42.4	42.7	53.3	51.6	53.7	46.5
	RRT	-17.5	71.1	43.2	-13.3	83.9	49.0	12.4	-97.7

Table 2: Mean Fixation Time (ms) in the Adverbial Phrase Depending on Verb, Object, and Adverb
Note. FPT = first pass time; TT = total time; M = mean; s.e. = standard error; RRT = residual reading time.

4 Discussion

We considered two hypotheses about the role of the verb in aspectual interpretation. The specification hypothesis maintains that boundedness information appears in the representation of verbs. This hypothesis proposes that the processor adopts an aspectual interpretation when it recognizes a verb. If subsequent information conflicts with this interpretation, the processor shifts its interpretation to agree with the (conflicting) new information. The specification hypothesis predicts that verb semantics interacts with conflicting information during sentence processing.

The under-specification hypothesis maintains that the mental representations of verbs do not contain information about boundedness. The processor forms an aspectual interpretation by integrating non-aspectual meanings at the end of the sentence. The under-specification hypothesis predicts no interactions between verb, object and adverbial phrase.

The present data support the view that boundedness information appears in the representation of verbs. When the verb was bounded, total time on the adverbial phrase was longer for indefinite objects than for definite objects. When the verb was unbounded, this difference did not occur. A similar but non-significant trend appeared in first pass time. These results support the view that aspectual interpretation occurs during sentence processing.

Our design suggests caution in concluding that recognition of a verb immediately establishes an interpretation of boundedness. The optimal comparisons for evaluating this claim is first pass time on definite vs. indefinite objects or on the following region. To increase the naturalness of our materials, we allowed the content of object phrases to vary across verbs. This feature of our design prohibits conclusive comparisons of fixation time on object phrases. In addition, our data showed only non-significant effects in first pass time on the following adverbial phrase. Although we cannot conclude that the processor adopts an aspectual interpretation at the moment of recognizing a verb, we can conclude that it does so at the time of processing the adverbial phrase. The appearance of significant effects in total time on the adverbial phrase suggests that the processor initiates aspectual re-interpretation on the adverbial phrase, and continues it in regressions and in re-reading the adverbial phrase. Thus, our data confirm that representations of verbs contain boundedness information and that the sentence processor uses this information during comprehension.

Acknowledgments

This research was supported by NIH 1R15HD055680-01A1. We are grateful to Chuck Clifton, Steve Seegmiller, and Mary Call for advice on the manuscript, to three anonymous reviewers for their valuable comments, and to the members of the Eye Tracking Lab at Montclair State University for assistance in testing participants.

References

Brennan, J., & Pylkkanen, L. (2008). Processing events: Behavioral and neuromagnetic correlates of aspectual coercion. *Brain & Language, 106*, 132-143.

Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems.* Cambridge: Cambridge University Press.

Declerck, R. (1979). Aspect and the bounded/unbounded (telic/atelic) distinction. *Linguistics, 17*, 761-794.

Dowty, D. R. (1991). Thematic proto-roles and argument selection*. Language, 67*, 547–619.

Dowty, D. R. (1979). *Word meaning and Montague grammar.* Dordrecht: Reidel.

Ferreira, F. & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*, 348-368.

Gennari, S., & Poeppel, D. (2003). Processing correlates of lexical semantic complexity. *Cognition, 89*, B27-B41.

Husband, E. M., Stockall, L., & Beretta, A. (2008). VP-Internal event composition. Unpublished manuscript, Michigan State University. World wide web: http://eyelab.msu.edu/people/matt/vp-internal_final_submitted.pdf (July 1, 2008).

Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.

Kennedy, A., & Murray, W. W. (1987). The components of reading time: Eye movement patterns of good and poor readers. In J. K. O'Regan & A. Levy-Schoen (Eds.). *Eye movements: From physiology to cognition.* (pp. 509-520). New York: Elsevier Science.

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.). *Eye guidance in reading and scene perception*. (pp. 55-75). Amsterdam: Elsevier Science Ltd.

Pickering, M. J., McElree, B., Frisson, S., Chen, L., & Traxler, M. J. (2006). Underspecification and aspectual coercion. *Discourse Processes, 42*, 131-155.

Pinango, M. M., Zurif, E., & Jackendoff, R. (1999). Real-time processing implications of enriched composition at the syntax-semantics interface. *Journal of Psycholinguistic Research*, 28, 395-414.

Proctor, A. S., Dickey, M. W., & Rips, L. (2004). The time-course and cost of boundedness inferences. *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, Chicago.

Pustejovsky, J. (1991). The syntax of event structure. *Cognition, 41*, 47-81.

Pylkkanen, L., & McElree, B. (2006). The syntax -semantics interface: On-line composition of sentence meaning. In M. J. Traxler & M. A. Gernsbacher (Eds.). *Handbook of psycholinguistics: 2nd Edition.* (pp. 539-579). Amsterdam: Elsevier.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Smith, C. S. (1991). *The parameter of aspect.* Dordrecht: Kluwer Academic Publishers.

Todorova, M., Straub, K., Badecker, W., & Frank, R. (2000). Aspectual coercion and the online computation of sentential aspect. *Proceedings of the annual conference of the Cognitive Science Society, 22*, 523-528.

Townsend, D. J., & Seegmiller, M. S. (2004). The linguistic representation and processing of event structure. *Journal of Cognitive Science, 5*, 157-244.

Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language, 47*, 530-547.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285-318.

Vendler, Z. (1957). Verbs and times. *Philosophical Review, 66*: 143-160.

Verkuyl, H. J. (1989). Aspectual classes and aspectual composition. *Linguistics and Philosophy, 12*: 39-94.

"I like work: I can sit and look at it for hours" Type clash vs. plausibility in covert event recovery

Alessandra Zarcone, Sebastian Padó
University of Stuttgart
Stuttgart, Germany
zarconaa,pado@ims.uni-stuttgart.de

Abstract

A range of event-subcategorizing verbs can combine with entity-denoting nouns, like *begin the newspaper*. The interpretation of such sentences typically involves the recovery of covert events (CE) which are not realized on the surface, as in *begin reading the newspaper*. We report on an ongoing study that scrutinizes two assumptions made by traditional accounts: (a) that the triggering of CEs can be ascribed to the object's ontological type; and (b), that one or two CEs can be retrieved for each noun. Preliminary evidence against both assumptions is presented.

1 Covert Events

There is a substantial class of more than a dozen verbs whose members have been argued to subcategorize for an event (Pustejovsky, 1995; Jackendoff, 1997), but which can also combine with an entity. This class comprises a number of high-frequency verbs, such *enjoy* or *begin*. These verbs do not pose problems when combined with event-denoting objects (EV, e.g. *the afternoon*), but when combined with entity-denoting objects (EN, e.g. *the newspaper*) they constitute a challenge for traditional compositional accounts of sentence meaning, because their interpretation seems to require the recovery of covert events (CE) which are not realized on the surface (*begin the newspaper* → *begin reading the newspaper*). The interpretation of such pairs seems to involve at least two specification steps: (1) the triggering of (the need for) a CE; (2) the recovery of a specific CE.

The main determinant of step (1) has been argued to be the ontological type of the object (EN vs. EV objects) and its type-clash with the event-subcategorizing verb (Pustejovsky, 1995; Jackendoff, 1997; Traxler et al., 2002). Step (2) is traditionally assumed to result in one or at most two CEs retrieved from the *qualia structure* (QS) of the lexical entry for the object (Pustejovsky, 1995).

Behavioral studies have grounded this binary distinction in higher processing costs for conditions that involve CE recovery (see Pylkkänen and McElree (2006) for a review). Traxler et al. (2002) compared EN conditions ("began the book") with EV conditions ("began the fight"), using both eye-tracking and self-paced reading, and detected higher processing costs for EN objects with event-subcategorizing verbs both at the target position (the object itself) and at the post-target position.

2 Open Issues

The goal of our work is to scrutinize two assumptions of the traditional account: the nature of the "trigger" and the range of possible CEs.

The trigger problem. The following examples illustrate our intuition that a type clash between verb and object cannot be the only factor responsible for evoking CEs:

(1) I like work: it fascinates me. I can sit and look at it for hours.[1]

(2) Mary began the translation → began the translation process (EV) OR began reading/revising/typing the translation (EN).

(3) a. John is a famous wrestler. He really enjoyed the fight last night.
b. John is a wrestling fan. He really enjoyed the fight last night. → enjoyed watching the fight.

The twist that turns (1) into a joke is exactly the interpretation of work as an event, which is nevertheless later modifier by the recovery of a CE inserted between the verb and the object. The second example introduces a whole category of cases which are problematic for an ontological trigger, namely sortally ambiguous nouns that can assume

[1] J. K. Jerome, *Three men in a boat*, 1889

Noun Type	Example			Interpretation	Paraphrase
EV	begin the **afternoon**	→	✓	begin(afternoon)	
		→	×	begin(**CE**(afternoon))	
EN	begin the **newspaper**	→	×	begin(newspaper)	
		→	✓	begin(**CE**(newspaper))	begin **reading** the newspaper
EN/EV	begin the **breakfast**	→	?	begin(breakfast)	
		→	?	begin(**CE**(breakfast))	begin **eating** the breakfast

Table 1: Interpretation of different noun types after event-subcategorizing verbs

both an EN and an EV reading (cf. Table 1). One possible prediction would be that if a reading without type clash (i.e., an EV reading) is available, it will be chosen. This prediction is contradicted by Example (3), which shows that preceding discourse context can determine the choice between EN and EV.

Evidence against the type clash hypothesis also comes from work on metonymy resolution (Markert and Hahn, 1997; Markert and Hahn, 2002), which rejects this hypothesis on the basis of computational and cognitive arguments, and from MEG studies (Pylkkänen and McElree, 2007; Pylkkänen et al., 2009), which showed different brain activity correlates for semantic anomaly and for CE constructions.

Corpus sentence	Interpretations
If you are going hungry, seek **help with food** right away	obtain, buy, get
One friend works in the kitchen, **helping with food**	prepare, cook
I need **help with** dog **food**	select, choose

Table 2: Examples of verb+EN noun pairs

The range problem. Another issue concerns the retrieved CEs. If we equate CEs with qualia roles, there should be one or two CEs associated with each noun. However, the examples in Table 2 indicate that a wider range of CEs might be available, as Vendler (1968) had also observed.

Also, as observed by Lapata and Lascarides (2003) and Shutova and Teufel (2009), CEs are to be considered not as single verb lexical items but rather as classes of events sharing semantic similarities: each entry in Table 2 can be interpreted with a set of synonymous verbs rather than with a single lexical item.

2.1 An alternative mechanism: Plausibility

The alternative hypothesis that we want to explore is that interpretation is basically *plausibility-driven*. This hypothesis is coherent with the results obtained by probabilistic models of logical metonymy (Lapata and Lascarides, 2003; Shutova and Teufel, 2009).

The trigger problem. Probabilistic models yielded interesting results in predicting CE interpretations, but they did not distinguish between contexts in which CE are retrieved and contexts in which they are not. In order to account for the trigger problem, we suggest that CEs are retrieved when the plausibility of the standard verb/noun combination is small compared to the plausibility of the verb/CE/noun combination[2].

The range problem. Assuming an important role of plausibility, there is also no reason why the range of CEs should be limited a priori; rather, the CE could be sampled from distributional knowledge about plausible predicate-argument structures (Padó et al., 2007); more than one or two clusters of meaning can be retrieved and ranked for their plausibility (Lapata and Lascarides, 2003).

Steps of interpretation. These are the operations that we assume to take place when a potentially metonymic construction v, o is processed, given a previous context c:

1. candidate retrieval: a number of CE interpretations ce are activated, showing high plausibilities $Plaus(v, ce, o|c)$;
2. CE triggering: $Plaus(v, e, o|c)$ for the selected interpretations are compared to $Plaus(v, o|c)$; if $Plaus(v, o|c)$ is high enough to warrant non-CE interpretation, then no CE is retrieved; if instead the most plausible interpretation involves a CE, then the CE interpretation is selected;
3. CE range: the most plausible CE interpretation for v, o given c is selected and the meaning of e is integrated into the sentence meaning.

[2]The plausibility of the verb/CE/noun combination (v, e, o) can be estimated as the joint probability of $P(v)$, $P(v|e)$ and $P(o|e)$ (Lapata and Lascarides, 2003).

We make four observations. (a) Traxler et al. (2005) and Frisson and McElree (2008) showed that higher processing costs in CE conditions are not due to the retrieval of the CE, but to the "building of an extended event sense of the complement", so the plausibility comparisons in step 2 alone do not lead to higher processing costs. (b) The model does not imply a strong rejection of the type-clash model, but rather its predictions capture "tendencies" of the model: EV nouns tend to show higher plausibilities for the verb/noun combination, EN nouns show an opposite tendency and therefore more often require the recovery of CEs. (c) The range in 3 can be wider or narrower depending on the skewedness of the distribution over covert events given the previous context c. (d) Less plausible interpretations can remain available, in case following context falsifies the selected interpretation.

Predictions from the model. EN/EV ambiguous nouns as objects provide a suitable test object for our hypothesis: with both readings available, we can test to what extent plausibility considerations can account for differences in reading times. We expect EN nouns to show longer reading times than EV nouns in metonymic contexts; as to EN/EV nouns, we expect their behavior to be highly lexically-determined and to correlate with plausibility estimations. We therefore plan a self-paced reading study involving EN/EV ambiguous nouns, which is described in Section 3.

As to the range problem, reading time studies cannot help us in regard to it, as the CEs do not form part of the information acquired from the subjects. Section 4 therefore presents web-based elicitation methods that serve both to select materials for the reading time study and to explore the correlation between speaker's categorization of objects into EN / EV and their CE interpretation.

3 A self-paced reading study

Our design mirrors the study in Traxler et al. (2002), with an additional level: together with EN and EV objects, we are going to analyze the interaction between entity-subcategorizing verbs and EN/EV ambiguous nouns. 10 triplets of EN - EV - EN/EV ambiguous nouns were selected. For each triplet, two verbs were chosen: an event-subcategorizing verb (*begin-verb*), and a verb which could categorize both for an event and an entity (*spot-verb*). See an example triplet:

EN: Keith enjoyed/approved the automobile on the premises of the company.

EV: Daniel enjoyed/approved the conference on the premises of the company.

EN/EV: Walter enjoyed/approved the translation on the premises of the company.

Objects were matched within each triplets for length, frequency (Francis and Kucera, 1967), and co-occurrence frequency with the begin-verb and the spot-verb (ukWaC corpus, Ferraresi et al. (2008)), as a rough indicator of plausibility (Lapata and Lascarides, 2003). The 10 triplets were selected after threefold annotation, to evaluate our annotation of the nouns as EN, EV or EN/EV. Non-weighted Krippendorff's α (Krippendorff, 2004) for the selected triplets was .71, or good agreement. We also computed the weighted version of α, which incorporates the idea that EN vs. EV is a stronger disagreement compared to either of the types vs. the ambigous EN/EV type.[3].

Weighted α is =.79 – that is, determining the appropriate reading is not trivial, but doable.

4 Web experiments

The experiments were delivered using the crowdsourcing paradigm (Snow et al., 2008), for fast and affordable collection of judgments.

4.1 Experiment 1

In Experiment 1, 14 annotators from the US re-annotated the 30 nouns from the 10 triplets selected for the self-paced reading study for their readings (EN, EV, EN/EV). The aim of Experiment 1 was to check for non-expert annotation of the materials for the self-paced reading study, and to verify that this annotation did not change with different PP contexts.

Each noun appeared with a begin-verb and with a spot-verb and in three contexts: without the PP ("Keith enjoyed the automobile"), with the first part of the PP ("Keith enjoyed the automobile on the premises"), and with the complete sentence ("Keith enjoyed the automobile on the premises of the company"). We found a reasonably good agreement among annotators for a crowdsourcing experiment (weighted $\alpha = .52$)[4] and were able to rule out potential meaning changes caused by the

[3] We assigned a weight of 1 to EN-EV and a weight of 0.5 to EN-EN/EV and EV-EN/EV.

[4] Note that 14 annotations allow us to compute a reliable "majority vote" so that the practical reliability is higher.

post-nominal PPs: higher processing costs in the self-paced reading study will only be ascribed to CE recovery.

4.2 Experiment 2

It is not unusual for works on logical metonymy to include off-line norming studies, which can involve estimation of plausibilities for given CEs in a metonymical construction (Lapata and Lascarides, 2003) or the elicitation of a CE in a cloze completion task (McElree et al., 2001; Lapata et al., 2003). Nevertheless, the very same design of these experiments neglected the two aspects we are focusing on: cloze completion and plausibility estimation do not explore differences betweeb CE and no-CE interpretation (trigger problem) and limit the range of elicitations to only one CE (range problem). The aim of Experiment 2 is to evaluate the role of EN, EV, EN/EV nouns in triggering CE interpretations, to elicit more than one CE interpretation and to explore their range.

4.2.1 Experiment 2: materials and design

Experiment 2 was conducted with the same materials and procedure of Experiment 1, but this time participants were asked to choose between a CE interpretation and a simple compositional interpretation (*does the sentence involve an additional activity that is not mentioned in the sentence?*). Two options were given (*additional activity* vs. *no additional activity*), some examples are provided, and when a participant answered *additional activity*, she or he was asked to provide instances of possible activities. EN and EV interpretations were not mentioned in the experiment's instructions.

4.2.2 Experiment 2: results

The results from Experiment 2 involve two aspects 1) the CE/no-CE answer; 2) the elicited CEs.

CE/no-CE. Agreement for Experiment 2 was rather low ($\alpha = .35$)[5], but the majority vote showed a good agreement with the Gold Standard ($\alpha = .60$).

A binomial logistic regression on the CE/no-CE answers ($answer \sim obj\_type * verb\_type$) yielded a significant effect of the type of the object (binomial $p < 0.001$), and of the verb type ($z = -8.322; p < 0.001$), with interaction (binomial $p < 0.001$). These effects seem to confirm the type-clash hypothesis, but consider Table 3: 38% of begin-verb/EN-noun combinations did not elicit CEs, while 18% of begin-verb/EV-noun combinations did.

condition	% CE	% no-CE
begin,EN	0.63	0.38
spot,EN	0.11	0.89
begin,EN/EV	0.39	0.61
spot,EN/EV	0.06	0.94
begin,EV	0.18	0.82
spot,EV	0.06	0.94

Table 3: CE and no-CE answers in Experiment 2

condition	V-N pair	% CE	% no-CE
begin,EN	begin the newspaper	0.89	0.11
begin,EN/EV	begin the breakfast	0.81	0.19
begin,EN	enjoy the automobile	0.50	0.50
begin,EN	endure the brandy	0.42	0.58
begin,EN/EV	enjoy the translation	0.39	0.61
spot,EN	remember the brandy	0.34	0.66
begin,EV	enjoy the conference	0.24	0.76
begin,EV	begin the afternoon	0.20	0.80
spot,EV	remember the revolt	0.10	0.90
spot,EN/EV	remember the shower	0.08	0.92
begin,EN/EV	endure the shower	0.07	0.93
spot,EV	approve the conference	0.07	0.93
begin,EV	endure the revolt	0.03	0.97
spot,EN	approve the automobile	0.00	1.00
spot,EN/EV	approve the translation	0.00	1.00
spot,EN	organize the newspaper	0.00	1.00
spot,EN/EV	organize the breakfast	0.00	1.00
spot,EV	organize the afternoon	0.00	1.00

Table 4: CE and no-CE answers for single items in Experiment 2

The type-clash hypothesis seems to capture a tendency in the data rather than to predict the participants' answers in every single case. As shown by examples in Table 4, an item-wise analysis shows a continuum of behaviors rather than clear-cut separate categories: 1) EN nouns tend to have a strong majority of CE answers with begin-type verbs; 2) EV nouns tend to have a strong majority of no-CE answers with begin-type and spot-type verbs, but exceptions are possible (e.g. *enjoy the conference*) 3) not all the spot-type verbs block CE interpretations (e.g. *remember the brandy*); 4) the behavior of EN/EV ambiguous nouns is highly lexically determined (contrast ad example *begin the breakfast*, *enjoy the translation* and *endure the shower*).

Range of CEs. Per each V-Obj combination each participant elicited on average 1.4 CEs (range 1-6). Although we did not limit the number of CEs to be elicited, eliciting only one CE appears to be a common behavior. Nevertheless, if we only look

[5] $\alpha = .36$ when excluding EN/EV ambiguous nouns, showing that the low agreement was not due to their presence

at the cases when participants elicited not more than one CE, a variety of different CEs per VP was given (average 3.2, range 1-7). In several cases the elicited CEs cover a broader set than the one given by the telic and agentive qualia:

EN: consider the butter → 8 CEs: eat (x4), add, buy, churn, cook with, eat, make, melt

EN/EV: prefer the collection → 6 CEs: view (x3), buy, discuss, polish, study, watch

EV: start the semester → 3 CEs: spend, teach, join

Even within a theory of extended qualia (Busa et al., 2001), CEs like *buy* or *melt* are difficult to account for with the QS of *butter*.

The average of elicited CEs per each verb-object combination across all participants was 5 (range 1-15). Consider the following examples from the elicited CEs:

EN: start the portrait → 9 CEs: paint (x20), draw (x4), critique (x3), hang (x2), model (x2), sketch (x2), admire, pose for, review

EN/EV: finish the harvest → 15 CEs: gather (x5), collect (x4), plan (x3), reap (x3), sell (x3), load (x2), store (x2), cook, eat, enjoy, jar, package, pick, pull, ship

EV: enjoy the conference → 4 CEs: attend (x3), hold (x2), participate in, watch

Again, ascribing the sets of verbs for an EN-noun like *portrait* to the QS of the noun seems to be an unsatisfying solution, at least if the qualia are understood as specific verbs, rather than concepts (like, e.g., the agentive quale of *portrait* is *to paint*): the sets of elicited CEs form *semantically motivated verb classes* structured by semantic relations (synonymy, hyponymy), which can be understood as classes of plausible events. Among the elicited CEs there are also events which do not fall under the categories of agentive quale or telic quale: *hang*, *model*, *review*. As to EV objects (e.g. *conference*), they can also elicit CE readings (*enjoy attending/holding a conference*), and for EN/EV ambiguous objects like *harvest* both readings often give rise to elicited events. Note also that the elicited CEs include not only light verbs (*performing a translation*), which would be semantically largely transparent, but also full verbs (*reading / completing a translation*).

Table 5 reports on the amount of CEs which can be accounted for by a QS-based theory. The annotation was performed by the authors by assigning an agentive quale and a telic quale to each noun and comparing them with the elicited CEs. We considered qualia as classes of meaning, in order to cover also synonyms of the annotated qualia. Almost half of the elicited CEs did not fall in either the agentive quale category or in the telic quale category.

	tot	QS CEs		other CEs
		agentive	telic	
elicited CEs (tokens)	542	132 24.3%	162 29.9%	248 45.8%
elicited CEs (types)	205	31 15.1%	25 12.2%	149 72.7%

Table 5: CEs accounted for by a QS-based theory vs. other CEs

5 Conclusions

We are proposing an alternative mechanism for the recovery of covert events, according to which CEs are activated when the overt form cannot be given a plausible interpretation. We use a combination of self-paced reading and web-based elicitation to explore our hypothesis: the former detects processing costs differences, while the latter provides access to the range of CEs understood by speakers.

Results from a web elicitation study showed that the type-clash and the QS hypothesis are not enough to predict elicited CEs in a given context: CEs are elicited also for EV and EN/EV nouns, and in general the triggering of a CE seems to be highly lexically determined. Recovered CEs seems to fall in a wider range than those captured by the QS, and this range is also fairly wide when participants only give one answer.

While challenging the type-clash model, a plausibility-driven model can still retain the descriptive power of the sortal trigger hypothesis by subsuming it as a general tendency: EV nouns "tend to" show higher plausibilities for the verb/CE/noun combination, EN nouns show an opposite tendency and therefore more often require the recovery of CEs. Also, in a plausibility-driven model there is no reason why the range of CEs should be limited a priori: more than one of two clusters of meaning can be retrieved and ranked for plausibility.

Acknowledgments

The authors thank Berry Claus and Ulrike Padó for helpful discussions concerning this work.

References

F. Busa, N. Calzolari, and A. Lenci. 2001. Generative Lexicon and the SIMPLE model; developing semantic resources for NLP. In F. Busa and P. Bouillon, editors, *The Language of Word Meaning*, pages 333–349. Cambridge University Press.

A. Ferraresi, E. Zanchetta, M. Baroni, and S. Bernardini. 2008. Introducing and evaluating ukWaC, a very large web-derived corpus of English. In *Proceedings of the WAC4 workshop at LREC 2008*, Marrakech. ELRA.

W. N. Francis and H. Kucera. 1967. *Computational Analysis of Present-Day American English*. Brown University Press, Providence.

Steven Frisson and Brian McElree. 2008. Complement coercion is not modulated by competition: evidence from eye movements. *Journal of Experimental Psychology*, 34:1–11.

R. Jackendoff. 1997. *The Architecture of the Language Faculty*. MIT Press.

K. Krippendorff. 2004. *Content Analysys: An introduction to its methodology (2nd ed.)*. Sage, Thousand Oaks, CA.

Mirella Lapata and Alex Lascarides. 2003. A probabilisitic account of logical metonymy. *Computational Linguistics*, 29(2):263–317.

Mirella Lapata, Frank Keller, and Christoph Scheepers. 2003. Intra-sentential context effects on the interpretation of logical metonymy. *Cognitive Science*, 27(4):649–668.

Katja Markert and Udo Hahn. 1997. In support of the equal rights movement for literal and figurative language: A parallel search and preferential choice model. In *Proceedings of CogSci 1997*, pages 289–294.

Katja Markert and Udo Hahn. 2002. Understanding metonymies in discourse. *Artificial Intelligence*, 135:145–198.

B. McElree, M. J. Traxler, M. J. Pickering, R. E. Seely, and R. Jackendoff. 2001. Reading time evidence for enriched composition. *Cognition*, 78:B17–B25.

S. Padó, U. Padó, and K. Erk. 2007. Flexible, corpus-based modelling of human plausibility judgements. In *Proceedings of EMNLP*, pages 400–409, Prague, Czech Republic.

J. Pustejovsky. 1995. *The Generative Lexicon*. MIT Press, Cambridge MA.

L. Pylkkänen and B. McElree. 2006. The syntax-semantic interface: On-line composition of sentence meaning. In Mattew Traxler and Morton Ann Gernsbacher, editors, *Handbook of Psycholinguistics*, pages 537–577. Elsevier, 2nd edition.

Liina Pylkkänen and Brian McElree. 2007. An MEG study of silent meaning. *Journal of Cognitive Neuroscience*, 19(11):1905–1921.

Liina Pylkkänen, Andrea E. Martin, Brian McElree, and Andrew Smart. 2009. The anterior midline field: Coercion or decision making? *Brain and Language*, 108(3):184–190.

E. Shutova and S. Teufel. 2009. Logical metonymy: Discovering classes of meanings. In *Proceedings of DiSCo 2009 Workshop*, Prague, Czech Republic.

R. Snow, B. O'Connor, D. Jurafsky, and A.Y. Ng. 2008. Cheap and fast - but is it good? Evaluating non-expert annotations for natural language tasks. In *Proceedings of EMNLP 2008*.

M. J. Traxler, M. J. Pickering, and B. McElree. 2002. Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language*, 47:530–547.

Matthew J. Traxler, Brian McElree, Rihana S. Williams, and Martin J. Pickering. 2005. Context effects in coercion: Evidence from eye movements. *Journal of Memory and Language*, 53:1–25.

Z. Vendler. 1968. *Adjectives and Nominalizations*. Mouton, The Hague, The Netherlands.