

Andrej Malchukov and Bernard Comrie (Eds.)

**Valency Classes in the World's Languages**

Vol. 1

# **Comparative Handbooks of Linguistics**

---

Edited by  
Edith Moravcsik and Andrej Malchukov

## **Volume 1.1**

# Valency Classes in the World's Languages

---

Volume 1

Introducing the Framework, and  
Case Studies from Africa and Eurasia

Edited by  
Andrej Malchukov  
Bernard Comrie

**DE GRUYTER**  
MOUTON

ISBN 978-3-11-033294-0  
e-ISBN (PDF) 978-3-11-033881-2  
e-ISBN (EPUB) 978-3-11-039527-3  
ISSN 2364-4354

**Library of Congress Cataloging-in-Publication Data**

A CIP catalog record for this book has been applied for at the Library of Congress.

**Bibliographic information published by the Deutsche Nationalbibliothek**

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.dnb.de>.

© 2015 Walter de Gruyter GmbH, Berlin/Boston  
Cover image: Jupiterimages/PHOTOS.com/thinkstock  
Typesetting: Meta Systems Publishing & Printservices GmbH, Wustemark  
Printing and binding: CPI books GmbH, Leck  
♻️ Printed on acid-free paper  
Printed in Germany

[www.degruyter.com](http://www.degruyter.com)

# Contents

## Volume 1

**Abbreviations — ix**

**Acknowledgments — x**

**List of authors — xi**

## **Part I: The Leipzig Valency Classes Project: Introducing the Framework**

Bernard Comrie, Iren Hartmann, Martin Haspelmath, Andrej Malchukov,  
and Søren Wichmann

**1 Introduction — 3**

Andrej Malchukov and the Leipzig Valency Classes Project team

**2 Leipzig Questionnaire on valency classes — 27**

Martin Haspelmath and Iren Hartmann

**3 Comparing verbal valency across languages — 41**

Andrej Malchukov

**4 Valency classes and alternations: parameters of variation — 73**

Martin Haspelmath

**5 Transitivity prominence — 131**

Damian E. Blasi

**5a Assessing transitivity prominence from a statistical perspective:**

**A commentary on Martin Haspelmath's "Transitivity prominence" — 149**

Søren Wichmann

**6 Statistical observations on implicational (verb) hierarchies — 155**

## **Part II: Case Studies**

### **Africa**

Martina Ernszt, Tom Güldemann, Alena Witzlack-Makarevich

**7 Valency in Nilng — 185**

Denis Creissels

**8 Valency properties of Mandinka verbs — 221**

Ronald P. Schaefer and Francis O. Egbokhare

**9 Emai valency classes and their alternations — 261**

Joseph Dele Atoyebi

**10 Valency classes in Yorùbá — 299**

Csilla Kász

**11 Valency properties of verbs in Modern Standard Arabic — 327**

## **Eurasia**

Jóhanna Barðdal

**12 Icelandic valency classes: oblique subjects, oblique ambitransitives and the actional passive — 367**

Michela Cennamo

**13 Valency patterns in Italian — 417**

Michael Daniel and Victoria Khurshudian

**14 Valency classes in Eastern Armenian — 483**

Bernard Comrie, Madzhid Khalilov and Zaira Khalilova

**15 Valency and valency classes in Bezhta — 541**

Andrej Malchukov and Igor V. Nedjalkov

**16 Valency classes in Even (North Tungusic) in a comparative Tungusic perspective — 571**

Edward J. Vajda

**17 Valency properties of the Ket verb clause — 629**

Robert Schikowski, Netra Paudyal, and Balthasar Bickel

**18 Flexible valency in Chintang — 669**

Bingfu Lu, Guohua Zhang and Walter Bisang

**19 Valency classes in Mandarin — 709**

Hideki Kishimoto, Taro Kageyama and Kan Sasaki

**20 Valency classes in Japanese — 765**

Anna Bugaeva

**21 Valency classes in Ainu — 807**

**Language index — 855**

**Subject index — 861**

## Volume 2

Andrej Malchukov and the Leipzig Valency Classes Project team

**Leipzig Questionnaire on valency classes — xi**

## Part II: Case Studies (cont.)

### Austronesia and the Pacific

Masayoshi Shibatani and Ketut Artawa

**22 Balinese valency classes — 877**

Thomas Connors, John Bowden and David Gil

**23 Valency classes in Jakarta Indonesian — 941**

Sebastian Nordhoff

**24 Case frames in Sri Lanka Malay — 987**

Claire Moyse-Faurie

**25 Valency classes in Xârâcùù (New Caledonia) — 1015**

Nicholas Evans

**26 Valency in Nen — 1069**

Eva Schultze-Berndt

**27 Complex verbs, simple alternations: valency and verb classes in Jaminjung — 1117**

### Americas

Osahito Miyaoka

**28 Valency classes in Central Alaskan Yupik, an Eskimoan language — 1165**

Richard Rhodes and Rand Valentine

**29 Transitivity in Ojibwe — 1205**

Iren Hartmann

**30 Valency Classes in Hooçak (Ho-Chunk) — 1265**

Honoré Watanabe

**31 Valency classes in Sliammon Salish — 1313**

Zarina Estrada-Fernández, Mercedes Tubino and Jesús Villalpando

**32 Valency classes in Yaqui — 1359**

Eric Campbell

**33 Valency classes in Zenzontepec Chatino — 1391**

Christian Lehmann

**34 Valency classes in Yucatec Maya — 1427**

Frank Seifart

**35 Valency classes in Bora — 1481**

Fernando Zúñiga

**36 Valency classes in Mapudungun — 1515**

**Part III: Theoretical Outlook**

Christian Lehmann

**37 Situation types, valency frames and operations — 1547**

Tasaku Tsunoda

**38 The hierarchy of two-place predicates: its limitations and uses — 1597**

Beth Levin

**39 Verb classes within and across languages — 1627**

Cliff Goddard

**40 Verb classes and valency alternations (NSM approach), with special reference to English physical activity verbs — 1671**

**Language index — 1703**

**Subject index — 1709**



# Abbreviations

1	first person	INF	infinitive
2	second person	INS	instrumental
3	third person	INTR	intransitive
A	agent-like argument of canonical transitive verb	IPFV	imperfective
ABL	ablative	IRR	irrealis
ABS	absolutive	LOC	locative
ACC	accusative	M	masculine
ADJ	adjective	N	neuter
ADV	adverb(ial)	N-	non- (e.g. NSG nonsingular, NPST nonpast)
AGR	agreement	NEG	negation, negative
ALL	allative	NMLZ	nominalizer/nominalization
ANTIP	antipassive	NOM	nominative
APPL	applicative	OBJ	object
ART	article	OBL	oblique
AUX	auxiliary	P	patient-like argument of canonical transitive verb
BEN	benefactive	PASS	passive
CAUS	causative	PFV	perfective
CLF	classifier	PL	plural
COM	comitative	POSS	possessive
COMP	complementizer	PRED	predicative
COMPL	completive	PRF	perfect
COND	conditional	PRS	present
COP	copula	PROG	progressive
CVB	converb	PROH	prohibitive
DAT	dative	PROX	proximal/proximate
DECL	declarative	PST	past
DEF	definite	PTCP	participle
DEM	demonstrative	PURP	purposive
DET	determiner	Q	question particle/marker
DIST	distal	QUOT	quotative
DISTR	distributive	RECP	reciprocal
DU	dual	REFL	reflexive
DUR	durative	REL	relative
ERG	ergative	RES	resultative
EXCL	exclusive	S	single argument of canonical intransitive verb
F	feminine	SBJ	subject
FOC	focus	SBJV	subjunctive
FUT	future	SG	singular
GEN	genitive	TOP	topic
IMP	imperative	TR	transitive
INCL	inclusive	VOC	vocative
IND	indicative		
INDF	indefinite		

## Acknowledgments

The project on Valency Classes in the World's Languages, some of whose results are reported in this volume, was funded by the Deutsche Forschungsgemeinschaft (DFG) as part of the DFG-Forschergruppe 742 (Grammar and Processing of Verbal Arguments), and was based at the Department of Linguistics of the Max Planck Institute for Evolutionary Anthropology, Leipzig. This volume partly originated from two workshops (2010, 2013) and a conference (2011) held in Leipzig and a workshop held at the 44<sup>th</sup> Annual Meeting of the Societas Linguistica Europaea at the Universidad de la Rioja, Logroño, Spain. We would like to thank the contributors to the volume and other participants in the conference and workshops for useful discussion. We further thank Max Planck staff members Claudia Bavero, Claudia Büchel, Julia Cissewski, and Peter Fröhlich for their administrative and technical support in organizing the Leipzig conference and workshops. We benefited greatly from the contribution of our highly motivated and reliable student assistants in editing and typesetting this volume; thanks are due especially to Diana Schackow and Tina Gregor. Finally, we thank Claudia Bavero for her invaluable work preparing the indexes and correcting the proofs.

Andrej Malchukov and Bernard Comrie  
Leipzig, 15 May 2015

# List of authors

**Ketut Artawa**, Udayana University, E-mail: artawa56@yahoo.com  
**Joseph Dele Atoyebi**, Collège Boréal, E-mail: joeatoyebi@gmail.com  
**Jóhanna Barðdal**, Department of Linguistics, Ghent University, E-mail: johanna.barddal@ugent.be  
**Balthasar Bickel**, Universität Zürich, E-mail: balthasar.bickel@uzh.ch  
**Walter Bisang**, Johannes Gutenberg University, Mainz, E-mail: wbisang@uni-mainz.de  
**Damian E. Blasi**, Max Planck Institute for Mathematics in the Sciences & Max Planck Institute for Evolutionary Anthropology, E-mail: damian.blasi@mis.mpg.de  
**John Bowden**, Max Planck Institute for Evolutionary Anthropology, E-mail: john.bowden@jakartafield.org  
**Anna Bugaeva**, National Institute for Japanese Language and Linguistics, E-mail: bugaeva@ninja.ac.jp  
**Eric Campbell**, University of California, Santa Barbara, E-mail: ecampbell@linguistics.ucsb.edu  
**Michela Cennamo**, University of Naples Federico II, E-mail: micennam@unina.it  
**Bernard Comrie**, Max Planck Institute for Evolutionary Anthropology & University of California, Santa Barbara, E-mail: comrie@linguistics.ucsb.edu  
**Tom Connors**, University of Maryland, E-mail: tom.connors@gmail.com  
**Denis Creissels**, Université Lyon 2, E-mail: denis.creissels@univ-lyon2.fr  
**Michael Daniel**, Higher School of Economics / Moscow State University, E-mail: misha.daniel@gmail.com  
**Francis O. Egbokhare**, University of Ibadan, E-mail: fo.egbokhare@mail.ui.edu.ng  
**Martina Ernst**, Max Planck Institute for Evolutionary Anthropology, E-mail: martinaernst@freenet.de  
**Zarina Estrada-Fernández**, Universidad de Sonora, E-mail: zarinaef@gmail.com  
**Nicholas Evans**, Australian National University, E-mail: nicholas.evans@anu.edu.au  
**David Gil**, Max Planck Institute for Evolutionary Anthropology, E-mail: gil@shh.mpg.de  
**Cliff Goddard**, Griffith University, E-mail: c.goddard@griffith.edu.au  
**Tom Güldemann**, Humboldt-Universität zu Berlin & Max Planck Institute for Evolutionary Anthropology, E-mail: tom.gueldemann@rz.hu-berlin.de  
**Iren Hartmann**, Max Planck Institute for Evolutionary Anthropology, E-mail: iren.hartmann@gmail.com  
**Martin Haspelmath**, Max Planck Institute for Evolutionary Anthropology, E-mail: haspelmath@shh.mpg.de  
**Taro Kageyama**, National Institute for Japanese Language and Linguistics & Kwansei Gakuin University, E-mail: tkage@ninja.ac.jp  
**Csilla Kász**, Christian-Albrechts-Universität zu Kiel, E-mail: ckasz@isfas.uni-kiel.de  
**Madzhid Khalilov**, G. Tsadasa Institute for Language, Literature and Art of the Daghestan Scientific Center of the Russian Academy of Sciences & Max Planck Institute for Evolutionary Anthropology, E-mail: madjid-kh@mail.ru  
**Zaira Khalilova**, Max Planck Institute for Evolutionary Anthropology, E-mail: zaira.khalilova@gmail.com  
**Victoria Khurshudian**, Centre national de la recherche scientifique, E-mail: vkhurshudian@gmail.com  
**Hideki Kishimoto**, Kobe University, E-mail: kishimot@lit.kobe-u.ac.jp  
**Christian Lehmann**, Universität Erfurt, E-mail: christian.lehmann@uni-erfurt.de  
**Beth Levin**, Stanford University, E-mail: bclevin@stanford.edu  
**Bingfu Lu**, Nanchang University, E-mail: lubingfu@yahoo.com

- Andrej Malchukov**, Institute for Linguistic Research, Russian Academy of Sciences,  
St. Petersburg & Johannes Gutenberg University, Mainz & Max Planck Institute for  
Evolutionary Anthropology, E-mail: malchuko@uni-mainz.de
- Osahito Miyaoka**, Kyoto University, E-mail: miyaoka@cup.ocn.ne.jp
- Claire Moyse-Faurie**, Centre national de la recherche scientifique, E-mail: moyse@vjf.cnrs.fr
- Igor Nedjalkov**, St. Petersburg State University & Institute for Linguistic Research, Russian  
Academy of Sciences, St. Petersburg, E-mail: igor\_nedjalkov@mail.ru
- Sebastian Nordhoff**, Freie Universität Berlin, E-mail: sebastian.nordhoff@langsci-press.org
- Netra Paudyal**, Universität Leipzig, E-mail: netrapaudyal@yahoo.com
- Richard A. Rhodes**, University of California Berkeley, E-mail: rrhodes@berkeley.edu
- Kan Sasaki**, Sapporo Gakuin University, E-mail: ksasaki@sgu.ac.jp
- Ronald P. Schaefer**, Southern Illinois University Edwardsville, E-mail: rschaefer@siue.edu
- Robert Schikowski**, Universität Zürich, E-mail: robert.schikowski@uzh.ch
- Eva Schultze-Berndt**, The University of Manchester,  
E-mail: eva.schultze-berndt@manchester.ac.uk
- Frank Seifart**, Universiteit van Amsterdam, E-mail: F.C.Seifart@uva.nl
- Masayoshi Shibatani**, Rice University, E-mail: shibo@rice.edu
- Tasaku Tsunoda**, National Institute for Japanese Language and Linguistics,  
E-mail: tsunoda@ninjal.ac.jp
- Mercedes Tubino**, Western Michigan University, E-mail: mercedes.tubino-blanco@wmich.edu
- Edward J. Vajda**, Western Washington University, E-mail: edward.vajda@wwwu.edu
- J. Randolph Valentine**, University of Wisconsin-Madison, E-mail: jrvalent@wisc.edu
- Jesús Villalpando**, University of Colorado Boulder, E-mail: Jesus.Villalpando@colorado.edu
- Honoré Watanabe**, Research Institute for Languages and Cultures of Asia and Africa, Tokyo  
University of Foreign Studies, E-mail: honore@aa.tufs.ac.jp
- Søren Wichmann**, Max Planck Institute for Evolutionary Anthropology,  
E-mail: wichmannsoeren@gmail.com
- Alena Witzlack-Makarevich**, Christian-Albrechts-Universität zu Kiel,  
E-mail: awitzlack@isfas.uni-kiel.de
- Guohua Zhang**, Research Center of Chinese Language History, Zhejiang University & School of  
Foreign Languages, Zhongnan University of Economics and Law, E-mail: huntergh@126.com
- Fernando Zúñiga**, Universität Bern, E-mail: fernando.zuniga@isw.unibe.ch

---

Part I: **The Leipzig Valency Classes Project:  
Introducing the Framework**



Bernard Comrie, Iren Hartmann, Martin Haspelmath,  
Andrej Malchukov, and Søren Wichmann

# 1 Introduction

## 1 General information

The present volume is one of the main products of the project “Valency classes in the languages of the world”, carried out by the authors of this chapter. This project grew out of an earlier project “Ditransitive constructions in the languages of the world”, which gave rise, among other publications, to Malchukov et al. (2010). The topic of the later project is clearly related to that of the earlier one, but it is also significantly different. The Valency Classes project has aimed to investigate the argument-structure properties of verbs of different valency classes in a typological perspective. It thus continued the line of research of the ditransitive project, but focused on lexical properties of verbs and extended the research questions to a variety of other valency classes, thus making a contribution to the nascent field of lexical typology. The present volume presents the results of an empirical study of a relatively large set of core verb meanings (70) for a relatively small set of languages (30);<sup>1</sup> it can be viewed as a compromise between Levin’s (1993) classification and semantically based classifications proposed in the typological literature. The other main product of the project is the on-line database Valency Patterns Leipzig (Hartmann et al. 2013).

## 2 Background

### 2.1 Approaches to the study of valency classes

All verbs in a language have different meanings, but with respect to their syntactic patterning, they show striking similarities and fall into a relatively circumscribed number of valency classes whose members behave alike. By “syntactic patterning”, we refer not just to the coding patterns (i.e. the ways in which the verbal arguments are flagged and indexed), but also to the behavior of verbs with respect to alterna-

---

<sup>1</sup> The accompanying database increases this to 80 verb meanings and 36 languages. Some of the extra 10 verb meanings, and also further verb meanings selected as interesting by individual authors, are included in the chapters of Part II.

tions such as causative, passive, applicative and other constructions that are not fully productive for all verbs (e.g., secondary predicates, certain word order phenomena). It has often been noted that these syntactic verb classes typically correlate with semantic classes. The syntactic properties of verbs can be studied separately from the semantic classes, so the strength of the correlation is an empirical question.

The literature on syntactic verb classes is vast, yet with few exceptions it is limited to relatively few languages, most of them European. A classic study in this area is Levin's (1993) book on English verb classes, where she argues that verb classes are semantically based and can be identified in terms of their argument structures, possible argument structure alternations, as well as further syntactic diagnostics (such as middle alternation, unspecified object deletion, *there*-insertion, etc.). Unlike earlier studies, which divided the verbal lexicon into a few highly general classes (e.g., stative vs. active verbs, intransitive vs. transitive vs. ditransitive, or, for intransitives, unaccusatives and unergatives), Levin attempts a much more fine-grained classification, which is constructed in a bottom-up fashion, and where verb classes are defined in terms of their overall syntactic distribution. Levin's study (as well as subsequent work with M. Rappaport Hovav; e.g., Levin & Rappaport Hovav 2005) has been highly influential not only in the theoretical work on lexical semantics, but also in computational linguistics, and underlies verb ontologies in WordNet (Fellbaum 1998) and extensions thereof such as VerbNet (see Kipper-Schuler 2005 for references) and FrameNet (see, e.g., Fillmore et al. 2003; see also Schulte im Walde 2003 for discussion of verb classes in German).

But neither Levin's study nor the pioneering study *Experimental Investigation of the Russian verb* by Apresjan (1969), where verb classes were also established on syntactic grounds, have been extended cross-linguistically. Jones (1994) is a small collection of working papers dealing with verb classes in English, German, Korean and Bangla, stemming from an MIT-based project explicitly aiming to extend the Levin-style classification to other languages. There are also occasional studies dealing with other languages (cf. Fukui et al. 1985 on Japanese; Vogel 2003 on Jarawara). Even contrastive studies devoted to a single verb type, such as the study of interaction verbs in English, German, Hungarian and Maori in Blume (1998), are rare.

This lack of an extension to more languages does not seem to be accidental. In spite of its merits, Levin's approach faces a number of problems which become evident once one attempts to extend it beyond English. First, since Levin's classes are constructed on syntactic criteria, they are not always semantically coherent. This is a serious drawback for typological studies, where the phenomena to be investigated have to be defined in semantic terms to make a comparison possible. The main question is: Which aspects of this classification are universal and which are language-particular? Clearly, the details cannot be universal as the study refers to specific language forms (e.g., encoding of arguments through specific case



forms, prepositions, etc.). Similarly, the encoding of alternative constructions, as well as syntactic diagnostics like the English Middle alternation are clearly not universal. Yet, it is expected that universal cross-linguistic patterns do exist, insofar as both cross-linguistically recurrent coding properties as well as the availability of certain alternations have a semantic motivation. This has already been anticipated in a work on transitivity alternation by Pinker (1989), whose approach is close to Levin's, but additionally tries to provide explanations for encoding options in terms of semantic properties of verb classes on the one hand and the semantics of the alternation on the other hand (thus, for example, the middle alternation targets EFFECT verbs like *break*, not CONTACT verbs like *hit*).

A different tradition in the research on verb classes (or, valency classes) takes its origin in the work on Case Grammar, different versions of which were developed in the 1970s by Fillmore, Gruber, Cook, J. Anderson, Jackendoff, and Chafe, among others. In this tradition, verb classes are identified in terms of the semantic roles of the verbal arguments. A related approach has been developed (particularly in France and Germany) in the work by Tesnière, Gross, Helbig, and others (see Ágel 2006 for a comprehensive bibliography of valency research and Herbst & Götz-Votteler 2007 for a representative sample of contemporary approaches to valency research). In this approach, however, verbal valency types are defined more in terms of formal than semantic criteria (see, e.g., Somers 1987 for a comparative treatment). In the subsequent literature, the argument structure of verb classes has played an important role in linguistic theories of different persuasions (see, for instance, Levin & Rappaport Hovav 2005 and Butt 2006 for overviews and discussion), yet this research has rarely been carried out systematically. In the generative literature the issue of subcategorization frames of different verb classes has been present for a long time, but it has not been addressed systematically. Thus, while there is a large literature on individual verb classes found to be of particular theoretical interest (cf. the work by Grimshaw 1990; Pesetsky 1995 and others on emotion verbs, which present challenges for argument linking), comprehensive studies of verb classes in other languages have not been attempted.

While the empirical basis of mainstream generative grammar has not been very broad until recently, other theories like Role and Reference Grammar (Van Valin & Lapolla 1997; Van Valin 2001) and Lexical Decomposition Grammar (Wunderlich 2006; Stiebels 2000) have developed a strong typological orientation. These studies pioneered systematic research into semantic argument types of languages of different alignment (in particular, the work by Van Valin has contributed to the study of argument alternations cross-linguistically). Yet, these studies operate in terms of broad valency-based classes rooted in aspectual properties and lexical decomposition, and never reach the level of granularity of Levin's classification (see, for instance, the study of verb classes in Wunderlich 2006; cf. Joppen-Hellwig 2001). This is even true for valency studies, which have been specifically designed for capturing lexical variation in syntactic properties. Valency dictionaries remain few

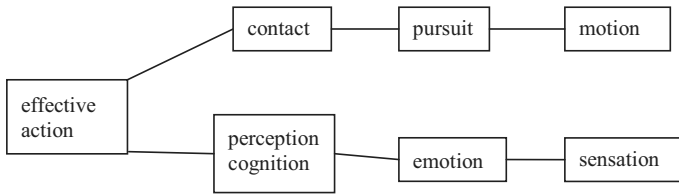
and are mostly confined to European languages (see Schumacher 2006a, b; Busse 2006 for references). It is also instructive that the monumental handbook of dependency and valency edited by Ágel (2006), while containing some contrastive studies of valency, does not provide a systematic overview even for European languages.

In typology, the cross-linguistic study of verb classes rooted in the Case Grammar and valency traditions was advanced especially by Lazard (1998), Lehmann (1991), and Drossard (1991). This research has contributed much to the understanding of how different verb types pattern in languages of different alignment, yet, like much of the research discussed earlier, it deals with rather broad classes. A partial exception is Lehmann's work, which achieves a finer cross-classification through the use of different criteria: aspectual (Vendler-style) criteria, verbal valency, and properties of arguments. Another approach, which is typological in nature, although it has been applied to English, is that of Dixon (1991, 2005). Dixon's taxonomy of verb classes, which predates Levin's classification, is different inasmuch as Dixon's is semantic in nature – the syntactic properties (even verbal valency) are secondary for his classification. The established classes are very general (although some are further subdivided) and include verbs of different valency; thus the MOTION type includes verbs such as *run* and *take*, CORPOREAL includes verbs such as *eat* and *die*, and so on. Yet, this classification, like Lehmann's, is of obvious interest to typology because it is semantically based and thus can be more easily extended to the study of other languages.

One general drawback of the typological work mentioned above is that it remains taxonomic in nature and has not led to broader generalizations. The only exception, to our knowledge, is the work by Tsunoda (1981, 1985), who proposed the following hierarchy of verb types that predicts the distribution of intransitive and transitive patterns in individual languages:

Effective action > Perception > Pursuit > Knowledge > Feeling > Relation

This hierarchy represents a scale stretching from the more transitive verb types on the left to the less transitive verb types on the right. It is called a hierarchy since it predicts that if a verb type lower in the hierarchy allows for a transitive case frame (NOM-ACC in accusative languages or ERG-ABS in ergative languages), so do verb types higher in the hierarchy. Tsunoda's approach is generally considered an important contribution to the study of verb classes (cf. Lazard 1994; Lehmann 1991). Yet until recently, this line of research has not been pursued further in typological work, as it faces a number of empirical problems (see Malchukov 2005 for discussion and references). Malchukov (2005), however, suggests that counterexamples can be accounted for by decomposing Tsunoda's hierarchy into the two dimensions of decreased patienthood of P (from 'break' to 'look for' and 'go to'), and decreased agenthood of A (from 'break' to 'like' and 'hurt'), recasting Tsunoda's hierarchy as two-dimensional:



**Fig. 1:** Two-dimensional Transitivity Hierarchy (semantic map).

Importantly, this hierarchy can be seen as a part of a larger semantic map showing how different verb types pattern in the semantic space from transitive to intransitive verbs, as explained in 2.2 (see Haspelmath 2003 for a general discussion of the semantic map approach).

Among recent contributions to the issues of verb classification and valency alternation one can mention the collection of papers Kratochvíl et al. (2011), focusing on the concept of transitivity and on transitivity mismatches (syntactic vs. morphological transitivity, etc.), another collection of papers on issues relating to verb classification from various theoretical perspectives (De Clerck et al. 2013), as well as Croft's (2012) monograph providing a comprehensive discussion of argument structure and argument alternations from a cross-linguistic perspective. These studies are representative of the state of art of the research into the issues of verb classification and argument alternations, yet they do not attempt to offer a comprehensive typology of valency classes.

As is clear from the brief discussion of the research on valency classes, even though each of the approaches has made important contributions to the study of verb classes, they all have certain drawbacks as well. The Valency Classes project selectively built on the advantages of different approaches, in particular the following desiderata:

- the taxonomy should be fine-grained enough, as in Levin's approach;
- it should be semantically based such that it can be applied to other languages, as in the different typological approaches (Lehmann, Dixon, and others);
- since the taxonomy is semantically based, the syntactic properties of these classes can be studied without the danger of circularity (cf. the work on ditransitive constructions reported in Malchukov et al. 2010);
- the general purpose is not just taxonomic, but aims to uncover universal and language-particular properties of valency classes in terms of coding properties and behavior (as in Tsunoda's hierarchies/semantic maps).

## 2.2 Semantic maps

In the Valency Classes project, as in the earlier Ditransitives project, cross-linguistic similarities between valency classes are captured in the form of semantic maps.

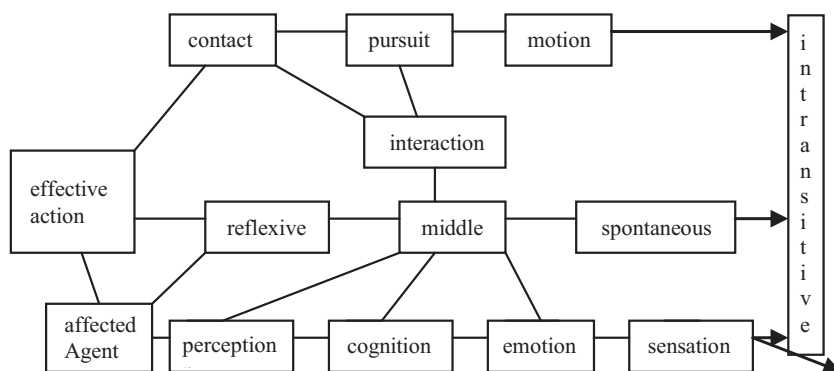
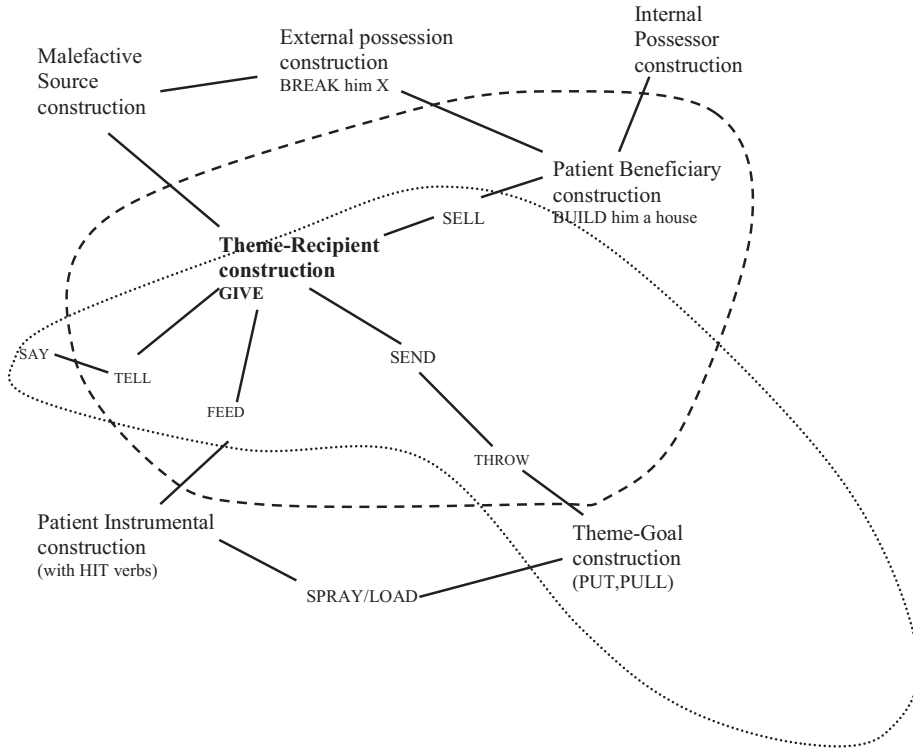


Fig. 2: Transitivity reduction on a semantic map.

The semantic map method (as developed by Anderson 1986; Haspelmath 1997, 2003; and Croft 2001; cf. also Cysouw et al. 2010) aims at uncovering semantic similarities among individual categories on the basis of cross-linguistically recurrent polysemy patterns. It is assumed – because of iconicity (Haiman 1985) – that recurrent similarity in form reflects similarity in meaning. The categories that are similarly encoded are therefore put contiguously in a semantic space. The central assumption of the semantic map methodology is that the semantic space is universal; thus a semantic map, once established empirically, makes universal predictions about possible and impossible (or rather, probable and improbable) polysemy patterns. Semantic maps have been applied to different domains, including verb types. Thus, Malchukov (2005) integrates Tsunoda’s hierarchy into the semantic map in Figure 2.

This map has the immediate advantage, as compared to Tsunoda’s one-dimensional hierarchy, that it can incorporate different extensions of the transitive frame across different dimensions that are conflated by Tsunoda. For example, English differs from Japanese in extending the transitivity pattern further along the second subhierarchy (to emotion verbs like *like* or *fear*, which pattern intransitively in Japanese; cf. *osoroshi* ‘fear(ful)’), but is more conservative with respect to the first subhierarchy (PURSUIT verbs in Japanese, unlike in English, pattern transitively; cf. *matu* ‘wait for’). This map also includes some other verb types proposed in the literature, which, however, differ in granularity. Thus, it additionally includes the categories of INHERENTLY REFLEXIVE, MIDDLE, and SPONTANEOUS verbs representing the reduction of transitivity along the dimension of distinguishability between subject and object (cf. Kemmer 1993). It also includes INTERACTION verbs discussed by Blume (1998), which are linked to PURSUIT and CONTACT verbs on the map. On the one hand, interaction verbs show similarities to PURSUIT verbs (both are treated as subclasses of AIMING verbs by Lazard 1998). Clearly, this hierarchy is still not fine-grained enough, as many classes may need further decomposition; for example, it is instructive that both Tsunoda’s PURSUIT verbs and



**Fig. 3:** Semantic map for ditransitive verbs.

The lines indicate an approximate range of verbs participating in the English Double Object Construction and *to*-NP Construction; their intersection delimits verb types participating in a dative alternation; DOC -----, *to*-NP .....

Blume's INTERACTION verbs are distributed across several subclasses in Levin's (1993) study of English verb classes. Yet the hypothesis embodied in the hierarchy clearly deserves to be explored cross-linguistically, as it constrains possible extensions of specific case patterns across the verbal lexicon in individual languages. Semantic maps also played an important role in the project on the typology of ditransitive constructions. Thus, for the ditransitive domain, the semantic map shown in Fig. 3 is proposed in Malchukov et al. (2010 b).

For the present purpose it is important to note that such maps are well suited for the representation of specific argument patterns as well as for argument alternations. Thus, Fig. 3 shows extensions of two different strategies involved in a dative alternation across valency classes. The map is arrived at empirically, through the study of extensions of particular strategies across the verbal lexicon. It also incorporates the hierarchies proposed in the literature, for example the hierarchy of transfer verbs proposed by Croft et al. (2001): *give* > *send* > *throw*. In turn, the pattern underlying the semantic map is guided by semantic considerations, i.e.

the relative similarities between the verb types in terms of their argument structure. For instance, as claimed in earlier work (Levin & Rappaport Hovav 2005) SEND verbs are intermediate between transfer of possession verbs like GIVE and caused motion verbs like THROW. This ensures that – on iconicity assumptions – the extension of a coding strategy (construction) will occupy a contiguous region of the semantic map. In short, the semantic map methodology captures the patterns in diversity: even though different languages (and even a single language) may exhibit different constructions, the resulting configuration may make claims to universality. Similar maps can be established for other verbal domains; see, for example, Comrie & van den Berg (2006) for a semantic map of experiential verbs in Daghestanian languages.

As for the domain of syntactic properties (alternations), it is also expected that syntactic behavior will largely respect the map, since the alternations themselves are sensitive to different functional properties that correlate with lexical properties. The semantic properties are not specifically discussed in Levin's study, but have been extensively discussed elsewhere (e.g., by Goldberg 1995 and Pinker 1989).

A further advantage of the semantic map methodology is that maps from different domains can be integrated as long as they deal with common types. Thus, the map for the transitive domain (in Fig. 2) can eventually be integrated with the map for the ditransitive domain (in Fig. 3) (Malchukov et al. 2010b: 52), and the two together would represent a part of a more comprehensive map for valency classes. Finally, the semantic maps method can be further refined through the application of statistical scaling methods, which are well suited for representing relative frequencies of individual patterns (cf. the semantic maps of motion verbs generated from parallel texts in Wälchli & Cysouw 2012). Within the present volume, the chapters by Malchukov, Haspelmath et al. and Wichmann all make use of semantic maps and hierarchies to represent cross-linguistic generalization, either conventional maps ("implicational maps" in terms of Haspelmath) as illustrated in Fig. 2 and 3 above, or automatically generated maps through the use of clustering techniques (see the NeighborNet graphs in the chapter by Wichmann).

## 3 Objectives and results

### 3.1 Objectives

#### 3.1.1 General approach

In the project that led to this volume, the participants studied the syntactic valency properties of verbs for a relatively large number of verbs, 70 (for the database, 80), in a relatively small sample of languages, 30 (for the database, 36). For each verb, they looked at two kinds of properties:

- coding properties: flagging (= case and adpositional marking) and indexing (= “agreement/cross-referencing”);
- behavioral properties: agent demotion, agent (causee) addition, object rearrangement, indefinite object omission, reflexivization, reciprocalization (and possibly others).

Behavioral properties have not figured prominently in earlier discussions of European languages, which are predominantly dependent-marking, but they are indispensable in the discussion of head-marking languages. For example, different subject types in split intransitive languages (or languages with semantic alignment; Donohue & Wichmann 2008) are typically differentiated through indexing rather than flagging.

The valency alternations considered are similar to those studied for ditransitive constructions: passivization, reflexivization, reciprocalization, as well as object rearrangement (applicativization), indefinite object omission and causativization (see Malchukov et al. 2010b for a list of other behavioral properties addressed). It should be noted that many alternations are of course limited to a subset of verbs and verb classes. This is also evident in Levin’s work, where many alternations are relevant only to certain valency classes. For example, the Unspecified Object alternation or the Middle alternation are relevant for some subgroups of transitive verbs, but not for intransitive ones.

The syntactic alternations studied are independent of the degree of morphological synthesis of a language. As noted above, languages with a richer morphology than English tend to make use of morphological valency-changing operations (applicatives, causatives, anticausatives, etc.) instead of the *uncoded alternations* that are widely found in English (thus the dative alternation will involve the use of applicatives in Bantu languages; see, for instance, Van Valin 2001: 60–65). Thus in Even, a Tungusic language, the “middle alternation” is signaled by the Mediopassive marker, the “inchoative-causative alternation” by the Causative marker (in competition with the Mediopassive), and equivalents of English verbs allowing for a “reciprocal alternation” commonly involve a lexicalized Sociative marker (e.g., *baka-lda* [find-SOC] ‘meet’) (Malchukov 1995). For these languages the applicability of the valency-changing markers across the verbs must be established. Things get more complex when several markers compete for a certain domain, as in the case of causative and anticausative markers, which show particular preferences for specific verb types established by Haspelmath (1993a) and Nichols et al. (2004). Another complication concerns languages displaying several different constructions corresponding to one of our broad construction types. For example, Even distinguishes between Mediopassives and Adversative Passives (corresponding to our agent demotion construction), and Philippine languages display multiple passive-like voices targeting different roles (see Comrie 2008 for further discussion of passives). It is important to establish how these valency markers extend across the verbal

lexicon. In our research we aimed to establish not only availability of particular valency markers for certain verb types, but also their function in the case of polyfunctional markers. For example, the reflexive marker *-sja* in Russian (and other Slavic languages) is notoriously polysemous, and its meanings are sensitive to the verb type involved (see Geniušienė 1987). Similarly, in many languages the interpretation of the applicative markers depends on the verb type (cf. Gerdts & Kiyosawa 2005 on Salishan applicatives). Taking into account the function of polysemous markers allowed the identification of further clustering in the verbal lexicon.

On the other hand, the verbal lexicon of languages of the more isolating type (like English) can be investigated through the study of (morphologically) uncoded valency alternations (as defined above). The study of such languages must be informed by the previous literature, not only in the descriptive tradition, but also in the generative tradition. Some other alternations must be established on a language particular basis. For instance, in many languages the coding patterns will be more informative for lexical typology than they are in English. Thus, Sauerland (1994) shows that German makes finer coding distinctions among verb classes than English, which is due to the availability of verb-specific dative and genitive cases in German. Case and adpositional patterns will be an important diagnostic for lexical splits in the project. For instance, in Tlapanec (an Otomanguean language of Mexico discussed in Wichmann 2009), verbs such as GET and RUN, which encode the single animate participant in the Dative case, contrast with verbs such as KILL and APPEAR, which encode the case-marked participant in the Absolutive (for a transitive verb such as KILL the case-marked participant will be the undergoer and for an intransitive such as APPEAR it will be the actor). Case is a sufficiently frequent phenomenon that it can be used as a consistent diagnostic across many languages, especially on a broad interpretation of case-marking which includes marking by adpositions as well (flagging). There will be other phenomena, however, that are language-specific to a higher degree. To continue with the Tlapanec example, this language has a distinction between agentive and patientive verbs which cross-cuts case distinctions, and by this criterion GET and APPEAR cluster in that they are treated as patientive, while RUN and KILL cluster in that they are agentive (see Wichmann 1996). The systematic study of certain diagnostic constructions across the lexicon can also provide important insights into the properties of the constructions in question.

In explaining different syntactic preferences of arguments for certain alternations (e.g., preferential cross-coreferentiality of the recipient with the subject in reciprocal formation), the hypothesis to be pursued is the one advocated in Malchukov et al. (2010b) for the domain of ditransitive constructions, namely, that the preferences are motivated by functional (semantic and/or pragmatic considerations), as acknowledged in different frameworks (see for example Van Valin 1990; Kaufmann 1995; and Wunderlich 2006 on the semantic grounding of unaccusativity diagnostics).



Thus, there are universal patterns for syntactic characteristics, just as there are universal patterns for coding properties. Different alternations target different regions in the semantic map. General patterns emerge even where languages differ with regard to coding properties as well as with regard to the availability of alternations. Thus, Russian differs radically from English in that it does not allow an uncoded inchoative-causative alternation (= ambitransitive verbs). Instead, the Reflexive-Anticausative marker *-sja* signals intransitive uses (*lomat* ‘break (tr.)’ vs. *lomat’-sja* ‘break (intr.)’). Yet the extension of the two constructions across verb classes is comparable in the two languages. Another point of contrast between English and Russian or German (Sauerland 1994) is that alternations involving rearrangement of two arguments remain uncoded in English, while they require prefixation in the other two languages (cf. Russian *na-gruzit’ seno na telegu* [PREFIX load hay on cart] ‘to load the hay on the cart’ vs. *za-gruzit’ telegu senom* [PREFIX load cart hay.INSTR] ‘to load the cart with hay’; cf. German *laden* vs. *be-laden*). This raises an interesting question regarding consequences of structural properties for the syntactic component (cf. Hawkins 1986 concerning a trade-off between morphological case and syntactic versatility).

Thus, general patterns of valency classes can be predicted from their characteristics in terms of coding and behavioral properties. As noted above, such predictions hold insofar as the properties themselves are functionally grounded. Yet these functional motivations have their own limitations: A strategy may generalize beyond the core class to adjacent classes (e.g., extending the Allative pattern from caused motion verbs with a goal argument to beneficiaries with other verb types, as in Finnish). There may be other exceptions to the generalization which are rooted in functional, structural, and diachronic factors. As discussed in Malchukov (2005) and Malchukov (this volume), verb polysemy can be responsible for such exceptions. Thus in Lezgian (Haspelmath 1993b), the verb for ‘see’ retains its Dative-Absolutive pattern even when it is used as an active perception verb (‘look’). Another type of exception is due to structural factors. One example from the domain of ditransitive constructions, discussed in Malchukov et al. (2010a), is the use of a double object construction in Malayalam. In this language, ‘give’ takes a dative construction, while “less canonical ditransitives” like ‘entrust’ and ‘feed’ take a double object construction (Asher & Kumari 1997: 205). This goes against the general tendency for ‘give’ to be preferentially encoded in a double object construction (Kittilä 2006). This exception is due to a structural factor: The verbs occurring in the double object construction like ‘entrust’ and ‘feed’ are (lexicalized) causatives, and causatives of transitives in Malayalam regularly occur in the double object construction. Finally, as an example of a diachronic factor, consider the case of subject-experiencer constructions in ‘Standard Average European’ languages, discussed by Haspelmath (2001). Haspelmath suggests a historical explanation for the preference of subject-experiencer constructions in Standard Average European languages, noted by Bossong (1998). He shows that many emotion predicates his-

torically arise through metaphorical extension from verbs denoting a physical action; thus, *worry* derives its meaning from ‘strangle; seize by the throat’, *stun* from ‘deprive of consciousness with a blow’, etc. (Haspelmath 2001: 79). Thus, the verb has changed its meaning but retained the case frame. Diachronic explanations have also been invoked for other cases, for example, to explain idiosyncrasies of split intransitive languages, where the exceptional (minor) pattern of patientive verbs is derived historically from a transimpersonal (= transitive impersonal) construction (Malchukov 2008). At the synchronic level, however, the existence of idiosyncratic case patterns cannot be denied; these are often discussed under the label of ‘lexical case’ in the literature.

All these exceptions to the majority pattern which have plagued the research on verb classes make generalizations in lexical typology subject to an important qualification: Implicational universals in terms of verb type hierarchies or semantic maps can only be formulated in existential terms (“for some member of the class X”), not in universal terms (“for every member of the class X”). Exceptions will often be found, due to the interference of other factors. In Malchukov (2005), such interfering factors were taken into account by identifying them as competing functional motivations in the domain of case marking (or competing constraints in terms of Optimality Theory). In other cases, deviant behaviors of certain verbs may be due to their structural make-up as derived (e.g., the Malayalam case mentioned in the previous paragraph) or periphrastic. In the accompanying database structural information is included which allows identification of such cases where this plays a role. To avoid this problem, whenever possible, contributors were asked to choose simplex equivalents, cf. the use of proxies in Johanna Nichols’s work (e.g., Nichols 2008). The questionnaire also includes information on verb polysemy to control for possible polysemy effects. As for the diachronic issues, these are addressed by specialists in particular languages in the chapters devoted to those languages.

### 3.1.2 Methods and data

The general goal of the Valency Classes project was to approach the question of valency classes from a broad typological perspective, identifying general patterns in verb syntax through an empirical examination of a cross-linguistically comparable list of verbs. The procedure follows that of lexical typology as outlined by Nichols (2008), where translational equivalents of a sample of meanings are identified for all the sample languages. In actual practice, it included a lexical questionnaire containing sentences (“typical contexts”) rather than glosses, to provide more contextual support. As noted above, the characteristics of the present approach are:

- the ontology of valency classes is fairly fine-grained, like Levin's;
- the ontology is grounded semantically, rather than syntactically; only in this way can the syntactic properties be studied without circularity;
- it is limited to verb meanings which are likely to have equivalents across languages.

The list of 70 verb meanings that underlie the chapters in Part II is given in Malchukov et al. (this volume), which should be consulted for discussion of further methodological issues. It can be viewed as a compromise between Levin's classification and semantically based classifications (in particular those suggested by Dixon 1991 and Lehmann 1991). Verbs corresponding to these meanings were collected for a core sample of 36 languages from all parts of the world, for 30 of these languages leading to the chapters in Part II.

The data for this project come from language experts who responded to the questionnaire (see Malchukov et al., this volume), which elicited information on both coding properties and behavioral properties of the verbs (such as occurrence in causative, passive and applicative constructions), as well as language-specific constructions that differentiate between different valency classes.

## 3.2 Results

Other than this volume, the Valency Classes project has led to a single substantial product, the database Valency Patterns Leipzig (Hartmann et al. 2013), in addition to individual articles (Haspelmath 2014; Hartmann et al. 2014; Malchukov, *forthc.*; Wichmann *forthc.*), a special journal issue (Wichmann 2014), and an edited volume (Hellan et al. *in prep.*). The present volume and the database should be viewed as two aspects of a single enterprise aiming to document variation in valency classification across languages, which complement and inform each other. Within the present volume the database information is represented most directly (even if partially) in the summary tables accompanying individual chapters which are exported from the database and provide information about coding frames and a selection of alternations for individual verb meanings.

The results as presented in this volume are divided into three parts. Part I contains, in addition to this Introduction, other preparatory material essential to an understanding of valency classes and valency alternations cross-linguistically, in particular the Questionnaire and more detailed discussion of the cross-linguistic comparison of verbal valency and an overview of valency classes and alternations, followed by contributions outlining some of the generalizations that emerged (or that failed to be substantiated) in the course of the project.

Part II contains 30 chapters dealing with individual languages of the sample. Geographically, these languages cover all six continents that are permanently inhabited by humans, while genealogically, they belong to the following families

(following the classification of Glottolog <<http://glottolog.org/>>, consulted on 2014-Sep-30): Tuu, Mande, Atlantic-Congo, Afroasiatic; Indo-European, Nakh-Daghestanian, Tungusic, Yeniseian, Sino-Tibetan, Japonic, Ainu; Mirndi, Morehead-Wasur, Austronesian; Eskimo-Aleut, Algic, Salishan, Siouan, Uto-Aztecan, Otomanguean, Mayan, Boran, Araucanian.

Finally, part III provides general perspectives on the results of the project in their broader scientific environment from four scholars (Beth Levin, Christian Lehmann, Tasaku Tsunoda and Cliff Goddard) who have made major contributions to our understanding of valency classes and valency alternations. Christian Lehmann presents – in a nutshell – his multilevel conception of argument structure, semantic roles and valency change. Beth Levin, whose work was one of the major sources of inspiration for the present project, reviews some recent developments in her study of valency classes, highlighting the role of the manner verbs vs. result verbs dichotomy for constraining argument alternations. Tasaku Tsunoda, whose work was another major influence, supplies new evidence for the role of the Transitivity Hierarchy ('Hierarchy of Two-Place predicates') in explaining cross-linguistic variation in verbal valency. Finally, Cliff Goddard, who along with Anna Wierzbicka is one of the most prominent advocates of the Natural Semantic Metalanguage (NSM) approach (e.g., Goddard & Wierzbicka 2002), demonstrates a potential of the NSM-style analysis, relying on the radical decomposition of the verb meaning, for the study of argument alternations, and concludes with a plea for closer integration of syntactic and lexical typology in the study of verb classes and valency change.

The following are some of the main general conclusions that emerge from the project, and more specifically from this volume.

There is a certain functionally motivated comparability in distributionally defined verb classes identified through coding frames and alternations, marked or unmarked. The distinction between coding frames and alternations is not categorical because unmarked alternations may be also interpreted as multiple coding frames. Notably, those verbs which show an alternation in one language may show a pattern split in another language. Thus Japanese (Kageyama et al, this volume), Chatino (Campbell, this volume) and Even (Malchukov & Nedjalkov, this volume) permit only one of the frames allowed for load-verbs in English. Balinese (Shibatani & Artawa, this volume) by contrast, features (verb-marked) object alternations, but applicatives rather indicate a deviation from the basic construction; so constructional preferences are detected here as well.

The generalizations in this domain can take the form of implicational hierarchies (Wichmann, this volume; Malchukov, this volume). These hierarchies would be distinct for different types of alternations, as the functions of alternations determine the preferential domain of application (i.e. to which verb classes an alternation preferentially applies) (Wichmann, this volume). Many alternations show a correlation which is motivated by the role of transitivity and the transitivity hierarchy as anticipated by Tsunoda (this volume, and *passim*) (Wichmann, this vol-

ume). Verb types and coding frames can be arranged on a transitivity scale or on a transitivity map showing the relative propensity of individual verb meanings for transitive coding (see Haspelmath, this volume). Another general factor which defines relative propensity for subject alternation vs. object alternation is the result vs. manner verb dichotomy as proposed by Levin (this volume; cf. Malchukov, this volume).

Our data also supports the frequently made observation regarding a tendency towards complementarity of marked and unmarked alternations: for example, for Japanese unmarked alternations are used more restrictedly, since marked alternations are used instead. Hoocak (Hartmann, this volume), a radically head-marking language, does not feature unmarked alternations at all. On the other hand, this correlation is not deterministic: in some cases existence of an unmarked alternation does not exclude the use of the marked one; thus in Italian anticausatives and S=P labile verbs are rather in competition, which leads to aspectual differentiation (Cennamo, this volume).

The hierarchies of coding (i.e. to which verbs an alternation preferentially applies) and the hierarchies of interpretation (preferential interpretation on the part of the polysemous form or multifunctional alternation) correlate with each other (Malchukov, this volume)

Generally speaking, most languages seem to support the insight that syntactic distribution allows one to arrive at a semantic classification, moreover, this classification will be comparable across languages even though it may differ in granularity. In particular those chapters that enumerate verbs classes, such as the chapters on Icelandic, Eastern Armenian, Emai, Chintang and Ainu, are instructive in this respect. However, the structure of the language, namely availability of specific coding frames and alternations, will also determine the granularity of classification (see Malchukov, this volume). For example, Northern Tungusic languages with larger case inventories show more granularity in coding frames and valency classes, as compared to Southern Tungusic languages (Malchukov & Nedjalkov, this volume). Similarly in Xârâcùù (Moyse-Faurie, this volume), valency classes are more differentiated than in other Oceanic languages, due to grammaticalization of numerous oblique prepositions, some of which are specific to a small group of verbs.

Similarly, for the domain of alternations the relative size of verb classes depends on availability/type of alternation. For example, in Chintang (Schikowski et al., this volume), which is outstanding for its rampant lability, alternations define only a few larger classes (including the preeminent class of labile verbs) as well as a host of minor verb classes with idiosyncratic behavior. Similar issues for the valency classification arise in Bora (Seifart, this volume), where some (marked) alternations are unproductive, while others apply regularly across the lexicon. Some languages stand out as having a set of verbs with a deviant behavior, with respect to coding or alternations or both. For example, for Jaminjung (Schultze-Berndt, this volume), the distributional method revealed predicate classes corresponding to

rather language-specific (possibly culture-specific) lexical fields, such as that of manner of application of heat (participating in a causative-inchoative pattern) or direction of gaze (participating in an S=A alternation).

While application of distributional criteria (coding frames, alternations) produces a semantically coherent classification in many, if not most languages, in some languages it is less predictable. For example, in Ainu (Bugaeva, this volume), some verbs (particular verb meanings from the Leipzig list) seem to belong to the wrong class. Such inconsistencies are due to interference of structural factors (e.g., a verb may resist applicative or antipassive formation attested for a semantically related verb, if the verb itself is derived through the use of applicative or antipassive morphology). For marked alternations, the expression of the alternation may have repercussions on its applicability. If affixal, alternations may differ in productivity, while if expressed periphrastically they are generally less restrictive (and from that perspective of less interest for the issues of valency classification). At the other extreme there are languages that express valency change through stem alternation like Arabic; such types of alternation are more idiosyncratic, although one may try to achieve a more coherent classification by matching individual stem classes and controlling for the function of a particular alternation (see Kász, this volume, on Arabic, especially Appendix 1).

Some other languages seem to be more challenging for valency classification: in particular, this holds for two Malay varieties, which are otherwise quite different in their grammatical make-up: Jakarta Indonesian and Sri Lanka Malay. Jakarta Indonesian (Connors et al., this volume) shows unusual flexibility, verb meanings being “promiscuous” with respect to the coding patterns they take. Sri Lanka Malay (Nordhoff, this volume), on the other hand, does not reveal an obvious argument/adjunct distinction, so establishing verb classes beyond the transitive/intransitive distinction is problematic. Moreover, according to Nordhoff, this language hardly shows any alternations, whether marked or unmarked. As a result, application of the conventional criteria does not result in an obvious valency classification, so the role of valency as argument specification for a particular verb is also questionable (Nordhoff, this volume). This last conclusion, however, should be qualified in several respects. First, the distinction between argument and adjuncts may indeed be subtle for particular languages, so it will need additional tests (see, for example, the sophisticated diagnostics developed by Bisang for Mandarin Chinese to ascertain the status of a particular noun phrase as an argument or an adjunct; see Lu et al., this volume; cf. Bisang 2006). Second, absence of familiar types of alternations does not necessarily invalidate valency classification. Thus, Emai (Schaefer & Egbokhare, this volume), which generally lacks conventional alternations (either marked or unmarked) but makes use of valency extension (valency augmentation) through oblique arguments or through a serial verb construction, still shows a regimented division into valency classes by these distributional criteria. Third, even for languages which, like Indonesian, are exceptionally flexible in their distributional

characteristics, frequency criteria may help to differentiate between different valency classes. It is instructive that Connors et al. (this volume) note that while Jakarta Indonesian fails to distinguish between conventional valency classes, it does make a distinction between *valency preference classes*, which are “sets of lexical items which differ with respect to the ease and frequency with which they enter into constructions associated with various valency patterns”. They note further: “As evident from the accompanying database, the valency preference classes and more frequent coding frames of Jakarta Indonesian fall within the range of variation of valency classes and grammaticalized coding frames observed in most other languages.” Yet, on the whole, it is true that in some languages the notion of valency seems to play a less important role than in other languages. In some languages the distinction between transitive and intransitive verbs or between core and oblique arguments is very clear (cf. Lehmann on Yucatec Maya and Creissels on Mandinka, this volume). Also for Ket, it is noted that verbal valency is fixed (as instantiated by a particular morphological pattern) and cannot be altered through an alternation (Vajda, this volume). In other languages, as illustrated by Jakarta Indonesian and Sri Lanka Malay, valency classes are underdifferentiated. It remains to be seen what is the explanation for these differences, but it is worth noting that both Jakarta Indonesian and Sri Lanka Malay are high-contact varieties, so the absence of valency specification (such as idiosyncratic coding of arguments or general selectivity of verbs with respect of certain patterns) may well be another trait of reduced complexity found in languages exposed to intensive language contact.

With respect to alternations, the contributions to this volume have shown that in most cases these alternations can be grouped together into identifiable classes of valency-increasing, valency-reducing, etc. alternations as identified in the typological literature (see Lehmann, this volume; Malchukov, this volume, on the taxonomy of valency alternations). For example, many of the alternations identified by Levin and others for English find their parallels in Nlɪŋg (Ernszt et al., this volume), including varieties of object/oblique alternations with a conative or partitive meaning, as well as analogs of a locative alternation. As a result the valency classes to which they apply are comparable across languages. Yet in some cases, alternations are more idiosyncratic; for example, “generalized voice alternations” (in terms of Gil) in Jakarta Indonesian may change verbal valency but need not do so, even with the same verb. In many cases such voice polyfunctionality has a principled basis: that is, a polyfunctional marker may change interpretation depending on the valency type of verb to which it applies (see Malchukov, this volume, on markedness patterns in voice ambivalence). An instructive example of principled preferences for different patterns of (S=A vs. S=P) lability is found in languages such as CAY Eskimo (Miyaoaka, this volume) and Mandinka (Creissels, this volume), and is further discussed in Malchukov (this volume). For ambivalence in marked alternations, an instructive case is Bezhta (Comrie et al., this volume), where “most of the alternations (case-coded and verb-coded) have a semantic ef-

fect beyond mere valency change”. Generally, application of a morphological operation as such often yields very broad and incoherent verb classes. However, the more one takes into account the syntactic information associated with a particular morphological operation, the more coherent the valency classification becomes (see Malchukov & Nedjalkov, this volume, on Tungusic).

The factors conditioning distributional patterns of verbs (coding patterns but also uncoded alternations) are variegated, but most of them have to do with semantic roles (see Lehmann, this volume, for general analysis of features contributing to semantic roles, and the chapter on Yucatec Maya, by Lehmann, for an illustration). Moreover, the manner/result dichotomy, which is held responsible by Levin for availability of object alternations, may be also related to differences in argument structure (insofar as only a result verb may have an incremental theme as its argument). In some languages, animacy is a relevant factor; in Lehmann’s approach animacy (degree of empathy) is also a contributing factor to semantic role identification. Ojibwe (Rhodes & Valentine, this volume), which shows lexical doublets depending on whether S of an intransitive verb or P of a transitive verb is inanimate or animate, is a particularly clear example of this sort. In some languages further factors may be of importance; the role of information structure has been held responsible for the dative alternation in English, and the same has been suggested for Nlɪŋg. Finally, other factors such as verb polysemy or meaning change which may be accompanied by a persistence of the original argument structure may play a role here.

Our investigations have confirmed once again the importance of the transitive/intransitive dichotomy: The vast majority of languages, even those showing under-differentiation in other domains, display the transitive/intransitive distinction. The privileged status of transitive verbs is investigated in the chapters by Haspelmath (this volume), and Blasi (this volume) on “transitivity prominence”; in particular, the latter chapter provides statistical evidence for the claim that the transitive pattern is the single major pattern among bivalent verbs. The privileged status of transitive verbs may be observed for the domain of alternations, as well. This is brought out especially clearly for some languages like Chintang, which also shows a privileged status of transitive verbs with respect to alternations (all but two of them apply to transitive verbs exclusively). Yet cross-linguistically, this distinction can be undermined on several counts (Malchukov, this volume): at the lexical level for languages showing rampant ambitransitivity (like Chintang); at the syntactic level for languages with inherent complements (like Yorùbá).

Although it was not a focus of the volume, the languages examined differ with respect to the status of the transitive vs intransitive classes: some are intransitive based (intransitives as an open class), while others are transitive based. Ainu (Bugaeva, this volume), where most transitive verbs (except for statives) are derived (by causative and applicative affixes), belongs to the class of intransitive-based languages. Chatino (Campbell, this volume) and Mapudungun (Zuñiga, this vol-



ume) also belong to the transitivizing type. On the other hand, Nen (Evans, this volume), where most intransitives are derived through the use of (detransitivizing) middle morphology, belongs to the second class.

Another issue related to transitivity and valency in general is that different constructions appear with different frequency; whenever corpus data was available (see the chapters on NlIng, Chintang, Jakarta Indonesian) it indicates that the distinctions between valency classes are scalar rather than gradual, and even for labile verbs there is a cline showing different predisposition of different verb meanings for different kinds of lability (see in particular the chapter on Chintang). The gradualness of such distinctions becomes still more evident once one tries to generalize across languages, as captured by alternation hierarchies in Wichmann (this volume). On the other hand, as already mentioned, frequency data can not only blur somewhat the distinction between valency subclasses, but also help to discern a general picture, as in the case of Jakarta Indonesian, which otherwise shows exceptional distributional flexibility, but still allows identification of verb classes once frequency information (about preferred frames) is included.

In addition to structural factors accounting for exceptions to the functionally based hierarchies in coding and behavior, there are some deeper rooted discrepancies related to the lexical meaning of individual verbs, that is to differences in how certain events are lexicalized in particular languages. In Lehmann's approach (Lehmann, this volume; Lehmann et al. 2004), the variation here concerns the mapping between the universal referential level and the lexical/semantic level, rather than a mapping from semantics to morphosyntax. These issues pertaining more to matters of lexical typology have not been systematically investigated within the project, except for the observation that periphrastic expression of verb meaning is a frequent source of idiosyncratic coding patterns. Such cases are briefly discussed in terms of "excorporation" in Malchukov (this volume), and "apotaxis" by Haspelmath (this volume). In some cases such mismatches have a consistent character; Lehmann (this volume and *passim*; see Lehmann et al. 2004) speaks in this connection of typological strategies. A good illustration of the latter is a predisposition to encode arguments as possessors, characteristic of certain languages such as Yucatec Maya (Lehmann, this volume).

Finally, there are obviously idiosyncratic patterns of meaning which are responsible for particular deviations from the expected valency pattern. This aspect is underscored in Goddard's chapter which argues that a predictive theory should take into account the full analysis of verbal meaning (rather than only putative syntactically relevant parts of it), as well as provide an explicit semantic analysis of different constructions (including marked alternations, like the passive construction), in order to achieve a better understanding of the applicability of certain operations (constructions) to particular verb meanings. This approach, which is close to Construction Grammar approaches, needs to be pursued typologically to spell out its predictions and check these predictions against a cross-linguistic set

of data. Goddard's in-depth analysis implies that verbal behavior with respect to clustering is also more complex than suggested in the literature, as there will be constructions sensitive to other components of meaning. The predictions seem plausible in the light of the data assembled in this volume<sup>2</sup> but would need a broader follow-up study to corroborate them.

In conclusion, we would like to stress that the Leipzig Valency Classes Project is conceived as a first systematic large-scale effort to construct a comprehensive typologically informed typology of valency classes that paves the way for future cross-linguistic studies of verb classes and argument marking. Some issues could only be treated tangentially within this project, and will have to be pursued in follow-up projects. In fact, apart from the Leipzig Valency Classes Project, there are two other current projects concerned with typology of verb classes and argument marking: the Zurich-based project dealing with role clustering and variation in argument structure within the AUTOTYP research program (e.g., Bickel et al. 2014), and the St. Petersburg-based typological project primarily concerned with syntactic clustering of bivalent verbs in the languages of Europe (e.g., Say 2014). The projects differ somewhat in scope and also in certain methodological assumptions, but all share the general goal of making an empirical contribution to establishing a cross-linguistically valid typology of valency classes. Another project which should be mentioned in this context is an ongoing project dealing with argument alternations at the National Institute for Japanese Language and Linguistics (Tokyo) and coordinated by Taro Kageyama, which is mostly focused on Japanese, but also has typological extensions (see, in particular, contributions to Kageyama & Jacobsen (forthcoming), as well as an online typological database of transitivity alternations WATP edited by Prashant Pardeshi (<http://watp.ninjal.ac.jp/en/>). Clearly, some broader issues brought up in the Leipzig Valency Classes Project can only be resolved through the concerted efforts of a linguistic community involving different sub-fields of the discipline (typology, semantics, and corpus linguistics, to name a few); yet we hope that the results of the Leipzig Valency Classes Project in the form of the present volume and the ValPaL database will remain a valuable resource informing future studies of this topic.

## References

- Ágel, Vilmos (ed.). 2006. *Dependenz und Valenz: ein internationales Handbuch der zeitgenössischen Forschung*. Berlin: Mouton.
- Anderson, Lloyd B. 1986. Evidentials, paths of change, and mental maps: typologically regular asymmetries. In Wallace Chafe & Johanna Nichols (eds.), *Evidentiality: The Linguistic Encoding of Epistemology*, 273–312. Norwood: Ablex.
- Apresjan, Jurij D. 1969. *Èksperimental'noe issledovanie russkogo glagola*. Moskva: Nauka.

---

<sup>2</sup> This view is also compatible with the studies of verbal valency in the Construction Grammar tradition such as Boas (2006), as well as with findings of some corpus research (see Faulhaber

- Asher, Ronald E. & T. C. Kumari. 1997. *Malayalam*. London: Routledge.
- Bickel, Balthasar, Taras Zakharko, Lennart Bierkandt & Alena Witzlack-Makarevich. 2014. Semantic role clustering: an empirical assessment of semantic role types in non-default case assignment. *Studies in Language* 38(3). 485–511.
- Bisang, Walter. 2006. Widening the perspective: Argumenthood and syntax in Chinese, Japanese and Tagalog. In Daniel Hole, André Meinunger & Werner Abraham (eds.), *Datives and Other Cases. Between Argument Structure and Event Structure*, 331–381. Amsterdam: Benjamins.
- Blume, Kerstin. 1998. A contrastive analysis of interaction verbs with dative complements. *Linguistics* 36(2). 253–280.
- Boas, Hans C. 2006. A frame-semantic approach to identifying syntactically relevant elements of meaning. In Petra Steiner, Hans C. Boas & Stefan Schierholz (eds.), *Contrastive Studies and Valency: Studies in Honor of Hans Ulrich Boas*, 119–149. Frankfurt: Peter Lang.
- Bossong, Georg. 1998. Le marquage différentiel de l'objet dans les langues d'Europe. In Jack Feuillet (ed.), *Actance et valence*, 259–294. Berlin: Mouton de Gruyter.
- Busse, Winfried. 2006. Valenzlexika in anderen Sprachen. In Vilmos Ágel (ed.), *Dependenz und Valenz: ein internationales Handbuch der zeitgenössischen Forschung*, 1424–1435. Berlin: Mouton.
- Butt, Miriam. 2006. *Theories of Case*. Cambridge: Cambridge University Press.
- Comrie, Bernard. 2008. What is a passive? In Zarina Estrada-Fernández, Søren Wichmann, Claudine Chamoreau & Albert Álvarez González (eds.), *Studies in Voice and Transitivity*, 1–18. Munich: Lincom Europa.
- Comrie, Bernard & Helma van den Berg. 2006. Experiencer constructions in Daghestanian languages. In Ina Bornkessel, Matthias Schlesewsky, Bernard Comrie & Angela D. Friederici (eds.), *Semantic Role Universals and Argument Linking: Theoretical, Typological, and Psycholinguistic Perspectives*, 127–154. Berlin: Mouton de Gruyter.
- Croft, William. 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.
- Croft, William. 2012. *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press.
- Croft, William, Jóhanna Barðdal, Willem Hollmann, Maike Nielsen, Violeta Sotirova & Chiaki Taoka. 2001. Discriminating verb meanings: The case of transfer verbs. Handout, LAGB Autumn Meeting, Reading.
- Cysouw, Michael, Martin Haspelmath & Andrej Malchukov (eds.). 2010. Semantic maps: theory and applications. [special issue] *Linguistic Discovery* 8(1).
- De Clerck, Bernard, Timothy Coleman & Dominique Willems (eds.). 2013. Thoughts on Typologies: Different Perspectives on Verb Classifications. [special issue] *Linguistics* 51(4).
- Dixon, Robert M. W. 1991. *A New Approach to English Grammar, on Semantic Principles*. Oxford: Clarendon Press.
- Dixon, Robert M. W. 2005. *A Semantic Approach to English Grammar*. Oxford: Oxford University Press.
- Donohue, Mark & Søren Wichmann (eds.). 2008. *The Typology of Semantic Alignment*. Oxford: Oxford University Press.
- Drossard, Werner. 1991. Verbklassen. In Hansjakob Seiler & Waldfried Premper (eds.), *Partizipation: das sprachliche Erfassen von Sachverhalten*, 150–182. Tübingen: G. Narr.
- Faulhaber, Susen. 2011. *Verb Valency Patterns: A Challenge for Semantics-based Accounts*. Berlin: Mouton De Gruyter.

---

2011). Moreover, typological studies operating with a larger set of verb meanings such as Say (2014) and Bickel et al. (2014) suggest that more caution is required for generalizations in terms of semantic classes, and, further, that certain verb meanings may be more representative for individual classes than some other meanings.

- Fellbaum, Christiane (ed.). 1998. *WordNet: an Electronic Lexical Database*. Cambridge, Mass.: MIT Press.
- Fillmore, Charles J., Christopher R. Johnson & Miriam Petruck. 2003. Background to FrameNet. *International Journal of Lexicography* 16(3). 235–250.
- Fukui, Naoki, Shigeru Miyagawa & Carol Tenny. 1985. Verb classes in English and Japanese: a case study in the interaction of morphology, syntax and semantics. *Lexicon Working Papers*, 3. Center for Cognitive Science, MIT.
- Geniušienė, Emma. 1987. *The Typology of Reflexives*. Berlin: Mouton de Gruyter.
- Gerdts, Donna & Kaoru Kiyosawa. 2005. Halkomolem psych applicatives. *Studies in Language* 29(2). 329–362.
- Goddard, Cliff & Anna Wierzbicka (eds.). 2002. *Meaning and Universal Grammar – Theory and Empirical Findings. Vols I and II*. Amsterdam: John Benjamins.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.
- Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, Mass.: MIT Press.
- Haiman, John (ed.) 1985. *Iconicity in Syntax*. Amsterdam: Benjamins.
- Hartmann, Iren, Martin Haspelmath & Bradley Taylor (eds.). 2013. *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.  
<http://valpal.info>.
- Hartmann, Iren, Martin Haspelmath & Michael Cysouw. 2014. Identifying semantic role clusters and alignment types via microrole coexpression tendencies. *Studies in Language* 38(3). 463–484.
- Haspelmath, Martin. 1993a. More on the typology of inchoative/causative verb alternations. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and Transitivity* (Studies in Language Companion Series 23), 87–120. Amsterdam: Benjamins.
- Haspelmath, Martin. 1993b. *A Grammar of Lezgian*. (Mouton Grammar Library, 9). Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 1997. *Indefinite Pronouns*. (Oxford Studies in Typology and Linguistic Theory). Oxford: Clarendon Press.
- Haspelmath, Martin. 2001. Non-canonical marking of core arguments in European languages. In Alexandra Y. Aikhenvald, Robert M. W. Dixon & Masayuki Onishi (eds.), *Non-Canonical Marking of Subjects and Objects*, 53–83. Amsterdam: Benjamins.
- Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In Michael Tomasello (ed.), *The New Psychology of Language*, vol. 2, 211–243. New York: Lawrence Erlbaum Associates Publishers.
- Haspelmath, Martin. 2014. Arguments and adjuncts as language-particular syntactic categories and as comparative concepts. *Linguistic Discovery* 12(2). 3–11.
- Hawkins, John. 1986. *A Comparative Typology of English and German: Unifying the Contrasts*. Austin: University of Texas Press.
- Hellan, Lars, Michela Cennamo & Andrej Malchukov (eds.). in prep. Contrastive studies in verbal valency.
- Herbst, Thomas & Katrin Götz-Votteler (eds.). 2007. *Valency: Theoretical, Descriptive and Cognitive Issues*. Berlin/New York: de Gruyter.
- Jones, Douglas (ed.). 1994. Working Papers and Projects on Verb Class Alternations in Bangla, German, English, and Korean. *MIT AI Memo* 1517.
- Joppen-Hellwig, Sandra. 2001. *Verbklassen und Argumentlinking*. Tübingen: Niemeyer.
- Kageyama, Taro & Wesley M. Jacobsen (eds.). Forthcoming. *Transitivity and Valency Alternations: Studies on Japanese and Beyond*. Berlin: De Gruyter Mouton.
- Kaufmann, Ingrid. 1995. O- and D-predicates: a semantic approach to the unaccusative-unergative distinction. *Journal of Semantics* 12. 377–427.

- Kemmer, Suzanne. 1993. *The Middle Voice*. Amsterdam: Benjamins.
- Kipper-Schuler, Karin. 2005. *VerbNet: a broad coverage, comprehensive verb lexicon*. Ph.D. dissertation, University of Pennsylvania.
- Kittilä, Seppo. 2006. The anomaly of the verb 'give' explained by its high (formal and semantic) transitivity. *Linguistics* 44(3). 569–612.
- Kratochvíl, František, Alexander R. Coupe & Randy J. LaPolla (eds.). 2011. Studies in Transitivity: Insights from Language Documentation. [special issue] *Studies in Language* 35(3).
- Lazard, Gilbert. 1994. *L'Actance*. Paris: Presses Universitaires de France.
- Lazard, Gilbert. 1998. *Actancy*. Berlin: Mouton de Gruyter.
- Lehmann, Christian. 1991. Predicate classes and PARTICIPATION. In Hansjakob Seiler & Waldfried Premper (eds.), *Partizipation: das sprachliche Erfassen von Sachverhalten*, 183–239. Tübingen: G. Narr.
- Lehmann, Christian, Yong-Min Shin & Elisabeth Verhoeven. 2004. *Direkte und indirekte Partizipation. Zur Typologie der sprachlichen Repräsentation konzeptueller Relationen*, 2<sup>nd</sup> revised edn. Erfurt: Seminar für Sprachwissenschaft der Universität (ASSidUE, 13).
- Levin, Beth. 1993. *English Verb Classes and Alternations*. Chicago: University of Chicago Press.
- Levin, Beth & Malka Rappaport Hovav. 2005. *Argument Realization*. Cambridge: Cambridge University Press.
- Malchukov, Andrej. 1995. *Even* (Languages of the World/Materials, vol. 12). München: Lincom Europa.
- Malchukov, Andrej. 2005. Case pattern splits, verb types, and construction competition. In Mengistu Amberber & Helen de Hoop (eds.), *Competition and Variation in Natural Languages: The Case for Case*, 73–117. Amsterdam: Elsevier.
- Malchukov, Andrej. 2008. Split intransitives, experiencer objects and 'transimpersonal' constructions: (re-) establishing the connection. In Mark Donohue & Søren Wichmann (eds.), *The Typology of Semantic Alignment*, 76–101. Oxford: Oxford University Press.
- Malchukov, Andrej. Forthcoming. "Ambivalent voice": markedness effects in valency change. In Taro Kageyama & Wesley M. Jacobsen (eds.), *Transitivity and Valency Alternations: Studies on Japanese and Beyond*. Berlin: De Gruyter Mouton.
- Malchukov, Andrej, Martin Haspelmath & Bernard Comrie (eds.). 2010a. *Studies in Ditransitive Constructions: A Comparative Handbook*. Berlin: De Gruyter Mouton.
- Malchukov, Andrej, Martin Haspelmath & Bernard Comrie. 2010b. Ditransitive construction: a typological overview. In Andrej Malchukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in Ditransitive Constructions: A Comparative Handbook*, 1–64. Berlin: De Gruyter Mouton.
- Nichols, Johanna. 2008. Why are stative-active languages rare in Eurasia? A typological perspective on split-subject marking. In Mark Donohue & Søren Wichmann (eds.), *The Typology of Semantic Alignment*, 121–139. Oxford: Oxford University Press.
- Nichols, Johanna, David A. Peterson & Jonathan Barnes. 2004. Transitivity and detransitivizing languages. *Linguistic Typology* 8(2). 149–211.
- Pesetsky, David. 1995. *Zero Syntax*. Cambridge, Massachusetts: The MIT Press.
- Pinker, Steven. 1989. *Learnability and Cognition. The Acquisition of Argument Structure*. Cambridge, Massachusetts: The MIT Press.
- Say, Sergej. 2014. Bivalent verb classes in the languages of Europe: A quantitative typological study. *Language Dynamics and Change* 4. 116–166.
- Sauerland, Ulrich. 1994. German diathesis and verb morphology. In Douglas Jones (ed.), *Working Papers and Projects on Verb Class Alternations in Bangla, German, English and Korean*, 37–92. MIT AI Memo 1517.
- Schulte im Walde, Sabine. 2003. *Experiments on the automatic induction of German semantic verb classes*. PhD Thesis. University of Stuttgart.

- Schumacher, Helmut. 2006a. Deutschsprachige Valenzwörterbücher. In Vilmos Ágel (ed.), *Dependenz und Valenz: ein internationales Handbuch der zeitgenössischen Forschung*, 1396–1424. Berlin: Mouton.
- Schumacher, Helmut. 2006b. Kontrastive zweisprachige Valenzwörterbücher. In Vilmos Ágel (ed.), *Dependenz und Valenz: ein internationales Handbuch der zeitgenössischen Forschung*, 1435–1447. Berlin: Mouton.
- Somers, Harold L. 1987. *Valency and Case in Computational Linguistics*. Edinburgh: Edinburgh University Press.
- Stiebels, Barbara. 2000. Linker inventories, linking splits and lexical economy. In Barbara Stiebels & Dieter Wunderlich (eds.), *Lexicon in Focus*, 211–245. Berlin: Akademie Verlag.
- Tsunoda, Tasaku. 1981. Split case-marking in verb types and tense/aspect/mood. *Linguistics* 19. 389–438.
- Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of Linguistics* 21. 385–396.
- Van Valin, Robert D. 1990. Semantic parameters of split intransitivity. *Language* 66(2). 221–260.
- Van Valin, Robert D. 2001. *An Introduction to Syntax*. New York: Cambridge University Press.
- Van Valin, Robert D. & Randy Lapolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.
- Vogel, Alan R. 2003. *Jarawara verb classes*. Ph.D. dissertation, University of Pittsburgh.
- Wälchli, Bernhard & Michael Cysouw. 2012. Lexical typology through similarity semantics: Toward a semantic map of motion verbs. *Linguistics* 50(3). 671–710.
- Wichmann, Søren. 1996. The degrammaticalization of agentivity in Tlapanec. In Elisabeth Engberg-Pedersen, Michael Fortescue, Peter Harder, Lars Heltoft & Lisbeth Falster Jakobsen (eds.), *Expression and Structure: Studies in Danish Functional Grammar* (Studies in Language Companion Series), 343–360. Amsterdam/ Philadelphia: John Benjamins.
- Wichmann, Søren. 2009. Case relations in a head-marking language: verb-marked cases in Tlapanec. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford Handbook of Case*, 797–807. Oxford: Oxford University Press.
- Wichmann, Søren (ed.). 2014. Arguments and Adjuncts Cross-Linguistically. [Special issue]. *Linguistic Discovery* 12(2).
- Wichmann, Søren. Forthcoming. Quantitative tests of implicational verb hierarchies. In Taro Kageyama & Wesley M. Jacobsen (eds.), *Transitivity and Valency Alternations: Studies on Japanese and Beyond*. Berlin: De Gruyter Mouton.
- Wunderlich, Dieter. 2006. Towards a structural typology of verb classes. In Dieter Wunderlich (ed.), *Advances in the Theory of the Lexicon*, 57–166. Berlin: Mouton de Gruyter.

## 2 Leipzig Questionnaire on valency classes

### Preliminaries

The present questionnaire was compiled by Andrej Malchukov with feedback from the other Leipzig Valency Classes Project members (Bernard Comrie, Martin Haspelmath, Iren Hartmann and Søren Wichmann) at an early stage of the project and distributed to contributors to the volume. Since 2010 the conception of the project as reflected in the questionnaire has undergone certain changes, as described in the Database Manual (Haspelmath & Hartmann 2013, available at: [http://www.eva.mpg.de/lingua/valency/files/database\\_manual.php](http://www.eva.mpg.de/lingua/valency/files/database_manual.php)). The Database Manual can be seen as a follow-up to the questionnaire but is intended for contributions to the database rather than for book chapters. Moreover, it differs somewhat in scope (see in particular the advanced part of the Leipzig Questionnaire, which addresses a broader set of questions not implemented in the database) and also provides the database contributors with a succinct introduction to the framework and terminological conventions used in the project. The present questionnaire does not pursue this latter goal, which is largely fulfilled within the present volume by the chapter by Haspelmath and Hartmann (this volume). Thus, the Leipzig Questionnaire largely follows the original format reflecting its use to guide contributions to the volume rather than database contributions, except for one important update. It includes an updated list of 70 core meanings rather than the original list of 64 verbs. The definitive list was agreed on the basis of the input from the project participants, as well as the feedback from the contributors.

### Introduction

The present questionnaire deals with a typology of valency classes, or verb types, in terms of Levin (1993). Levin (1993) is a seminal study of syntactic classes of verbs in English, which shows that a semantic classification of verbs can be achieved through applying syntactic diagnostics. Yet, this study, as well as an earlier study by Apresjan (1969) on Russian, has not been followed up cross-linguistically, which leaves open the question of which aspects of these classifications are universal and which are language particular. Similarly, valency dictionaries are few in number and mostly deal with European languages, thus they cannot fill the gap. The questionnaire has been compiled by participants of the DFG funded project on valency classes<sup>1</sup> and is designed to obtain a consistent set of data from a representative set

---

<sup>1</sup> See <http://email.eva.mpg.de/~haspelmt/ValencyClasses.pdf> for the project description.

Tab. 1: The 70 verb meanings.

meaning label	role frame	typical context
RAIN	(it) rains	It rained yesterday.
BE DRY	S is dry	The ground is dry.
BURN	S burns	The house is burning.
SINK	S sinks	The boat sank.
ROLL	A rolls	The ball is rolling.
BE A HUNTER	S is a hunter	This man is a hunter.
BE HUNGRY	E is hungry	The baby is hungry.
BE SAD	E is sad	The little girl was sad.
DIE	S dies	The snake died.
FEEL COLD	S is cold	I'm cold.
FEEL PAIN	E feels pain in X	My arm is hurting. = I'm feeling pain in my arm.
SCREAM	S screams	The man screamed.
LAUGH	S laughs	The little girl laughed.
PLAY	S plays	The child is playing.
LIVE	S lives somewhere (L)	The old people live in town.
LEAVE	A left L	The boy left the village.
GO	S goes somewhere (L)	The woman went to the market.
SING	S sings	The boy sang (a song).
JUMP	A jumps	The girl jumped.
SIT DOWN	S sits down (somewhere (L))	The children sat down on the bench.
SIT	S sits somewhere (L)	The children sat on the floor.
RUN	A runs	The horse is running.
CLIMB	A climbs (up L)	The men climbed (up) the tree.
COUGH	S coughs	The old man coughed.
BLINK	S blinks	I blinked (my eyes).
SHAVE	A shaves (his beard/hair)	The man shaved his beard/cut his hair
DRESS	A dresses P	The mother dressed her daughter
WASH	A washes P	The mother washed the baby.
EAT	A eats P	The boy ate the fruit.
HELP	A helps X	I helped the boys.
FOLLOW	A follows X	The boys followed the girls.
MEET	A meets X	The men met the boys.
HUG	A hugs P	The mother hugged her little boy.
SEARCH FOR	A searches for X	The men searched for the women.
THINK	A thinks about X	The girl thought about her grandmother yesterday.
KNOW	A knows P	The girl knew the boy.
LIKE	E likes X	The boy liked his new toy.
FEAR	E fears X	The man feared the bear.
FRIGHTEN	A frightens P	The bear frightened the man.
SMELL	E smells X	The bear smelled the boy.
LOOK AT	A looks at P	The boy looked at the girl.
SEE	E sees X	The man saw the bear.
TALK	A talks (to X) (about Y)	The girl talked to the boy about her dog.



meaning label	role frame	typical context
ASK FOR	A asks (X) for Y	The boy asked his parents for money.
SHOUT AT	A shouts at X	The woman shouted at the children.
TELL	A tells (X) Y	The girl told the boy a funny story.
SAY	A says “...” (to X)	They said “no” to me.
NAME	A name X (a) Y	The parents called the baby Anna.
BUILD	A builds P (out of X)	The men built a house of wood.
BREAK	A breaks P (with I)	The boy broke the window with a stone.
KILL	A kills P (with I)	The man killed his enemy with a club.
BEAT	A beats P (with I)	The boy beat the snake with a stick.
HIT	A hits P (with I)	The boy hit the snake with a stick.
TOUCH	A touches P (with I)	The boy touched the snake with a stick.
CUT	A cuts P (with I)	The woman cut the bread with a sharp knife.
TAKE	A takes P (from X)	The man took the money from his friend.
TEAR	A tears P (from X)	The girl tore the page from the book.
PEEL	A peels (X off) P	The boy peeled the bark off the stick.
HIDE	A hides T (from X)	The boy hid the frog from his mother.
SHOW	A shows T (to R)	The girls showed pictures to the teacher.
GIVE	A gives T to R	We gave the books to the children.
SEND	A sends T (to X)	The girl sent flowers to her grandmother.
CARRY	A carries T (to X)	The men carried the boxes to the market.
THROW	A throws T somewhere (L)	The boy threw the ball into the window.
TIE	A ties P (to L) (with I)	The man tied the horse with a rope to the tree.
PUT	A puts T somewhere (L)	I put the cup onto the table.
POUR	A pours T somewhere (L)	The man poured water into the glass.
COVER	A covers P (with X)	The woman covered the boy with a blanket.
FILL	A fills P (with X)	The girl filled the glass with water.
LOAD	A loads T (onto L)	The farmer loaded hay onto the truck. = The farmer loaded the truck with hay.

of languages to be described in contributions to the edited volume. It starts with the study of a list of 70 verb meanings (Vs, for short) taken as representative of the verbal lexicon, as well as Levin’s taxonomy (but see below, in particular §6)<sup>2</sup>.

Since Vs might allow for different uses, the meanings are narrowed down through the use of example sentences to be translated into the respective languages (see the reference to ‘typical contexts’ in Table 1 above). The subsequent parts of the questionnaire address coding and syntactic properties of Vs. It begins with questions about coding properties in constructions formed by a V (in particular, case-marking of arguments) in order to determine the basic valency pattern. In

<sup>2</sup> This study focuses on lexical verbs rather than auxiliary verbs (with modal, aspectual and other uses).

accordance with a conventional usage the valency of a V is understood here as the list of its arguments with their coding properties (referred to as **coding frames**; see Haspelmath & Hartmann this volume, for definitions of terms). §3 deals with case alternations, which do not involve voice morphology (cf. different ‘transformations’ used to cross-classify the English verbal lexicon in Levin’s work). §4 addresses diathetic alternations (or verb-coded alternations), asking about availability of certain valency operations (like passives and applicatives) for Vs, as well as for the meanings expressed by the valency changing markers with Vs, in case they are polysemous. §3 and §4 can be seen as largely complementary, as what is a case alternation in one language will be coded as a diathetic alternation in another language (for example, many case alternations in English will be expressed by different valency operations in languages with richer morphology). The last (advanced) part of the questionnaire explores to what extent Vs are representative of lexical classes, i.e., which other verbs belong to the same valency class. This latter part cannot be fully reflected in individual contributions to the volume, which due to size limitations will just offer a summary of verb taxonomies starting from these 70 verb meanings.

## I Basic Questionnaire

### 1 Valency patterns basic examples

---

Please provide glossed examples of sentences containing the relevant Vs (see the prompt typical contexts in the Table 1 above).

---

The examples sentences exemplifying typical contexts are intended to elicit the verb meanings (Vs) introduced above. The author of a chapter is asked to provide either translational equivalents of the sentences above or other comparable constructions found with the Vs (possibly extracted from corpora). In either case, of special interest is the coding of verbal arguments, for this reason the arguments should be overtly expressed (at least in dependent-marking languages). For head-marking languages, constructions with pronominal (1<sup>st</sup>/2<sup>nd</sup> person) subjects and objects should also be considered, as 3<sup>rd</sup> person arguments are often not indexed on the verb. Alternatively the corresponding constructions with pronominal arguments should be described in §2.2 under indexing.

Examples might be somewhat modified to reflect cultural realities. In cases where a V has different translational equivalents, please choose the verbal lexeme which is more basic (i.e. more frequent and/or morphologically less complex); in the case of several basic items, please include all.

The following sections will provide further explanation of the glossed examples with respect to coding and behavioral properties of Vs.

## 2 Coding properties of valency patterns

Coding properties involve the following techniques (Haspelmath 2005; Malchukov et al. 2010):

- flagging (case or adposition marking)
- indexing (agreement, cross-referencing)
- word order (in the absence of other kinds of marking)

### 2.1 Flagging

---

How are the arguments of the verbs flagged (by a case or adposition)?

---

NB degree of differentiation will naturally depend on the number of cases available. Thus, German makes finer coding distinctions among verb classes than English, which is due to the availability of verb-specific dative and genitive case selection in German (Sauerland 1994).

**Further questions about flagging**, which may be addressed in case they interact in an interesting way with verb classification.

---

Does flagging differ for different kinds of nominals (animate/inanimate, definite/indefinite)?

---

Some languages show variation in case marking depending on nominal features such as animacy and definiteness; most commonly it has been observed for (direct) objects (cf. differential object marking in languages like Hindi; where P is marked if human, or definite in the case of inanimates). Other arguments may also differ in case marking depending on nominal features; cf. different cases for animate/inanimate locations, as in Dyirbal.

---

Do free pronouns show the same valency pattern? (Bound pronouns are considered in § 2.2).

---

This need not be the case, as witnessed by split-ergative languages of the Australian type.

---

What are other relevant factors affecting argument marking here?

---

In some languages, alignment patterns further depend on TAM-features, as is familiar from split-ergative languages like Hindi and Newari which have ergative

alignment in the perfective/past and accusative in the imperfective. Such alternations are relevant for the project to the extent they are sensitive to verb classes. Thus, in Newari, availability of an ergative pattern depends on tense, on the one hand, and on lexical class of the verb, on the other hand, so that 2-argument verbs deviating from the transitive prototype can take this pattern optionally.

## 2.2 Indexing

---

How are the arguments of Vs indexed (by agreement/cross-referencing)?

---

Some languages (head-marking languages), may distinguish valency patterns through indexing rather than flagging. Thus, Tlapanec has 4 different patterns: ergative vs. absolutive indexing patterns are used for canonical transitives, and the ‘pegative’-dative alternation is used for less canonical transitives (like ‘fear’ and ‘meet’). Note that the discussion of indexing (conjugation) markers that also signal voice distinctions (like the middle voice in Greek) should be postponed to §4 dealing with diathetic alternations.

**Further questions about indexing;** which need to be addressed to the extent these patterns reveal verb classification.

---

Does indexing depend on the features of the nominal (see above)?

---

Indexing, like flagging, may depend on nominal features. For example, in many languages with object agreement only prominent (animate/definite) objects are indexed. Again, such cases will be relevant for our project insofar as these features further interact with the verb type. For example, in the Austronesian language Manam some experiencer verbs (‘like’, ‘know’, ‘be bad at’) use object indexing only when the object is prominent, while canonical transitives (like ‘break’) invariably index the object. Thus, here we observe effects of differential object marking (i.e. differential object indexing) for verb-types deviating from canonical transitives.

## 2.3 Word order

---

What are the word order patterns associated with Vs?

---

Most often word order depends on syntactic transitivity, but some languages make further distinctions depending on the valency class. Thus, in (Gao) Songhai, canonical transitives (‘break’, ‘kill’), have SOV order, while less canonical transitives (‘see’, ‘follow’, ‘love’), have SVO order. Questions of word order need to be ad-

dressed to the extent word order interacts in an interesting way with verb classification.

### 3 Argument alternations

In this section only argument alternations (or uncoded case alternations) are considered; discussion of verb-marked diathetic alternation is postponed to the next section. The two sections should be seen as largely complementary, as alternations will be coded by dedicated markers in some languages with a richer morphology and left uncoded in other languages (like English).

---

Do Vs allow for an alternative construction (valency pattern)?

---

For example, for English, Levin (1993) mentions, in particular, the following alternations. (The list below mentions only fairly productive alternations; and does not include verb-coded diathetic alternations like the passive alternation):

- a) the inchoative-causative alternation (*John broke the stick ~ The stick broke*)
- b) the “middle” alternation (*John cut the bread ~ The bread cuts easily*)
- c) the reflexive deletion (*John washed himself ~ John washed*)
- d) the reciprocal transformation (*John married Mary ~ John and Mary married*)
- e) the dative alternation (*Mary gave the book to John ~ gave John the book*)
- f) the locative alternation (*John loaded the truck with hay ~ the hay onto the truck*)
- g) the conative alternation (*John cut the bread ~ cut at the bread*)
- h) the object deletion alternation (*John ate the bread ~ John ate*).
- i) the preposition dropping alternation (*John climbed up the hill ~ climbed the hill*).

Given that Levin’s list includes a number of other more lexically restricted alternations, it is clear that alternation types need to be generalized before they can be applied cross-linguistically.

We will distinguish between the following general types of case alternations, which also find equivalents among voice-alternations to be considered in the next section:

- a) subject-demoting/deleting case alternations

This type will include subject-demoting alternations frequently discussed under the heading of differential subject marking (e.g., genitive of negation in Russian, or ergative/oblique alternation related to volitionality in some ergative languages), but also subject-deleting alternations, as in the case of S/P labile verbs (like *break*; see Levin’s “inchoative-causative alternation”).

## b) object-demoting/deleting case alternations

Object-demoting alternations refer, in particular, to varieties of differential object marking (e.g., asymmetric alternations such as the ACC/NOM alternation in Hindi, but also symmetric ones, such as the ACC/PART alternation in Finnish). Object-deleting alternations will refer, in particular, to A/S-labile verbs (like *eat*; see Levin's "object deletion alternation").

## c) object rearranging case alternations

These include, in particular, varieties of dative and locative alternations (cf. (e, f) above), also found in other languages.

So the first question to be addressed is:

---

What are the major varieties of case alternations in your language (subject-demoting/deleting; object-demoting/deleting; object rearranging)?

---

Of course, there may be further varieties of argument alternations, not listed above (for example, object incorporation). These varieties, as well as the alternations listed above, are relevant to the project to the extent they are sensitive to verb classification (e.g., some varieties of differential object-marking apply to any transitive verb, which does not yield an interesting clustering of verb types). More generally for our purposes most relevant are those alternations which are fairly productive (not restricted to a few lexical items), but – most importantly – are sensitive to lexical classes. That is, we are interested in alternations which are distinctive for the verbal lexicon (as sampled here) rather than in those which apply across the board or apply to just few items.

After delimiting in this way the set of most relevant case-alternations in your language, the question is to be addressed is:

---

To which Vs in your language do these alternations pertain?

---

For example, if your language features labile verbs, which of the verbs from the list are labile: S/P labile (cf. the causative-inchoative alternation), and S/A labile (cf. the unspecified object-deleting transformation)?

## 4 Diathetic alternations and valency changing operations

As mentioned above, languages with richer morphology use diathetic alternations for many argument alternations left uncoded in English. Thus in Even, a Tungusic language, the "middle alternation" is signaled by the mediopassive marker, the "inchoative-causative alternation" is signaled by the causative marker (in competi-

tion with the mediopassive), while equivalents of English verbs allowing for a “reciprocal alternation” commonly involve a lexicalized sociative marker (e.g. *baka-lda* [find-SOC] ‘meet’).

It is convenient to use the same (or similar) taxonomy for the domain of verb-coded diathetic alternations, as we adopted for (uncoded) case alternations. Also in this case we will distinguish between the following types:

- a) Subject demoting/deleting (anticausatives, etc.)
- b) Subject-Object rearranging (passives)
- c) Object demoting/deleting (antipassives)
- d) Object-Object rearranging (applicatives)

In addition we address valency increasing alternations:

- e) Subject addition (causatives)
- f) Object addition (applicatives)

It should be noted that in some cases the distinction between these subvarieties may be problematic (especially between subvarieties of valency rearranging vs. valency increasing applicatives). It is also convenient to treat Subject-Object rearranging operations (passives) together with subject demoting/deleting (anticausatives; reflexives, etc.), as they frequently employ the same markers. From this perspective, the major distinction will be between valency reducing vs. valency increasing diathetic alternations.

Obviously, the set of valency/voice markers varies across languages; some of these languages distinguish between several such markers (e.g., anticausatives vs. reflexives), while other languages use the same polyfunctional marker. For such cases it is important to state both availability and the meaning of particular markers for certain Vs.

## 4.1 Valency-reducing operations

Valency reducing operations come in several subtypes, as illustrated below:

- a) Subject demoting/deleting voice alternations
  - anticausative (cf. a) in § 3 above)
  - middle (cf. b) in § 3 above)
  - reflexive (cf. c) in § 3 above)
  - reciprocal (cf. d) in § 3 above)
- b) Object demoting/deleting voice alternations
  - antipassive (cf. g), h) in § 3 above)
- c) Subject-Object rearranging voice alternations
  - passive (differs from anticausative in that A may be expressed, or is implied)

---

Does your language have the voice alternations listed above? If not, how are these functions expressed (this information should be given in § 3 if the functions are expressed through argument alternations). If so, what Vs do these markers apply to?

---



---

What are the functions of these markers when applied to different Vs?

---

For example, in Russian the “reflexive” suffix *-sja* can be used in a reflexive function with some verbs (*myt’-sja* ‘wash’), while with other verbs it has anticausative (*slomat’-sja* ‘break’), passive (*stroit’-sja* ‘be built’), or reciprocal (*vstrečat’-sja* ‘meet’) function.

## 4.2 Valency-increasing and valency-rearranging operations

The most important valency-increasing operations are causatives and applicatives; the latter however may rearrange rather than increase the valency. In some languages, the same marker is used both in causative and applicative functions; also for such cases it is important to determine which Vs select for which function.

---

If your language has causative markers, which Vs can they apply to? Does the meaning of the causative marker differ with the V involved?

---

Note that some languages have several causative markers, for example, for building intransitive vs. transitive causatives. These can be used to test for transitivity of less prototypical transitive verbs.

---

If your language has applicative markers, which Vs can they apply to? Does it have several applicative markers used with different Vs? Does the meaning of the applicative marker differ with the V involved?

---

There may be several subtypes of applicatives, depending on which object is promoted (for example, in Hoocak (Siouan), there are 4 different applicative markers, including the benefactive applicative, the instrumental applicative and two types of locative applicatives). On the other hand, the general applicative in Salish has been claimed to have different meanings depending on the verb’s class. Applicatives may be used to render many of the alternations listed in § 3, including the dative (cf. e) in § 3 above), locative (cf. f)), but also preposition dropping (cf. i)). Other languages may use directional markers to code some of these alternations (cf. Russian *na-gruzit’ seno na telegu* [PREF-load hay on cart] ‘to load the hay on the cart’ vs. *za-gruzit’ telegu senom* [PREF-load cart hay.INSTR] ‘to load the cart with hay’; German *laden* vs. *be-laden*).



### 4.3 Other valency/voice categories

Of course, it is impossible to foresee (let alone, list here) all language-particular voice categories. While choosing to address certain voice alternations, one should again be reminded that voice constructions will be relevant to the project to the extent that they interact in an interesting way with the verb lexicon (in particular, are neither restricted to few lexical items, nor apply across the board to all verbs).

---

Does your language have other voice categories? Which Vs do they apply to?

---

For example, some (Austronesian) languages show a variety of “voice” (or “focus”) forms (“actor focus”, “goal focus”, etc), used for ‘promotion’ of different objects to the subject position; for these languages it will be relevant which Vs allow for which voice constructions. On the other hand, head-marking languages of the “hierarchical type” show a direct-inverse alternation triggered by the relative prominence of the A and P arguments. In that case it is relevant to study the use of direct-inverse alternations with different groups of two and three argument verbs (in the latter case, it is also relevant which of the object arguments takes part in the alternation; e.g., Theme or Recipient of a ditransitive verb). But also for the domain of monotonitives some languages may show further differentiation; e.g., some languages (like Tlapanec) may have different inverse forms for different subtypes of 2-argument verbs.

## II Advanced Questionnaire

### 5 Further properties of individual verbs

#### 5.1 Morphological issues: complexity

---

Indicate which of the Vs are morphologically complex?

---

In §1 the contributors were prompted to select for the basic (nonderived) equivalent of verbs on the list. In some cases, however, this is impossible, as in the case where all ditransitives including GIVE are derived (e.g. applicative, as in Tzotzil). Therefore it is important to provide information about morphological complexity of Vs. This question is relevant insofar as morphological make-up may determine availability of a certain valency pattern. For example, in Malayalam, only derived ditransitives (causatives of transitives), take a double object construction, while basic (underived) ditransitives take a dative construction.

## 5.2 Semantic issues: polysemy

---

Do Vs have other meanings?

---

This question is relevant insofar as the valency pattern may be motivated through one of the meanings of the polyfunctional item. For example, in some languages, which use the same verbs for both ‘hit’ and ‘throw’, this verb follows the allative pattern, as expected for caused motion verbs.

## 5.3 Semantic issues: etymology

---

Do you know etymology of Vs?

---

The motivation for this question is the same as for the previous one: a verb may inherit the valency pattern from its original meaning.

# 6 Further properties of lexical classes

## 6.1 Lexical issues: open and closed valency classes

---

Which other verbs belong to the same valency pattern as individual Vs? Is it an open or a closed class? For an open class, please specify which verbs belong to this class (in terms of relevant semantic or formal features). For a closed class, please list other verbs in this class.

---

For example, if some of the Vs are labile (see §3 above), please give the list of other (S/P and S/A) labile verbs.

---

Do other verbs, semantically similar to a V, participate in the same alternations as this V?

---



---

Do other verbs, semantically similar to a V, show the same diathetic alternations as this V?

---

## References

- Apresjan, Jurij D. 1969. *Ėksperimental'noe issledovanie russkogo glagola*. Moskva: Nauka.
- Haspelmath, Martin. 2005. Ditransitive Constructions: The Verb ‘Give’. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 426–429. Oxford: Oxford University Press.
- Levin, Beth. 1993. *English Verb Classes and Alternations*. Chicago: University of Chicago Press.

- Malchukov, Andrej, Martin Haspelmath & Bernard Comrie. 2010. Ditransitive constructions: A typological overview. In Andrej Malchukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in Ditransitive Constructions*, 1–65. Berlin: Mouton de Gruyter.
- Sauerland, Ulrich. 1994. German diathesis and verb morphology. In Douglas A. Jones (ed.), *Working Papers and Projects on Verb Class Alternations in Bangla, German, English, and Korean*, 37–92. MIT AI Memo 1517.



### 3 Comparing verbal valency across languages

#### 1 Introduction

That different verbs may have different valencies even when they are semantically similar has long been well known (at least since Tesnière 1959), and it is the reason why dictionaries contain (or should contain) valency information for each verb. For example, we need to specify that English *wait* takes a *for*-complement, while *expect* takes a direct object.

That “the same” verbs across languages also often differ in valency is not quite so well known, though typological-comparative and contrastive works in linguistics have often discussed valency mismatches of the type illustrated in (1)–(2).

- (1) a. English  
*I<sub>1</sub> miss you<sub>2</sub>.*
- b. French  
*Vous<sub>2</sub> me<sub>1</sub> manquez.*  
‘I miss you.’ (Literally something like ‘You are missing to me.’)  
(Tesnière 1959: § 123.2)
- (2) a. English  
*She<sub>1</sub> filled the container<sub>2</sub> with water<sub>3</sub>.*
- b. Chintang (Tibeto-Burman, Nepal)  
*Huĩsa-ŋa<sub>1</sub> cuwa-Ø<sub>3</sub> gagri-be<sub>2</sub> phatt-e.*  
she-ERG water-NOM container-LOC fill-PST  
‘She filled the container with water.’  
(Literally ‘She filled the water into the container.’)  
(Bickel et al. 2010: 387)

Tesnière called such cross-linguistic mismatches “metataxis” (Tesnière 1959; cf. Koch 1994), and it is such differences between languages that we are primarily interested in here.

But what does it mean for the valencies of two verbs to fail to match? How can we compare valencies across languages? These are not trivial questions, but answers to them are a prerequisite for any comparative or contrastive research on valency. This chapter will explain how we approached this problem, and what decisions we took for the Valency Patterns Leipzig (ValPaL) database (Hartmann et al. 2013).

The basic principle is that we decided to match verbs across languages on the basis of verb meaning, and to match arguments across languages on the basis of

individual argument meaning, i.e. argument microroles (see §5, and also Hartmann et al. 2014).

## 2 Valency: coding frames and role frames

The **valency** of a verb is the range of syntactic properties of other elements of the clause that depend on the particular choice of verb, i.e. that are verb-specific. These other elements of the clause are called **arguments**. The most salient argument properties are the **coding properties**, i.e. flagging (case or adpositional marking) and indexing (i.e. bound person marking associated with the verb).<sup>1</sup> Examples of argument-coding elements are Nominative and Accusative case and the preposition *of* in (3) from English, and Dative and Nominative case as well as 3<sup>rd</sup> person singular Subject indexing in (4) from German.

(3) English

[*They*]<sub>NOM</sub> *accused* [*her*]<sub>ACC</sub> [*of plagiarism*].

(4) German

[*Den*        *Kindern*]    *gefällt*    [*der*        *Schneemann*].  
the.PL.DAT child.PL.DAT please.3SG the.SG.NOM snowman.SG.NOM  
'The children like the snowman.'

Everything else in the clause is independent of the verb: A clause can contain locational and temporal setting adverbials (e.g. 5), or it can contain manner adverbials or illocutionary adverbials (e.g. 6), regardless of the kind of verb that is chosen. Such verb-independent elements are called **adjuncts**.

(5) ***Last week** they accused her of plagiarism **in her school**.*

(6) ***In all frankness**, they accused her **with verve**.*

A clause can contain subordinators or diverse particles, it can exhibit special word order, and it can occur in different tenses (all illustrated by (7) from German), independently of its verb.

(7) ***weil**        *der Schneemann*        *dem Kind*        **doch**    *gefiel**  
because the snowman.SG.NOM the child.SG.DAT after.all please.PST.3SG  
'since the child liked the snowman after all'

---

<sup>1</sup> See Haspelmath (2013) for the term "person indexing" and its relation to traditional terms like "agreement" and "bound pronoun".



of greatest interest for the comparative study of valency, because they also tend to show the greatest cross-linguistic variability.

In addition to coding properties, verbs may also determine other properties of their arguments, such as the ability to be the antecedent of a reflexive pronoun, to be the target of omission under coreference, or the ability to be passivized. These properties are generally called **behavioural properties**. Arguments that are coded in the same way may nevertheless show different behaviours. For example, German *wiegen* ‘have a weight of (a measurement)’ takes an Accusative argument that cannot be passivized (cf. 11a–b), while *wiegen* ‘determine the weight of (an object)’ takes an Accusative argument that can be passivized (cf. 12a–b).

- (11) a. *Der Sand wiegt einen Zentner.*  
           the sand weighs one hundredweight  
           ‘The sand has the weight of one hundredweight.’  
           (passive)  
       b. \**Ein Zentner wurde von dem Sand gewogen.*  
           one hundredweight was by the sand weighed
- (12) a. *Frau Müller wiegt den Sand.*  
           Ms. Müller weighs the sand  
           ‘Ms. Müller determines the weight of the sand.’  
           (passive)  
       b. *Der Sand wird von Frau Müller gewogen.*  
           the sand is by Ms. Müller weighed  
           ‘The weight of the sand is determined by Ms. Müller.’

Thus, one can say that the two homonymous verbs *wiegen* have two different **valency frames**, perhaps notated as <Arg1/Subject V.SUBJ[1] Arg2/Extent> and <Arg1/Subject V.SUBJ[1] Arg2/Object>, respectively. In other words, we might use **syntactic-function** labels like **subject**, **object** and **extent** in valency frames, thus taking into account not only coding properties of arguments, but also behavioural properties. However, in our comparative study of valency, we have not been able to take these other properties into account systematically, so we generally use the label **coding frame** rather than **valency frame**. (Word order is intermediate between coding properties and behavioural properties; see § 11 below on word order.)

### 3 Notes on terminology

Before we go on to compare languages, we offer a few notes comparing linguistic traditions, especially with respect to terminology. The term “valency” was intro-



duced by Tesnière (1959) and has been used in English at least since Svartvik (1966), but other terms have been widely used as well.<sup>2</sup> The terms in (13) all have basically the same meaning. A terminological distinction between the more abstract “valency” and the more concrete “valency frame” is sometimes made, but is often redundant.

- |      |                       |  |
|------|-----------------------|--|
| (13) | a. complementation    | e.g. Quirk et al. (1985: 1069–71)                    |
|      | b. subcategorization  | Chomsky (1965)                                       |
|      | c. argument structure | e.g. Goldberg (1995)                                 |
|      | d. government model   | Mel’čuk (1974) (Russian <i>model’ upravljenija</i> ) |
|      | e. clause blueprint   | Grebe (1959) (German <i>Satzbauplan</i> )            |

Those nominals in a clause that are determined by (i.e. depend on the choice of) the verb are now generally called “arguments”, while the verb-independent elements are called “adjuncts”. Again, this distinction is widely recognized, but in the past the terminology varied quite a bit:

- |      |               |                   |                   |
|------|---------------|-------------------|-------------------|
| (14) | a. argument   | adjunct           | (here)            |
|      | b. complement | modifier, adjunct | e.g. Vater (1978) |
|      | c. actant     | circumstant       | Tesnière (1959)   |
|      | d. argument   | satellite         | Dik (1997: 86–90) |

What we call the “role frame”, i.e. the semantic representation of the verb meaning with the participant variables, has been variously called “logical structure” (e.g. Van Valin & LaPolla 1997), “lexical-conceptual structure” (e.g. Jackendoff 1990), or “semantic valency”.

Throughout this chapter, we will work with the distinction between descriptive categories of particular languages and the comparative concepts of cross-linguistic research (Haspelmath 2010). In the practice of linguists, verbal valency is primarily used as a descriptive concept, needed to characterize the behaviour of particular verbs in particular languages. Thus, we need to develop a number of comparative concepts that allow us to compare valency patterns across languages.

The most important concepts are the semantic concepts **comparison meaning** (meanings that are used to compare verbs across languages, e.g. ‘miss’ and ‘fill’ in (1)–(2)) and **microrole** (meanings that are used to compare arguments with similar roles across languages, e.g. ‘liker’ and ‘likee’ in (10)), and concepts for formal coding elements, in particular **flags** (cases and adpositions) and **indexes** (see § 6 below). We have also tried to work with a comparative concept of argument, but as we will see in the next section, this is more difficult to apply consistently across languages than the comparison meanings, the microroles, and the coding elements.

---

<sup>2</sup> Note also that the alternative form “valence” is also common, especially in American English (e.g. Abraham 1978; Langacker 1988).

## 4 Delimiting valency: arguments and adjuncts

A lot of research on valency has been concerned with the practical task of compiling valency dictionaries, i.e. dictionaries of verbs that give a valency frame for every verb. This forces researchers to delimit the notion of valency very clearly, and to distinguish sharply between arguments (included in the valency frame of the verb) and adjuncts. But sometimes it is not quite clear whether a nominal is specific to the verb or not. Consider the bracketed phrases in (15).

- (15) a. *She put the book [on the table].*  
 b. *He lives [in Wisconsin].*  
 c. *She is sitting [on the sofa].*  
 d. *They cut the meat [with the knife].*  
 e. *He broke the window [with his fist].*

These are locational and instrumental phrases which could be taken either as arguments or as setting or manner adverbials and thus as adjuncts. Is there a way to tell whether they are arguments or adjuncts?

One way in which the distinction between arguments and adjuncts has often been framed is by considering the verb's meaning. If a participant is **entailed** by the meaning of a verb, this participant is an argument, otherwise it is an adjunct (e.g. Van Valin 2005; Bickel 2011). According to this criterion, *with the knife* in (15d) would be an argument, because *cut* means 'sever with a sharp instrument', while *with his fist* in (15e) would not be an argument, because *break* only means 'do something so that something becomes broken'. However, not all entailed participants can be overtly expressed, not all entailed participants would be considered arguments, and not all elements that are generally considered arguments are entailed participants.

First, not all entailed participants can be overtly expressed. For example, the English verb *tell* allows the expression of both the addressee and the content, but *lie* does not allow the expression of the content:

- (16) a. *She told her address to her friend.*  
 b. *She lied (\*her address) to her friend.*

Second, not all entailed participants would normally be considered arguments. For example, the verb *break* means 'do something<sub>1</sub> so that something<sub>2</sub> becomes broken'. The first entailed participant, the action that causes the breaking event, can be expressed in a *by*-phrase (*He broke the window by hitting it with his fist*), but such a *by*-phrase would not normally be regarded as an argument. More generally, all verbs that denote a spatiotemporal event entail a time and a location, but these are the most typical adjuncts, not arguments.

Third, some arguments are not entailed by the verb's meaning, e.g. the beneficiary *me* in (17).

(17) *Roland baked me a cake.*

As has been widely discussed, *bake* does not entail a beneficiary (one can bake a cake without having someone specific in mind), but the beneficiary is generally regarded as an argument.

Thus, because there is no close match between entailed participants and what are normally considered arguments, we do not adopt the entailment-based definition of *argument*.

We think that the notion that best captures the intuition that lies behind the argument-adjunct distinction is the notion of **verb-specificity**. Elements that are verb-specific are arguments, and elements that are not verb-specific are adjuncts. Perhaps the clearest case of verb-specificity is **coding-specificity**, i.e. where a verb determines idiosyncratic coding of its arguments, so that the coding is not predictable on the basis of the verb's meaning. This can best be seen in minimal pairs of verbs with very similar or identical meaning but different argument coding. Thus, the semantic role of *the boy* or *der Junge* is not very different in the (a) and (b) examples in (18)–(19), but only the (b) examples require a dative preposition or case. This is something that needs to be learned in addition to the meaning of the verb, and it is thus no accident that systematic valency research began in the context of language teaching (Helbig & Schenkel 1969).<sup>3</sup>

(18) English

- a. *I showed **the boy** the solution.*
- b. *I demonstrated the solution **to the boy**.*

(19) German

- a. *Ich unterstützte **de-n** Junge-n.*  
I.NOM support.PST.1SG the-ACC boy-ACC  
'I supported the boy.'
- b. *Ich half **de-m** Junge-n.*  
I.NOM helped the-DAT boy-DAT  
'I helped the boy.'

Another example of this type was given in the very first paragraph of this paper (English *wait for someone* vs. *expect someone*).

---

<sup>3</sup> There was of course valency research before Helbig & Schenkel (1969), but they published the first valency dictionary, which was perhaps also the first book specifically on valency patterns in a language (incidentally, this work was carried out at Leipzig University and published in Leipzig).

Another fairly clear case of verb-specificity is **obligatoriness**. In many languages, some kinds of nominals must occur when a given verb is chosen. This is true not only of subjects (which are fairly trivial cases of arguments and therefore sometimes even ignored by valency researchers), but also of many objects, and sometimes even of oblique arguments. For example, the object in (20a) cannot be omitted (contrasting with 20b), and the prepositional *auf*-argument in (21a) cannot be omitted (contrasting with 21b).

(20) English

- a. *The dragon devoured **the princess**.*  
     (\**The dragon devoured.*)
- b. *The customer ate (**the fish**).*

(21) German

- a. *Sie hat mich **auf den Fehler** hingewiesen.*  
     she has me on the mistake pointed  
     ‘She pointed me to the mistake.’
- a’. (\**Sie hat mich hingewiesen.*)
- b. *Sie hat (**auf Godot**) gewartet.*  
     she has on Godot waited  
     ‘She waited for Godot.’

But there are many arguments that exhibit neither coding specificity nor obligatoriness. In fact, in many languages no argument is ever obligatory, because all arguments can be omitted when they can be reconstructed from the context.<sup>4</sup> And in many cases, the absence of an argument may be odd for pragmatic reasons, not for any grammatical reasons:

- (22) a. ??*He lives.*
- b. ??*She is sitting.*

Hearing about someone living or sitting is not informative, so the location is usually specified to get a pragmatically felicitous utterance. Thus, obligatoriness is not as useful for delimiting valency as is often thought, not even for those languages where it does play a role.

---

<sup>4</sup> However, often one can make a clear distinction between an anaphoric and an existential interpretation of argument omission, and when argument absence implies an anaphoric interpretation, this could be taken as evidence of verb-specificity and argumenthood. Thus, in *She found out* the omitted argument has to be definite and anaphoric (= *She found out about it*), whereas in *She ate* the omitted argument has to be indefinite and non-anaphoric (= *She ate something*).

The crucial criterion for argumenthood is whether a nominal is limited in its cooccurrence options to a restricted and semantically arbitrary set of verbs (i.e. whether it is verb-specific), or whether it can occur with any verb, or at least with a large and semantically coherent class of verbs (i.e. whether it is verb-free). We can call this **specificity of occurrence**. Thus, while the boldfaced arguments in (23) are not coded in a (highly) verb-specific way and are not obligatory, they are still considered arguments because they cannot occur freely with any verb, as shown in (24).

- (23) a. *He called **his brother**.*  
       b. *She gave **the shopkeeper** too much money.*  
       c. *He ran **to the house**.*
- (24) a. *\*She laughed **her brother**.*  
       b. *\*He spent **the shopkeeper** too much money.*  
       c. *\*She sang **to the house**.*

One widely cited test for (non-)argumenthood that reflects specificity of occurrence is the ‘happen’ test: If a sentence has a paraphrase in which a phrase from the original sentence is removed and occurs in an anaphoric ‘happen’ (or ‘do so’) clause, then this phrase is an adjunct:

- (25) a. *She called her brother **in the morning**.*  
       b. *She called her brother, and this happened **in the morning**.*
- (26) a. *They accused her of plagiarism **in her school**.*  
       b. *They accused her of plagiarism, and this happened **in her school**.*  
       (OR: *and they did so in her school*)

This test shows clearly that temporal and locational setting adverbials are adjuncts, while obligatory and coding-specific nominals are arguments:<sup>5</sup>

- (27) a. *She called **her brother**.*  
       b. *\*She called, and this happened **her brother**.*
- (28) a. *They were waiting **for Godot**.*  
       b. *\*They were waiting, and this happened **for Godot**.*<sup>6</sup>

---

<sup>5</sup> In the ValPaL manual, we asked the contributors to determine argumenthood by means of a test of this kind. But several participants told us that no such test is available in their language (e.g. Seifart 2013 on Bora). Thus, this test is not readily available as a cross-linguistically applicable way of determining arguments (see also Haspelmath 2014).

<sup>6</sup> This sentence is acceptable with an irrelevant sense (‘this happened for the sake of Godot’).

The test can be applied to many languages, because anaphoric verbal expressions like ‘do so’ or ‘this happened’ are widely found. However, with (stationary)<sup>7</sup> locational phrases and instrument nominals, the result of the test is not so clear:

- (29) a. *She was sitting **on the sofa**.*  
       b. *?She was sitting, and this happened **on the sofa**.*
- (30) a. *They cut the meat **with the knife**.*  
       b. *?They cut the meat, and this happened **with the knife**.*

Quite generally, locational phrases and instrument nominals are hard to classify uniquely as arguments or adjuncts. The concept of valency and the argument/adjunct distinction is simply not very useful for these kinds of phrases.<sup>8</sup>

In addition to coding specificity and obligatoriness, quite a few other language-particular criteria for distinguishing between arguments and adjuncts have been cited for different languages in the literature (cf. Haspelmath 2014). However, since we are interested in cross-linguistic comparison, language-particular argument criteria are not of interest to us.

Thus, for quite a few cases we did not have a unique way of distinguishing between arguments and adjuncts, and the ValPaL database is therefore not consistent in this regard. If the verb ‘break’ is said to have three arguments in one language and two arguments in another language, this does not mean that there is a typological difference between the two languages. This may just be due to an arbitrary or at least idiosyncratic difference between the decisions taken by the authors. As a result, the number of arguments is not a kind of information that should be taken as important for cross-linguistic comparison, especially when the difference concerns locational phrases and instrument nominals.

## 5 Comparison meanings: verbs and microroles

For cross-linguistic comparison of valencies, or more specifically coding frames, we need to be able to compare verbs (the valency-bearers), participant roles, and coding elements (flags and indexes, see § 6 below). For example, a comparison as in (2a–b), repeated here from above, presupposes that we know that both English

---

<sup>7</sup> By contrast, directional locational phrases are clearly arguments: *He ran to the house* / *\*He ran, and this happened to the house*.

<sup>8</sup> Note also that the ‘happen’ test cannot be applied to subjects in English; and the criterion of verb-specificity might be taken to indicate that English subjects are not arguments, because they occur with all verbs. (However, nobody has suggested that subjects are not arguments, in English or other languages, so this has no practical consequences.)

*fill* and Chintang *phatt-* mean ‘fill’, that the verbs in addition to an agent role have a substance role (in this example, the water) and a location role (in this example, the container), and that the substance role is coded by a preposition (*with*) in English, and with Nominative case in Chintang.

- (31) a. English  
*She<sub>1</sub> filled the container<sub>2</sub> with water<sub>3</sub>.*
- b. Chintang (Tibeto-Burman, Nepal)  
*Huĩsa-ŋa<sub>1</sub> cuwa-Ø<sub>3</sub> gagri-be<sub>2</sub> phatt-e.*  
 she-ERG water-NOM container-LOC fill-PST  
 ‘She filled the container with water.’  
 (Literally ‘She filled the water into the container.’)  
 (Bickel et al. 2010: 387)

Thus, our ValPaL database contains entries such as those in (32a–b). The first line contains the verb form (its citation form), the second line contains the coding frame, and the third line contains the microroles, i.e. the participant roles relevant for this particular verb.

- (32) a. English *fill*  
 <1-NOM V.subj[1] 2-ACC with+3>  
 (1=filler, 2=filled container, 3=filling material)
- b. Chintang *phatt-*  
 <1-ERG V.subj[1] 2-NOM 3-LOC>  
 (1=filler, 3= filled container, 2= filling material)

In other words, the verb entries in our database contain the kinds of information that we saw above in (9)–(10), though in a slightly abbreviated notation, with microrole labels instead of full role frames with index numbers (thus, “1=filler, 2= location, 3=substance” is equivalent to “Arg1 fills Arg2 with Arg3”).

The 80 pre-defined comparison meanings that we used to gather comparable verb forms for the database are identified by their English counterpart (which is also their label) and by a typical context. For example, the meaning ‘cover’ is identified by the label COVER as well as the typical context “The woman covered the boy with a blanket.”. We did not try to provide a more detailed semantic description, as we did not expect this to lead to greater comparability of verbs across languages.<sup>9</sup>

<sup>9</sup> We are aware that the use of English labels to elicit counterpart verbs, as well the use of English as a general metalanguage of our project, may well have biased the set of comparison meanings toward the kinds of meanings that tend to have simple expression in English and related languages. Unfortunately, we felt that there was nothing we could do about this, other than avoiding verb

Languages differ not only in their forms, but also in the kinds of meanings that they have words for, so the verbs that are given as counterparts to our 80 comparison meanings are not always perfect matches. For this reason, we do not call them “equivalents”, but “counterparts”. We asked the contributors to find the semantically closest verb in their language. Moreover, we said that the verb should have a “basic” flavour, i.e. verbs that are used very rarely should be avoided if a more common verb with similar meaning is available. Sometimes a basic verb may have a somewhat different meaning from the English label that serves as comparison meaning; in such cases, we said that deviation from the comparison meaning was less important than basicness. The reason for this was that we felt that there was no need to insist on exact matches, because this was not achievable anyway in many situations. And it must be kept in mind that the purpose of our database is not to provide exact translations from each language to each other language, but to compare languages with respect to their valency patterns.<sup>10</sup>

The relationship between the pre-defined comparison meanings and counterpart verbs can be many to many, so that if there were two basic verbs in the language that corresponded to a given meaning, both could be included. For example, in Sliammon, there are two verbs corresponding to EAT, *ʔiltən* and *məkʷt*, with two different coding frames (Watanabe 2013). Conversely, when a single verb corresponded to two different meanings, it was not necessary to enter it twice into the database. For example, the Mandinka verb *bori* is a counterpart both of RUN and of ROLL (Creissels 2013). Homonymous verbs are distinguished by a number, e.g. Nllng *nllaa* (1) (‘dress’) and *nllaa* (2) (‘live’).<sup>11</sup>

The argument variables in the coding frames are represented by integers (1, 2, 3, ...) whose order does not have any significance (though normally the linearly first argument gets variable 1, and so on). Each argument variable is linked to a microrole, so that we know what role the argument plays in the verb meaning. With typical transitive verbs, as in (33), the microroles are not so important, because they could normally be replaced by “agent” and “patient”, and with single-argument verbs, as in (34), they are completely redundant.

---

meanings for which we were aware from the beginning that there is something peculiar about their English counterparts.

**10** Every translation distorts the original text to some slight extent, but translations are still eminently useful. We do not expect our cross-linguistic comparison to be any better (or worse) than the average professional translation.

**11** Since a verb must have a unique coding frame in our database (this was a decision we took to simplify the database), if a verb had two slightly different meanings but different coding frames, it had to be broken up into two different verb entries that are treated as homonymous. For example, Yaqui *chaae* has the coding frame <1-NOM V> when it renders ‘scream’, but <1-NOM 2-ACC-DIR V> when it means ‘shout at’ (Estrada-Fernández et al. 2013). ‘Scream’ and ‘shout’ would be sufficiently similar to count as one meaning, and no doubt most linguists would say that we are dealing with the same verb here, but since there are two possible coding frames, two different ValPaL database entries are required.



- (33) a. EAT            eater            eaten food  
       b. WASH        washer        washed entity  
       c. HELP        helper        helpee
- (34) a. COUGH      cougher  
       b. SINK        sunken entity

(Microroles sometimes have unusual-sounding labels such as *helpee*; we hope that readers and database users will quickly get used to them. We find transparent labels much more practical than opaque abbreviations or numbers.)

The microroles are important for those verb meanings where cross-linguistic metataxis is not uncommon. We saw some examples in (1)–(2/31) above, and here is another contrasting pair of verbs that are counterparts of the same meaning from two languages:

- (35) Sri Lanka Malay (Nordhoff 2013)  
       a. verb:                *mintha*- ‘ask for’  
       b. coding frame:    <1 2-ACC 3-LOC V>  
                                   (1=asker, 2=requested thing, 3=askee)  
       c. example:  
           *see baapa=ka car=yang su-mintha*  
           1SG father=LOC car=ACC PST-beg  
           ‘I asked my father for the car.’ (Lit. ‘I asked the car at my father.’)
- (36) Jaminjung (Schultze-Berndt 2013)  
       a. verb:                *yanggi ganarrany* ‘ask for’  
       b. coding frame:    <1-ERG 2-ABS 3-DAT SBJ[1].OBJ[2].V>  
                                   (1=asker, 2=askee, 3=requested thing)  
       c. example:  
           *gurrany yanggi ya-wun-karra=yinyag mangarra-wu*  
           NEG ask IRR-DU.A:1.P-put=1DU.EXCL.P plant.food-DAT  
           ‘Don’t ask us two for food, you two!’ (Lit. ‘Don’t ask us to food.’)

The microroles ‘askee’ and ‘requested thing’ do not clearly map onto the general roles ‘patient’, ‘source’ and/or ‘recipient’, but in our approach this does not matter: The coding frames in (35b) and (36b) are perfectly clear and comparable on the basis of the microroles. Intuitively, the Jaminjung sentence is more like English (using a dative case or the requested thing), while the Sri Lanka Malay sentence is literally ‘I asked the car at my father’, which is rather different from English, and more like Russian *poprosit’* (*ja poprosila mašinu u otca* [I asked the.car at father]). So there is a metataxis relationship between (35) and (36).

## 6 Coding elements

We saw above (§ 2) that valency has both a coding and a behavioural component, but here we limit ourselves to coding frames. And while behavioural properties of arguments are quite heterogeneous, there are just two kinds of **coding elements**, which are quite straightforward to represent schematically: **flags** (adpositions or case markers) on arguments, and **indexes** on the verb that match arguments. Let us consider three additional examples, which illustrate the conventions that we use:

(37) Japanese (standard) (Kishimoto & Kageyama 2013)

- a. verb: *kowasu* ‘break’
- b. coding frame: <1-NOM (2-INSTR) 3-ACC V>
- c. example:  
*Ken-ga (hanmaa-de) mado-o kowasi-ta.*  
 Ken-NOM hammer-INS window-ACC break-PAST  
 ‘Ken broke the window (with a hammer).’

(37) Hoocąk (Hartmann 2013)

- a. verb: *hožu* ‘load’
- b. coding frame: <1 2 LOC3 UND[2].ACT[1].V>
- c. example:  
*wiiraruti=ra kšee ho<ha>žu*  
 trailer=DEF apple <1E.A>put.in  
 ‘I loaded apples on the trailer.’

(39) Icelandic (Barðdal 2013)

- a. verb: *stela* ‘steal’
- b. coding frame: <1-NOM V.AGR[1] 2-DAT (*frá*+3-DAT)>
- c. example:  
*Þjófur-inn stal peningum frá gömlu konu-nni.*  
 thief-the.NOM stole money.DAT from old.DAT woman-the.DAT  
 ‘The thief stole money from the old lady.’

The following conventions are used to represent coding elements (flags and indexes) in coding frames in ValPaL:

- (i) Cases are represented by (abbreviations of) their category labels, such as NOM (for nominative), DAT (for dative), following the argument variable and linked to it by a hyphen.

- (ii) Adpositions are represented by their form, linked to the argument variable by a plus (+) sign, preceding it (for prepositions) or following it (for postpositions).
- (iii) Index-sets are represented by their category labels, linked to the verb by a period (e.g. V.AGR in Icelandic). The index label is immediately followed by brackets which contain the argument variable number, so “V.AGR[1]” in Icelandic means that the agreement index-set corresponds to the argument that also bears nominative case (the stealer in the case of the verb *stela* ‘steal’), and “UND[2].ACT[1].V” in Hoocak means that arguments 1 and 2 are indexed by the Actor and Undergoer sets, respectively (the loader and the loaded theme in the case of *hožu* ‘load’).
- (iv) Optional arguments may be enclosed in brackets. (We did not enforce this, because optionality is a difficult concept, and we were not able to provide this information consistently.)

## 7 Kinds of arguments

In the most common case, verbal arguments are nominals (noun phrases, adpositional phrases) or person indexes on verbs. Such arguments are represented in our coding frames by free-standing integers (possibly accompanied by adpositions or with case labels attached to them) or by integers in brackets following the index-set label which is attached to the verb variable.

But there are certain other kinds of arguments, in particular locational arguments, clausal arguments, and utterance arguments.

A locational argument need not have any particular categorial form. It is most often expressed by an adpositional phrase, but the nature of the adposition is determined by the spatial meaning, not by the verbal valency. In many languages, locational adverbs may also be expressed by spatial adverbs which are neither adpositional phrases nor noun phrases. For example, a locational argument occurs with Italian *sedere* ‘sit’, illustrated in (40a–c).

(40) Italian (Cennamo & Fabrizio 2013)

- a. *Mario siede in seconda fila.*  
Mario sits in second row  
‘Mario is sitting in the second row.’
- b. *Gli anziani sede-va-no intorno a-l fuoco.*  
the old.man-PL sit-IPFV-3PL around to-the fire  
‘The old men were sitting around the fire.’
- c. *Mario sied-e lì.*  
Mario sits there  
‘Mario is sitting there.’

In ValPaL's coding frames, locational arguments are notated with LOC, followed by the argument number. Thus, the coding frame of Italian *sedere* is <1 V.SUBJ[1] (LOC2)>. (A locational argument also occurs in the coding frame of Hoocak *hožu* 'load', see (38b) above.)

Clausal arguments are complement-clause arguments such as *that*-clauses and infinitival clauses in English (e.g. *I hope that you'll be there; I hope to be there*). These were excluded from ValPaL, as the typology of complement clauses is an entirely separate domain of study. The only comparison meanings that frequently require a clausal argument are 'know' and 'want', as illustrated by (41b). However, we asked our contributors for contexts like (41a), where the 'wanted thing' argument is a nominal rather than a clause. Thus, Eastern Armenian *uzel* 'want' has the coding frame <1-NOM 2-NOMDAT V.SUBJ[1]> in ValPaL.

(41) Eastern Armenian (Khurshudian & Daniel 2013)

- a. *Jes uz-um em ajs akžka-n.*  
 I.NOM want-CVB.IPFV AUX.1SG this girl.DAT-DEF  
 'I want this girl.'
- b. *Jes uz-um em gn-al.*  
 I.NOM want-CVB.IPFV AUX.1SG go-INF  
 'I want to go.'

There is another comparison meaning that takes a special kind of argument: 'say' usually takes what we call an **utterance argument** (and marginally 'tell' and 'ask' can have such an argument as well). Like locational and clausal arguments, utterance arguments cannot be treated like nominal arguments because they are not coded by the usual flags (cases, adpositions) or indexes. Most commonly, they are simply juxtaposed (as in 42a–b), but occasionally there are special quotative markers, as in Japanese (see 42c).

(42) a. Jakarta Indonesian (Connors & Gil 2013)

*Nadia omong kita mendingan pergi.*  
 Nadia say 1PL better go  
 'Nadia said it would be best if we go.'  
 (OR: 'Nadia said: "We better go".')

b. Ket (Vajda & Kryukova 2013)

*Qar' da'ŋa bara: "eddi".*  
 DEM.M 3SG.M.DAT 3.M.SG.say alive-1SG  
 'He said to him: "I'm alive".'

c. Japanese (standard) (Kishimoto & Kageyama 2013)

*Ken-ga Mari-ni "hai" to it-ta.*  
 Ken-NOM Mari-DAT yes QUOT say-PST  
 'Ken said "yes" to Mari.'

Utterance arguments are notated with UTT in ValPaL, followed by the argument number (thus, Jakarta Indonesian *omong* has the coding frame <1 V (UTT2)>).

## 8 Prominence-based splits

Coding frames contain a unique coding element, i.e. a unique case or adposition associated with a free-standing argument number, or a unique index-set on the verb that indexes an argument. But sometimes languages exhibit prominence-based splits, i.e. depending on the argument's inherent or contextual prominence features, different flags are employed. A well-known case of this is differential object marking in Spanish, where human direct objects have the preposition *a* (e.g. *veo a Juan* 'I see Juan'), while nonhuman objects lack this preposition (e.g. *veo la casa* 'I see the house'). In such cases, one would not say that the verb *ver* 'see' is associated with two different coding patterns, because the coding contrast does not depend on the verb, but on the nature of the argument.

But an even better-known case is English, where a Nominative-Accusative distinction is made only with personal pronouns (*I/me, he/him, she/her, we/us, they/them, who/whom*), i.e. with nominals that are higher in prominence. All other nominals make no distinction. When a distinction is only partial and depends on the kind of nominal, then the usual solution adopted by linguists is to encode the distinction in the syntactic rule and to make a separate (morphological) statement about the lack of the distinction in certain kinds of nominals. Thus, in Russian we would say that there is a general Nominative-Accusative distinction, even though it is syncretized in many inanimate nouns in the singular (e.g. NOM/ACC *kniga/knigu* 'book', but *stol/stol* 'table', *mesto/mesto* 'place'), and in all inanimate nouns in the plural (e.g. NOM/ACC *knigi/knigi* 'books'). Likewise, we decided to say that English has a general Nominative-Accusative distinction, even though it is syncretized in all nominals that are not personal pronouns (and in the personal pronouns *you* and *it*). Thus, the English verb *see* has the coding frame <1-NOM > V.SUBJ[1] > 2-ACC> in ValPaL.

Another case of a prominence-based split is Eastern Armenian, where a direct object is consistently in the Dative case when it is human, but in the Nominative case when it is inanimate (Khurshudian & Daniel 2015). In the glosses in the Eastern Armenian examples, the case labels DAT and NOM are used, e.g.

(43) Eastern Armenian (Khurshudian & Daniel 2013)

- a. *Mard-ə tesa-v tʰa-ji-n.*  
 man[NOM]-DEF see.AOR-AOR.3SG boy-DAT-DEF  
 'The man saw the boy.'
- b. *Jes cʰamakʰ em tesn-um.*  
 I.NOM land[NOM] AUX.1SG see-CVB  
 'I see the land.'

However, since the Dative and the Nominative occur in complementary distribution with direct objects, they are treated as a single case (“Dative-Nominative”) for the purposes of valency information, and the coding frame of *tesnel* ‘see’ in the database is <1-NOM 2-NOMDAT V.SUBJ[1]>.

## 9 Verb forms

Following our guidelines, the verbs in our database have to be conventional expressions in the language, but they do not have to be basic or simple verbs, and they do not even have to be “verbs” in a very strict sense.

The requirement that the entries are conventional (or FIXED) expressions means that we did not want the contributors to provide ad-hoc paraphrases. For example, in Bezhta the verbal meaning ‘peel’ is rendered by ‘take (off)’ plus ‘skin’ (Comrie & Khalilova 2013).

- (44) *Kibba k’atu-wa-s beš b-ayo-yo.*  
 girl.ERG potato-OBL-GEN1 skin(III).ABS III-take-PST  
 ‘The girl peeled the potato.’ (Lit. ‘The girl took (off) the skin of the potato.’)

Another example is the meaning ‘sing’, which is rendered by *donkilóo láa* [song say] in Mandinka. This is a fully regular combination of a verb (‘say’) and its object (‘song’). Other objects would be equally possible (e.g. ‘tell a story’). Thus, the database has “no counterpart” for SING in Mandinka and PEEL in Bezhta.

Since it is very difficult to distinguish between fixed expressions and freely created expressions, our database may not be homogeneous in this regard. For example, for ‘rain’, we have *bi-ga o-da* [rain-NOM come-DECL] in Korean, but in Bezhta, the very similar *wodo guu-s* [rain come-PRS] was not counted as a conventional expression. This may not correspond to a real difference in the fixedness of the two expressions.

The counterparts of our comparison meanings are not always really verbs in the languages, but we wanted to avoid the issue of the verb-adjective distinction and thus basically ignored the language-particular status of the counterparts. We were exclusively interested in the valency properties of expressions corresponding to meanings such as ‘be afraid’, ‘like’, ‘feel cold’, ‘be sad’, ‘be dry’, not in their word-class assignment. It is apparently the case that all counterparts of dynamic comparison meanings are actually verbs in all languages, but counterparts of meanings such as ‘be afraid’ and ‘be sad’ (and especially ‘be dry’) are very often called “adjectives” in language descriptions. In such cases, the entries often include copulas, but these copulas should be regarded as irrelevant to our underly-

ing question of valency (we did not try to be consistent in including or excluding copulas).<sup>12</sup>

While fully regular ad-hoc paraphrases are not included, as just noted, **complex verbal expressions** are allowed in the database, and were in fact used in many cases (more than a quarter). They may be quite diverse formally: They may consist of a verb plus an additional locative particle (e.g. English *sit down*), or a verb plus a reflexive pronoun (e.g. German *sich hinsetzen* ‘sit down’), or a verb plus an incorporated noun (e.g. Bezhta *komak bowal* ‘help’, lit. ‘do help’), or two verbs (e.g. Yoruba *mú ... wá* ‘bring’, lit. ‘take ... come’). They may also include valency-changing affixes, especially causative affixes, e.g. Mandinka *nikin* ‘teach’, which derives from *nikin* ‘learn’.

It is important to be aware that by including complex verbal expressions in the database, we are on the one hand maximizing the amount of data that we can include for each language, but on the other hand we are introducing a possible bias. The comparison meanings that we started out with almost all have simple counterparts in English (and closely related European languages), but not uncommonly, they do not have simple counterparts in other languages. Thus, the English verb set (Goddard 2013) consists of simple verbs, while the verb sets of other languages often include derived verbs such as causatives or other complex verbal expressions. Since complex verbal expressions may have different valency profiles than simple verbs, we may not get an unbiased picture of the general valency character of the language.

We chose to include complex verbal expressions in order to maximize the amount of data in the database, and because it is not clear that complex verbal expressions are less fundamental to a language’s character than simple, monomorphemic verbs. The question of how best to compare verbal syntax across languages in a systematic way is still quite open, in our view, so we did not want to limit the available data in advance. We did try to annotate all verbs in the database for simple vs. complex status,<sup>13</sup> so that users can filter out complex verbal expressions if they want.

With complex verbal expressions that include a nominal element, the question may arise what the relation of this nominal element is to the verb’s valency. Consider the examples in (45)–(47).

---

<sup>12</sup> One of our comparison meanings was ‘be a hunter’. Including this was motivated by the question whether the subject of the predicate ‘is a hunter’ is coded in the same way as the subject of other stative predicates such as ‘be sad’ or ‘be afraid’. We recognize that including such a meaning in a database on verbs is somewhat confusing, and as a result some contributors did not include a counterpart.

<sup>13</sup> More precisely, a *complex* verb in the database is a verbal expression that includes an element that is relevant for the verb’s valency, such as a causative marker, an applicative morpheme or an incorporated noun.

- (45) Bezhta *komak bowal* [help do] ‘help’  
*Kibba iyo-l komak b-oo-yo.*  
 girl.ERG mother-LAT help(III) III-do-PST  
 ‘The girl helped her mother.’  
 (Lit. ‘The girl did help to the mother.’)  
 (Comrie & Khalilova 2013)
- (46) Yucatec Maya *ch’a’ sahakil* [take fear] ‘be afraid’  
*Le wíinik-o’ t-u ch’a’-ah sahak-il ti’ le báalam-o’.*  
 DEM man-D2 PFV-SBJ.3 take-CMPL fear LOC DEM jaguar-D2  
 ‘The main became afraid of the jaguar.’  
 (Lehmann 2013)
- (47) Icelandic *finna lykt* [find smell] ‘smell’  
*Björn-inn fann lykt-ina af strák-num.*  
 bear-the.NOM found smell-the.ACC of boy-the.DAT  
 ‘The bear smelled the boy.’  
 (Lit. ‘The bear found the smell from the boy.’)  
 (Barðdal 2013)

In these cases, one might at first want to say that we are dealing with transitive verbs, and that the incorporated nouns (‘help’, ‘fear’, ‘smell’) are objects of these verbs. This would not be incorrect, and indeed at some level this needs to be said, because otherwise the case-marking (ergative marking of the helper in (45), accusative marking of the incorporated noun in (47)) and the agreement (gender III agreement of the verb with ‘help’ in (45)) could not be explained. Certainly the Bezhta ‘do’ verb, the Yucatec Maya ‘take’ verb and the Icelandic ‘find’ verb are transitive, and this explains some facets of the complex verbal expressions. But from the comparative perspective, it is these complex verbal expressions as a whole which are the counterparts of the comparison meanings, and thus their valencies are as in (48a–c).

- (48) a. Bezhta *komak bowal* [help do] ‘help’  
 <1-ERG 2-LAT V>  
 (1=helper, 2=helpee)
- b. Yucatec Maya *ch’a’ sahakil* ‘be afraid’  
 <SBJ[1].V 1 le+2>  
 (1=fearer, 2=fear stimulus)
- c. Icelandic *finna lykt* ‘smell’  
 <1-NOM V.AGR[1] af+2-DAT>  
 (1=smeller, 2=smelled entity)



These “composed valencies” may sometimes look a bit odd, especially for Bezhta: A verb with an ergative but no absolutive argument is not normally possible, and most verbs normally agree with one argument in gender. In purely language-specific descriptions, composed valencies are often left aside, but strictly speaking, they are necessary for language descriptions, too. That the action of helping is standardly expressed by ‘do help’ in Bezhta is not fully predictable (alternatively, one could say ‘give help’, ‘extend help’, etc., or there could be a completely unrelated simple verb), and that the smelled entity is coded by the preposition *af* in Icelandic is not fully predictable (cf. also English *take part*, where it is not predictable that the activity is coded with the preposition *in*: *take part in something*).

Thus, the general principle of the ValPaL database is that only arguments that can be varied freely are taken into account as variables in the coding frames. Elements which are fixed parts of complex verbal expressions are ignored. In the coding frame <1-ERG 2-LAT V> of Bezhta *komak bowal*, the variable V stands for the complex verbal expression. The internal structure of counterpart verbs (or verbal expressions) is not reflected in the database.

## 10 Argument mismatches: apotaxis

In the simplest case, which is very frequent in the ValPaL database, the arguments of the counterpart verbs have roles that correspond to the microroles that we anticipated, and our anticipated microroles have corresponding arguments.

But this is not always the case. What we anticipated may admittedly have been based on English and similar European languages. But languages may differ not only in the way in which arguments expressing the same roles are coded (this is metataxis, cf. §1 above). They may also differ in the roles that can be expressed as arguments (this can be called “apotaxis”). Let us illustrate this with the verb meaning ‘dig’, which involves the creation of a hole (i) in the ground (ii) by a digger (iii), possibly with the goal of digging up a thing (iv) that was hidden in the ground. (In addition, digging of course requires an instrument, minimally a body part, but since the coding of instruments exhibits little variability, this is left aside here.) In German, the digger and the ground can be expressed simultaneously with the hole or with the hidden thing:

(49) German

- a. *Sie gruben ein tiefes Loch in die Erde.*  
 they dug a deep hole in the.ACC earth  
 ‘They dug a deep hole in the earth.’
- b. *Sie gruben in der Erde nach dem Schatz.*  
 they dug in the.DAT earth after the treasure  
 ‘They dug for the treasure in the earth.’

But the hole and the hidden thing cannot be expressed simultaneously (\**Sie gruben ein Loch nach dem Schatz* ‘They dug a hole for the treasure’).

In other languages in our database, our contributors gave us verbs that take primarily the ground as the argument of digging (e.g. Bezhta in 50, Hoocək in 51), or primarily the hidden thing (e.g. Bora in 52).

- (50) Bezhta (Comrie & Khalilova 2013)  
*Öždi mäche b-äx-čä.*  
 boy.ERG ground(III).ABS III-dig-PRS  
 ‘The boy digs the ground.’

- (51) Hoocək (Hartmann 2013)  
*Məq=ra ha-k’e.*  
 earth=DEF 1E.A-dig  
 ‘I’m digging the ground.’

- (52) Bora (Seifart 2013)  
*Mújcuri ó tsehdí.*  
 barbasco.root(ACC) 1SG dig  
 ‘I dig for barbasco root.’

Another manipulation verb meaning that shows variability is ‘peel’. In Russian, both the peel and the peeled object can be expressed as arguments (as seen in 53), while in German, only the peeled object can be an argument (at least with the verb *schälen*, 54). In Yucatec Maya, finally, only the peel is an argument (as seen in 55).

- (53) Russian (Malchukov & Jahraus 2013)  
*Povarënok očisti-l kartošk-u ot kožur-y.*  
 kitchen.boy.NOM peel.PFV-PST.M.SG potatoes-ACC from skin-GEN  
 ‘The kitchen boy peeled the potatoes.’  
 (Lit. ‘The kitchen boy cleaned the potatoes of their skin.’)

- (54) German  
*Die Großmutter schälte die Kartoffeln.*  
 the grandmother.NOM peeled the potatoes.ACC  
 ‘Grandmother peeled the potatoes.’

- (55) Yucatec Maya (Lehmann 2013)  
*Le xibpal-o’ t-u lak-ah u sóol le che’-o’.*  
 DEM boy-D2 PFV-SBJ.3 detach-CMPL POSS.3 shell DEM wood-D2  
 ‘The boy removed the bark from the stick.’  
 (Lit. ‘The boy removed the stick’s shell.’)

These differences between languages must be kept in mind when interpreting the cross-linguistic data of the ValPaL database. While most verbs have arguments that correspond largely to arguments in other languages, there are sometimes apotactic mismatches of the type just seen.

Another example is the comparison meaning ‘shave’. A verb with this meaning can take the shaved person as its non-agent argument (e.g. in Chatino, in 56), or the body part (e.g. in Ainu, in 57), or the hair/the beard (e.g. in Balinese, in 58), or it can be intransitive, because the affected entity is understood from the context (e.g. in Xârâcùù in 59, and also in English).

- (56) Zenzontepec Chatino (Campbell 2013)

*Yū=wá nka=téē? =yu j=yū.*  
 3SG.M=DEM CPL.CAUS-get.shaved=3SG.M OBL=3SG.M  
 ‘He shaved (lit. himself).’

- (57) Ainu (Bugueva 2013)

*E=nan-u hu memke.*  
 2SG.A=face-POSS shave  
 ‘Shave (lit. your face).’

- (58) Balinese (Shibatani & Artawa 2013)

*Anak=e muani ento nguris jenggot=ne.*  
 person=DEF male that shave beard=3.POSS  
 ‘The man shaved (lit. his beard).’

- (59) Xârâcùù (Moyse-Faurie 2013)

*Nâ xii.*  
 he shave  
 ‘He shaved.’

Similar cases of apotaxis are also found with other verbs involving the human body, such as ‘hurt’, where either the affected person can be an argument of the verb (as in English *I feel pain*), or the affected body part (as in English *My head is hurting*).

In general, apotactic variability in valency patterns can be attributed to metonymic shift (e.g. Waltereit 1998), but there is a lot of interesting cross-linguistic variation here that deserves further study (see also Michaelis & APiCS Consortium 2013 on diverse ways of expressing having a headache in pidgin and creole languages).

## 11 Constituent order

Unlike argument flagging and argument indexing, constituent order is not a coding element that is generally available. The primary way in which constituent order may serve to code semantic roles in some languages is by requiring agents to precede the verb and patients to follow the verb, as in typical SVO languages like English or Yoruba. Since constituent order is salient in some languages, we allowed the contributors to provide this information, but since it is often difficult to say whether a particular order is required for coding a relation, we made this an optional feature. In (60) we give a few examples of coding frames that include constituent order information.

(60) English

a. *be afraid* ‘fear’ <1-NOM > V.SUBJ[1] > of+2>

Emai

b. *dia* ‘live’ <1 > V > *vbi*+2>

Mandarin Chinese

c. *zhuāng* ‘load’ <1 > (LOC2 >) V > 3>

Mandinka

d. *ma* ‘touch’ <1 > 2 > V > 3+*la*>

Mandarin Chinese and Mandinka are atypical in that they have fairly rigid constituent order, but have some arguments that must occur in preverbal position, while others must occur in postverbal position:

(61) Mandarin Chinese

(Zhang Guohua 2013)

*Gōngrén-mén wǎng kǎchē-shàng zhuāng gèncǎo.*

worker-PL PREP truck-LOC load hay

‘The workers loaded hay onto the truck.’

(62) Mandinka

(Creissels 2013)

*Kambaan-óo ye saá maa fál-oo la.*

boy-DEF PFV.POS snake.DEF touch stick-DEF OBL

‘The boy touched the snake with a stick.’

While constituent order is represented in the coding frames of these languages in ValPaL, order is not regarded as a coding element.

## 12 Alternations

In many languages, verbs may systematically be associated with different valencies under different circumstances. Such situations are called **valency alternations**.

Three well-known examples from English are the Passive alternation, the Dative alternation and the Causal-noncausal alternation (also known as “ambitransitive” or “labile” verbs).

(63) Passive ( $V \rightarrow be\ V'-en$ )

<1-NOM V 2-ACC>                       $\rightarrow$     <2-NOM V' (*by*+1)>

- a. *The bear caught a fish.*
- b. *A fish was caught by the bear.*

(64) Dative alternation (uncoded)

<1-NOM V 2-ACC *to*+3>                       $\rightarrow$     <1-NOM V' 3-ACC 2-ACC>

- a. *She gave the money to her brother.*
- b. *She gave her brother the money.*

(65) Causal-noncausal alternation (uncoded)<sup>14</sup>

<1-NOM V 2-ACC>                       $\rightarrow$     <2-NOM V'>

- a. *He opened the door.*
- b. *The door opened.*

Like valency frames, valency alternations generally do not affect all verbs equally and thus subclassify the verbal lexicon in a language. Alternations often introduce many additional complexities and thus could not be covered exhaustively in Val-PaL. The general guideline was that not more than ten alternations had to be entered if the language had more. Also, alternations which more or less apply across the board (as in some languages causatives do) did not have to be included either, as they would not be very useful in distinguishing verb classes.

Alternations are sometimes subdivided into voices (like passive, middle and antipassive) and valency-changing operations (like causative, applicative, desubjective), but this distinction cannot be made consistently across languages. It is sometimes correlated with inflectional vs. derivational status, but this distinction likewise cannot be made consistently across languages. Another frequently made association is between voice and information-structural function, but other alternations may also have informational-structural effects (e.g. Peterson 2007: Chapter 4). Thus, we do not distinguish between voices and other valency-changing alternations.

Alternations can be grouped into **coded alternations**, where the verb undergoes a change (as in the English Passive alternation), and **uncoded alternations**, where the form of the verb is the same with the two alternate valency frames (as in 65 for English). In addition, alternations may preserve the semantic roles, or

<sup>14</sup> This alternation has often been called “inchoative-causative” alternation. See Haspelmath et al. (2014) for the more general and more transparent term “causal-noncausal”.

they may change the available semantic roles (by removing a role as in (65), or by adding a role as in (66) below), as long as the same basic meaning of the verb is preserved.

When an alternation was unproductive and occurred only with a small number of verbs, we were less interested in the alternation. (When the valency-frame variation affects only a single verb, it is not really an alternation, because alternations require some systematicity.)

In (63)–(65) above, the alternations are presented as **directed**, with a **basic** alternant on the left-hand side of the arrow and a **derived** alternant on the right-hand side. In coded alternations, it is generally easy to distinguish the basic from the derived alternant, because in most cases only the derived alternant involves some overt valency-changing morphology. Two more standard cases of this kind from other languages are given in (66)–(67). Mapudungun has an applicative suffix *-l*, and Chintang forms reciprocal verbs by reduplication and an affix *-ka-*.

(66) Mapudungun (Zúñiga & Lienlaf 2013)

<V.subj[1].obj[2] 2 1> → <V'.subj[1].obj[3] 2 1 3>

- a. *Ti pichi malen ye-tu-y kiñe rayen colegio mew.*  
 ART little girl carry-TEL-IND one flower school LOC  
 'The girl carried a flower to school.'

- b. *Ti pichi malen ye-l-fi kiñe rayen ñi kimeltuchefe.*  
 ART little girl carry-APPL-3.OBJ one flower 3.POSS teacher  
 'The girl carried a flower to the teacher.'

(67) Chintang (Schikowski et al. 2013)

<1-ERG 2-ABS V.AGT[1].OBJ[2]> → <1-ABS V'.SUBJ[1]>

- a. *Cha-ŋa puchak lauri-ŋa ten-o-s-e.*  
 child-ERG snake stick-ERG hit-[3SA.]3[s]P-PRF-IND.PST  
 'The child has hit the snake with a stick.'

- b. *Teĩ-ka-teĩ lus-i-niŋ-kha.*  
 beat-RECP-beat AUX-1p[S]-NEG-BGR  
 'Let's not beat each other!'

But when the alternation is uncoded, as in the Dative alternation and the Causal-noncausal alternation in English, it is often difficult or impossible to determine which of the alternants is basic. But given the setup of our database, where alternations are associated with basic verbs rather than abstract verb roots, this decision must be taken, be it in some arbitrary way.<sup>15</sup>

<sup>15</sup> The alternative to this would have been a database with a list of verb roots (or stems), plus a list of coding frames that can be associated with the verb roots in a many-to-many fashion. Such a database would not have contained a data type "alternation" at all. We considered this alternative,

Another issue that sometimes arises with alternations is that the same alternation occurs twice in our database, but in different directions. For example, Mandinka has a Causative alternation, which turns intransitive verbs into transitive verbs, as in (68).

- (68) Mandinka (Creissels 2013)
- a. *Saatéw-o jani-ta.*  
village-DEF burn-PFV.POS  
'The village burned / was burned.'
- b. *Jáw-oo-lu yé saatéw-o jani.*  
enemy-DEF-PL PFV.POS village-DEF burn  
'The enemies burned the village.'

It also has a Noncausative alternation, which turns transitive verbs into intransitive verbs, as in (69).

- (69) Mandinka (Creissels 2013)
- a. *Kambaan-óo ye palantéer-oo teyi ber-óo la.*  
boy-DEF PFV.POS window-DEF break stone-DEF OBL  
'The boy broke the window with the stone.'
- b. *Kíl-oo teyí-ta.*  
egg-DEF break-PFV.POS  
'The egg broke.'

These two alternations are of course really the same alternation in Mandinka. The reason they are treated as two alternations for the purposes of the database is that the relevant comparison meanings happened to be 'burn (intr.)' and 'break (tr.)' and alternations had to be unidirectional for the sake of the database

## 13 Conclusion

Let us summarize briefly our discussion and our choices for the Valency Patterns Leipzig database (Hartmann et al. 2013). As in other areas of morphosyntax, comparison of verbal valency requires semantic comparative concepts as a basis of comparison, in particular verbal comparison meanings and microroles (§ 5). Distinguishing consistently between arguments and adjuncts across languages is diffi-

---

but decided that it would have been more difficult to extract the information that interested us from such a database.

cult, but it is not really necessary to capture the most important aspects of valency variation, because the intermediate cases (especially locational and instrumental arguments/adjuncts) show relatively little variation across languages (§4). To be comparable across languages, coding frames must primarily contain information about argument flags (cases and adpositions associated directly with the argument nominals) and about argument indexes (person forms mostly associated with the verb) (§6). It is primarily nominal arguments that are of interest for valency comparison (§7), and prominence-based splits are not encoded in the valency frame (§8). Counterpart verb forms need not be verbs in the strict sense, but can be adjectives and complex verbal expressions, but they must be fixed conventionalized expressions. Incorporated nouns that are part of such complex expressions do not count for the valency of the verbal expression (§9). Cases of apotaxis, i.e. the availability of different roles for argument expression in verbs with similar meanings in different languages, are a non-negligible difficulty for comparing valencies that must be kept in mind (§10). Alternations have to be regarded as directed in a database which asks for each verb whether it undergoes a certain alternation or not (§12).

## Acknowledgements

We are grateful to Bernard Comrie and Andrej Malchukov for helpful comments on this paper.

## Special abbreviations

ACT	actor inflection
AOR	aorist
BGR	background
CPL	completive aspect
D1/2/3	proximal/distal/anaphoric/demonstrative
POS	positive
PREP	preposition
UND	undergoer inflection
TEL	telic

## References

- Abraham, Werner (ed.). 1978. *Valence, Semantic Case, and Grammatical Relations*. Amsterdam: Benjamins.
- Barðdal, Jóhanna. 2013. Icelandic Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary



- Anthropology. (Available online at <http://valpal.info/languages/icelandic>, Accessed on 2014-07-25)
- Bickel, Balthasar. 2011. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*, 399–444. Oxford: Oxford University Press.
- Bickel, Balthasar, Manoj Rai, Netra P. Paudyal, Goma Banjade, Toya N. Bhatta, Martin Gaenszle, Elena Lieven, Ichchha Purna Rai, Novel Kishore Rai & Sabine Stoll. 2010. The syntax of three-argument verbs in Chintang and Belhare (Southeastern Kiranti). In Andrej Mulchokov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in Ditransitive Constructions: A Comparative Handbook*, 382–408. Berlin: Mouton de Gruyter.
- Bugaeva, Anna. 2013. Ainu Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/ainu>, Accessed on 2014-07-25)
- Campbell, Eric. 2013. Zenzontepec Chatino Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/zenzontepec-chatino>, Accessed on 2014-07-25)
- Cennamo, Michela & Claudia Fabrizio. 2013. Italian Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/italian>, Accessed on 2014-07-25)
- Chomsky, Noam A. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Comrie, Bernard & Zaira Khalilova. 2013. Bezhta Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/bezhta>, Accessed on 2014-07-25)
- Connors, Thomas J. & David Gil. 2013. Jakarta Indonesian Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/jakarta-indonesian>, Accessed on 2014-07-25)
- Creissels, Denis. 2013. Mandinka Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/mandinka>, Accessed on 2014-07-25)
- Dik, Simon C. 1997. *The Theory of Functional Grammar*. 2 vols. Berlin: Mouton de Gruyter.
- Estrada-Fernández, Zarina, Mercedes Tubino Blanco & Jesús Francisco Villalpando Quiñonez. 2013. Yaqi Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/yaqui>, Accessed on 2014-07-25)
- Goddard, Cliff. 2013. English Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/english>, Accessed on 2014-07-25)
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.
- Grebe, Paul. 1959. *Grammatik der deutschen Gegenwartssprache*. (Duden). Mannheim: Bibliographisches Institut.
- Hartmann, Iren, Martin Haspelmath & Bradley Taylor (eds.). 2013. *Valency Patterns Leipzig*. Leipzig: MPI for Evolutionary Anthropology. (Available online at <http://valpal.info>)
- Hartmann, Iren, Martin Haspelmath & Michael Cysouw. 2014. Identifying semantic role clusters and alignment types via microrole coexpression tendencies. *Studies in Language* 38(3). 463–484.

- Hartmann, Iren. 2013. Hoocak Valency Patterns In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/hoocak>, Accessed on 2014-07-25)
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.
- Haspelmath, Martin. 2013. Argument indexing: a conceptual framework for the syntax of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages Across Boundaries: Studies in Memory of Anna Siewierska*, 197–226. Berlin: De Gruyter Mouton.
- Haspelmath, Martin. 2014. Arguments and adjuncts as language-particular syntactic categories and as comparative concepts. *Linguistic Discovery* 12(2). 3–11.
- Haspelmath, Martin, Andreea Calude, Michael Spagnol, Heiko Narrog & Elif Bamyacı. 2014. Coding causal-noncausal verb alternations: A form-frequency correspondence explanation. *Journal of Linguistics* 50(3). 587–625.
- Helbig, Gerhard & Wolfgang Schenkel. 1969. *Wörterbuch zur Valenz und Distribution deutscher Verben*. Leipzig: Bibliographisches Institut.
- Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.
- Khurshudian, Victoria & Michael Daniel. 2013. Eastern Armenian Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/eastern-armenian>, Accessed on 2014-07-25)
- Khurshudian, Victoria & Michael Daniel. 2015. Eastern Armenian. In Andrej Malchukov & Bernard Comrie (eds.), *Valency Classes in the World's Languages*. Berlin: De Gruyter Mouton.
- Kishimoto, Hideki & Taro Kageyama. 2013. Japanese (standard) Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/japanese-standard>, Accessed on 2014-07-25)
- Koch, Peter. 1994. Verbvalenz und Metataxe im Sprachvergleich. In Werner Thielemann & Klaus Welke (eds.), *Valenztheorie – Werden und Wirkung*, 109–124. Münster: Nodus.
- Langacker, Ronald W. 1988. The nature of grammatical valence. In Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*, 91–125. Amsterdam: Benjamins.
- Lehmann, Christian. 2013. Yucatec Maya Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/yucatec-maya>, Accessed on 2014-07-25)
- Levin, Beth & Malka Rappaport Hovav. 2005. *Argument Realization*. Cambridge: Cambridge University Press.
- Malchukov, Andrej & Alexander Jahraus. 2013. Russian Valency Patterns In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/russian>, Accessed on 2014-07-25)
- Mel'čuk, Igor' Aleksandrovič. 1974. *Opyt teorii lingvističeskix modelej "smysl-tekst": semantika, sintaksis*. Moskva: Nauka.
- Michaelis, Susanne Maria & the APiCS Consortium. 2013. Experiencer constructions: headache. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The Atlas of Pidgin and Creole Language Structures*, 260–263. Oxford: Oxford University Press.
- Moyse-Faurie, Claire. 2013. Xârâcùù Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/xaracuu>, Accessed on 2014-07-25)

- Nordhoff, Sebastian. 2013. Sri Lanka Malay Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/sri-lanka-malay>, Accessed on 2014-07-25)
- Peterson, David A. 2007. *Applicative Constructions*. Oxford: Oxford University Press.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Schikowski, Robert, Balthasar Bickel & Netra Prasād Paudyāl. 2013. Chintang Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/chintang>, Accessed on 2014-07-25)
- Schultze-Berndt, Eva. 2013. Jaminjung Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/jaminjung>, Accessed on 2014-07-25)
- Seifart, Frank. 2013. Bora Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/bora>, Accessed on 2014-07-25)
- Shibatani, Masayoshi & Ketut Artawa. 2013. Balinese Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/balinese>, Accessed on 2014-07-25)
- Svartvik, Jan. 1966. *On Voice in the English Verb*. The Hague: Mouton.
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.
- Vajda, Edward J. & Elena Kryukova. 2013. Ket Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/ket>, Accessed on 2014-07-25)
- Van Valin, Robert D., Jr. & Randy J. LaPolla. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.
- Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.
- Vater, Heinz. 1978. On the possibility of distinguishing complements and adjuncts. In Werner Abraham. 1978. *Valence, Semantic Case, and Grammatical Relations*, 21–45. Amsterdam: Benjamins.
- Waltereit, Richard. 1998. *Metonymie und Grammatik: Kontiguitätsphänomene in der französischen Satzsemantik*. Tübingen: Niemeyer.
- Watanabe, Honore. 2013. Sliammon Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/sliammon>, Accessed on 2014-07-25)
- Zhang, Guohua. 2013. Mandarin Chinese Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/mandarin>, Accessed on 2014-07-25)
- Zúñiga, Fernando & Leonel Lienlaf. 2013. Mapudungun Valency Patterns. In Iren Hartmann, Martin Haspelmath & Bradley Taylor (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://valpal.info/languages/mapudungun>, Accessed on 2014-07-25)



Andrej Malchukov

## **4 Valency classes and alternations: parameters of variation**

### **1 Introductory remarks**

As explained in Comrie et al. (this volume), the goal of the Leipzig Valency Classes Project is to study cross-linguistic universals and variation with respect to valency classes, as identified through distributional properties, in particular, through coding frames and alternations. The Leipzig Valency Classes Project follows up on some in-depth studies of European languages such as Levin (1993) on English and an earlier study by Apresjan (1969) on Russian. These studies have demonstrated that it is possible to arrive at a semantic classification of verbs through the study of their syntactic behavior, in particular, their potential to take part in alternations (such as the English dative alternation, locative alternation and passive alternation discussed by Levin). These results are highly interesting, but they raise the question of the extent to which valency classes arrived at on syntactic grounds are language-particular or universal. Clearly, some aspects of valency classes should be language-particular, otherwise valency dictionaries would not be necessary, yet other aspects are likely to show regularities, mirroring regularities in argument structure (role frames). The Leipzig Valency Classes Project is the first systematic large-scale attempt to explore cross-linguistic regularities in this domain. We establish valency classes on the basis of both the basic pattern (basic coding frame) and alternations—verb-coded (like passive alternation) as well as uncoded. In this respect the present project follows Apresjan (1969), which studies both coding frames and alternations, more closely than Levin (1993), which focuses on alternations exclusively. The study of coding frames is further informed by the typological studies of transitivity, such as the work by Tsunoda (1981, 1985) and subsequent work.

Recently, the approaches like Levin's that claim that the verb meaning can be recovered from its syntactic distribution have been challenged with the advent of Construction Grammar (Goldberg 1995). Construction Grammar differs from Levin's work and related approaches inasmuch as it does not assume that the verbal semantics is (fully) recoverable from syntactic distribution, since syntactic context can also lead to reinterpretation (alternative "construal" in the terms of Croft (2012)) of verbal meaning. Yet, generally, these approaches are compatible with the proviso that for Construction Grammarians syntactic distribution defines both the core meaning of a verb and its possible extensions ("construals") determined by the syntactic context. Clearly, the verb's "core" meaning largely determines its possible construals, manifested through different constructions. In the present volume, Goddard's paper stressing the importance of "lexico-semantic frames" for verb meaning is most in line with the Construction Grammar approaches.

In this chapter I will provide an overview of valency classes as identified through coding frames (section 2) and alternations (section 3). The final section (section 4) addresses the phenomenon of ambitransitivity and voice ambivalence. Generally, the aim of the chapter is to introduce the framework for the comparative study of the valency classes and to summarize some of the findings in other chapters of the volume dealing with individual languages. While this chapter focuses on variation in the domain of valency classes, the chapter by Wichmann (this volume) summarizes some of the quantitative results leading to the formulation of verb hierarchies based on patterns of alternations.

## 2 Valency classes: coding frames

### 2.1 Introduction

Let us consider first the question of universality of valency classes as far as coding is concerned. On the one hand, there are clearly certain cross-linguistic regularities in coding of verbs within semantic classes reflecting similarities in argument structure (i.e. semantic roles licensed by the verbal lexeme). Thus a case marker called “allative” is used for marking Goals in very different languages, but may be extended to semantic roles which are similar, such as Recipients and Addressees. Also the transitive patterns will be used in language after language with verbs like KILL and BREAK describing canonical transitive actions with a typical agent and patient argument<sup>1</sup>, but can be extended to other types of bivalent verbs. Such regularities are expected on iconicity assumptions, which predict iconic marking of an argument, but also allow that other semantically similar (micro)roles are coded in the same way.

On the other hand, valency classes cannot be universal, since after all the coding frames (as manifested in case and agreement) are language-particular. It is easy to show that the availability of certain coding frames will have an impact on valency classes, insofar as the number of valency classes is partially dependent on the “resources” available (in our case, different coding frames). Thus, the Northern Tungusic languages Even and Evenki feature 12–14 cases, most of which are also used to code arguments, while in Eastern Tungusic languages (such as Nanai) their number reduces to 8, and in Southern Tungusic Manchu it is further reduced to 5 cases, of which only 4 can be used to code arguments (Malchukov & Nedjalkov, this

---

<sup>1</sup> The latter statement is circular if a transitive pattern is defined on such verbs as BREAK and KILL; but it is not circular if a transitive pattern is described as a major or default pattern with bivalent verbs, as suggested by Lazard (1994) and others (see Haspelmath, this volume; and Blasi, this volume, for further discussion).

volume). This has an obvious impact on diversification of valency classes, which correspondingly reduce in number as one moves from Northern Tungusic to Southern Tungusic. Thus, in Manchu, the Dative is used as a default oblique case for the second argument of bivalent intransitives with different argument structures, which would be coded distinctly in Northern Tungusic. For example, the NOM-DAT pattern in Manchu is found with verbs of motion (like ‘come’), static location (like ‘sit’), and emotion (like ‘fear’), which would take three distinct patterns in Even (with an allative, locative, and instrumental object, respectively). Still more instructive in that respect are languages like Vafsi (an Iranian language; Stilo 2009) with a general oblique case which can be used on both obliques and direct objects (when the latter are animate). Thus, diversification/neutralization of valency classes depends on the set of coding frames available (coding resources), although this dependency is not deterministic (one should take into account that some case systems may be elaborate for spatial relations, which are of less relevance for argument coding, and also that a reduced case system may be compensated for by adpositional use at least for some argument types).<sup>2</sup>

Perhaps the most important source of variation in coding patterns across languages concerns different extensions of the transitive frame for different verb types. This variation will be discussed in more detail below in connection with Tsunoda’s transitivity hierarchy. At this point, suffice it to say that this extension varies broadly across languages, which also has an immediate impact on differentiation of valency classes. To a certain extent this variation also depends on the argument-coding type: thus, the inventory of ‘flags’ (dependent role marking by case and adposition) tends to be richer than inventory of ‘index-sets’ (head-marking through agreement/cross-referencing). Note that minimal case systems are rather exceptional (see Arkadiev 2009 on two-term case systems), while elaborate indexing systems are equally rare (see Kratochvíl 2011 on Abui, which features no less than seven index-sets). As a result, head-marking languages tend to have an extension of a transitive frame across different types of bivalent verbs: thus in Hoočək (Hartmann, this volume) effectively all types of bivalent verbs are transitive (featuring subject-object agreement) apart from the verbs of motion and static location, which take spatial arguments. Also “no-marking” languages of the isolating type seem to be more liberal in extending the transitive frame. Thus, Mandarin Chinese (Lu et al., this volume) can accommodate into the postverbal position many more arguments, including locative ones, than is usual for languages relying on dependent marking (see also (1)–(2) from Indonesian below). The same is true of Nlɪŋg (Ernszt et al., this volume), where the vast majority of two-argument verbs are transitive.

---

<sup>2</sup> See Say (2014) for further discussion of this point and for confirming evidence from European languages.

Yet, while such examples clearly show that valency classes are subject to cross-linguistic variation, this does not mean that this variation (in particular variation conditioned by availability of resources) is unconstrained. In what follows variation in the number of valency classes will be discussed in terms of the two competing motivations of Iconicity and Markedness. As will be shown below, both factors can be taken into account within the semantic map approach, which allows both for capturing universal regularities in coding and also for representing cross-linguistic variation.

## 2.2 Variation in valency classes: valency and transitivity: a first look

The most basic distinction in valency classes is the transitive/intransitive distinction. Recent typological research has generally confirmed the universality of this distinction, while certain other subclasses have been shown to be more subject to cross-linguistic variation. Thus, many languages also have zero valent verbs as distinct from intransitives, but this pattern is not universal (for example, Atoyebi, this volume, notes that verbs with no arguments do not exist in Yorùbá). Similarly, some languages may feature a dedicated syntactic frame for ditransitive verbs (taking Agent, Theme and Recipient arguments), a double object construction, yet the latter pattern is not universal (Malchukov et al. 2010). The transitive-intransitive distinction is cross-linguistically most robust: thus, even Sri Lanka Malay (SLM), which shows little differentiation in verb classes (Nordhoff, this volume), still distinguishes between transitive and intransitive verbs insofar as only the latter may take an accusative object. As in other languages with extensive pro-drop (use of zero anaphora), a transitive verb in SLM is better defined as a verb which **can** be used in the transitive frame, rather than a verb that **must** be used in the transitive frame (cf. Schikowski et al., this volume, on Chintang).

Yet some languages seem to challenge the universality of the transitive-intransitive distinction as well. Thus, in Jakarta Indonesian (Connors et al., this volume; cf. Gil 2009 and passim on Riau Indonesian), verbs of different semantic classes are surprisingly liberal in allowing different valency frames. Particularly surprising is that seemingly any monovalent verb may appear in the transitive frame with a bare postverbal object.

- |     |                                |                               |
|-----|--------------------------------|-------------------------------|
| (1) | Indonesian                     | (Connors et al., this volume) |
|     | (a) <i>Beli rumah.</i>         |                               |
|     | buy house                      |                               |
|     | ‘(He) bought a house.’         |                               |
|     | (b) <i>Tidur rumah.</i>        |                               |
|     | sleep house                    |                               |
|     | ‘(He) is sleeping in a house.’ |                               |



Yet even Indonesian allows us to distinguish between the two classes if one takes into account alternative case frames for different verb types. Indeed, those verbs that deviate from the transitivity prototype (as defined by Hopper & Thompson 1980 and Tsunoda 1981) allow for an alternative pattern involving various prepositions. For example, an alternative pattern for verbs of location like ‘sit’ is a prepositional one:

- (2) Indonesian (Connors et al., this volume)  
*Mulaintar malem, tidur di rumah Ayu.*  
 beginning\_later night sleep in house Ayu  
 ‘Starting from tonight, you sleep at Ayu’s house.’

Within the approach adopted here this variation between unmarked bivalent (“transitive”) pattern and the prepositional one can be described as an (unmarked) alternation. Depending on which pattern is taken as basic, it can be further described either as preposition dropping or as preposition insertion. (As noted by Connors et al. the patterns differ in their frequencies, so for different patterns different coding frames can be considered as basic). In both cases it is clear that even though Indonesian underdifferentiates the notional transitives conforming to the transitivity prototype from other types of bivalent verbs, their neutralization is resolved through alternations. Also, the transitive-intransitive opposition can be weakened in certain languages such as Chintang (see (50)–(51) below) and Central Alaskan Yupik (44)–(45), which feature “rampant ambitransitivity” (i.e. pervasive lability of verbs). However, rampant ambitransitivity does not eliminate the transitive/intransitive distinction as such, since it does not carry over from the lexical to the syntactic level: syntactically the transitive and intransitive patterns are clearly distinguished in these languages.

Apart from the traditional distinction between transitives and intransitives (with further subgroups of impersonals, and syntactic ditransitives), some linguists have proposed a more elaborate classification. Thus, Dixon (Dixon 1994; Dixon & Aikhenvald 2000) proposes the following classification, where E stands for an extension (roughly, an oblique argument):

- a) intransitive (S)
- b) transitive (A + O)
- c) extended intransitive (S + E)
- d) extended transitive (A + O + E)
- e) ditransitive (A + O + O)

An advantage of this classification is that it clearly distinguishes verbal valency and transitivity: thus, a class of bivalent intransitives (extended intransitives) is distinguished from (mono)transitives, and trivalent monotransitives (extended transitives) are distinguished from ditransitives. It also allows us to capture further variation between valency classes. Thus, one can characterize languages like Chati-