

Valency



Trends in Linguistics

Studies and Monographs 187

Editors

Walter Bisang

(main editor for this volume)

Hans Henrich Hock

Werner Winter

Mouton de Gruyter
Berlin · New York

Valency

Theoretical, Descriptive and Cognitive Issues

edited by

Thomas Herbst

Katrin Götz-Votteler

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

© Printed on acid-free paper which falls within the guidelines
of the ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

Valency : theoretical, descriptive, and cognitive issues / edited by Thomas Herbst, Katrin Götz-Votteler.

p. cm. – (Trends in linguistics. Studies and monographs ; 187)

Includes bibliographical references and index.

ISBN 978-3-11-019573-6 (hardcover : alk. paper)

1. Dependency grammar. 2. Cognitive grammar. 3. Contrastive linguistics. 4. Computational linguistics. 5. Semantics. I. Herbst, Thomas. II. Götz-Votteler, Katrin, 1975–

P162.V345 2007

415–dc22

2007031827

ISBN 978-3-11-019573-6

ISSN 1861-4302

Bibliographic information published by the Deutsche Nationalbibliothek

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

© Copyright 2007 by Walter de Gruyter GmbH & Co. KG, D-10785 Berlin

All rights reserved, including those of translation into foreign languages. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording or any information storage and retrieval system, without permission in writing from the publisher.

Cover design: Christopher Schneider, Berlin.

Printed in Germany.

Preface: Valency – theoretical, descriptive and cognitive issues

Thomas Herbst and Katrin Götz-Votteler

As with most other concepts in linguistics, in the discussion of valency one must distinguish between the linguistic phenomenon of valency on the one hand and the use of the term *valency* and the development of theoretical frameworks associated with it on the other. As far as the former is concerned, it is obvious that valency phenomena have been treated in linguistics under a variety of different labels ranging from *government* or *Rektion* in traditional grammar to *subcategorization* in generative frameworks or comparatively neutral labels such as *complementation* in descriptive grammars such as the *Comprehensive Grammar of the English Language*. Obviously, up to a point the use of different terms suggests different ways of viewing the phenomenon in question.

The notion of valency as such is generally linked with Tesnière's dependency grammar, although similar concepts had been put forward for example by Bühler (1934) and de Groot (1949).¹ It is probably fair to say that very significant contributions to the development of a theory of valency have been made by German linguistics since the 1960s. It is particularly the work of Gerhard Helbig and the emergence of a number of German valency dictionaries (Helbig and Schenkel 1969; Engel and Schumacher 1976; VALBU 2004) that are of importance here. Both lexicographically oriented and theoretical work on valency have resulted in an extensive discussion of criteria for the distinction between complements and adjuncts and a distinction between different types of complements with respect to their various degrees of obligatoriness. In recent years, the term valency has increasingly been used for the description of English, sometimes with explicit reference to the European tradition of valency theory and the concepts and criteria developed there,² sometimes just as a new term for complementation phenomena.

This volume comprises articles which deal with both the theoretical notion of valency and the analysis of valency phenomena. The articles in the first section, theoretical and descriptive aspects of valency, discuss the valency concept in its theoretical context (Peter Matthews) and the question of how valency phenomena can be described most appropriately with refer-

ence to certain distinctions such as complement inventories or valency patterns or semantic or syntactic valency (Thomas Herbst, Katrin Götz-Votteler). Other papers focus on different concepts of grammaticalization (Lene Schøsler, Dirk Noël) and particular problems of valency in synchronic and diachronic descriptions (Mechthild Habermann, Michael Klotz, Ilka Mindt). Finally, this section contains an outline of the treatment of valency phenomena and the underlying theoretical concept in the Berkeley FrameNet project (Charles Fillmore).

Section II focuses on the important issue of the role of valency phenomena in cognitive linguistics (Gert Rickheit and Lorenz Sichelschmidt, Rudolf Emons), where the acquisition of valency structures is of course a particularly important aspect (Heike Behrens).

Section III contains a number of papers with a contrastive orientation, which ranges from descriptive issues comparing different aspects of valency in English and German (Klaus Fischer, Irene Ickler, Brigitta Mittmann) and English, German and Norwegian (Stig Johansson) to a more pedagogically oriented account of valency errors in the performance of German and English learners (Ian Roe).

Finally, Section IV is concerned with computational aspects of valency analysis, where possible ways of using existing valency descriptions such as the *Valency Dictionary of English* (2004) as the basis for programs of word recognition are demonstrated (Dieter Götz, Ulrich Heid) and other approaches towards the automatic analysis of valency structures in computational linguistics are outlined (Roland Hausser, Besim Kabashi, Günther Görz and Bernd Ludwig).

The volume comprises papers given at a conference entitled *Valency: Valenz – Theoretical, Descriptive and Cognitive Issues* held at the Friedrich-Alexander-Universität Erlangen-Nürnberg in April 2005, which was supported by the Deutsche Forschungsgemeinschaft and the Dr.-Alfred-Vinzl-Stiftung. The editors would like to thank these institutions for the generous support they gave to the conference, Dr. Anke Beck for attending the conference and her support of the present volume, David Heath for his help and advice in all matters linguistic and Susen Schüller for her work on the index. Above all, our thanks go to all participants of the conference.

Notes

1. Cf. de Groot (1949/1964: 114-115) and Matthews (1981: 117). For the history of the concept of valency see Ágel (2000); for valency models in German linguistics see Herbst, Heath, and Dederding (1980) and Helbig (1992).

2. See, e.g., Emons (1974), Allerton (1982) and VDE (Herbst et al. 2004).

References

- Ágel, Vilmos
2000 *Valenztheorie*. Tübingen: Gunter Narr Verlag.
- Allerton, David J.
1982 *Valency and the English Verb*. London/New York: Academic Press.
- Bühler, Karl
1934 *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Fischer Verlag.
- Engel, Ulrich, and Helmut Schumacher
1976 *Kleines Valenzlexikon deutscher Verben*. Tübingen: Gunter Narr Verlag.
- Emons, Rudolf
1974 *Valenzen englischer Prädikatsverben*. Tübingen: Max Niemeyer Verlag.
- de Groot, Albert W.
1964 Reprint. *Structurele Syntaxis*. The Hague: Servire. Original edition, The Hague: Servire, 1949.
- Helbig, Gerhard
1992 *Probleme der Valenz- und Kasustheorie*. Tübingen: Max Niemeyer Verlag.
- Helbig, Gerhard, and Wolfgang Schenkel
1969 *Wörterbuch zur Valenz und Distribution deutscher Verben*. Leipzig: VEB Verlag Enzyklopädie.
- Herbst, Thomas, David Heath, and Hans-Martin Dederding
1980 *Grimm's Grandchildren. Current Topics in German Linguistics*. London/New York: Longman.
- Herbst, Thomas, David Heath, Ian Roe, and Dieter Götz (eds.)
2004 *A Valency Dictionary of English. A Corpus-Based Analysis of the Complementation Patterns of English Verbs, Nouns and Adjectives*. Berlin/New York: Mouton de Gruyter. [VDE]
- Matthews, Peter
1981 *Syntax*. Cambridge: Cambridge University Press.
- Quirk, Randolph, Sydney Greenbaum, Geoffrey Leech, and Jan Svartvik
1985 *A Comprehensive Grammar of the English Language*. London: Longman.
- Schumacher, Helmut, Jacqueline Kubczak, Renate Schmidt, and Vera de Ruyter (eds.)
2004 *VALBU – Valenzwörterbuch deutscher Verben*. Tübingen: Gunter Narr Verlag [VALBU].
- Tesnière, Lucien
1959 *Éléments de Syntaxe Structurale*. Paris: Klincksieck.

Contents

Preface	v
Valency – theoretical, descriptive and cognitive issues <i>Thomas Herbst and Katrin Götz-Votteler</i>	
Section 1: Theoretical and descriptive aspects of valency	
The scope of valency in grammar <i>Peter Matthews</i>	3
Valency complements or valency patterns? <i>Thomas Herbst</i>	15
Describing semantic valency <i>Katrin Götz-Votteler</i>	37
The status of valency patterns <i>Lene Schøsler</i>	51
Verb valency patterns, constructions and grammaticalization <i>Dirk Noël</i>	67
Aspects of a diachronic valency syntax of German <i>Mechthild Habermann</i>	85
The valency of experiential and evaluative adjectives <i>Ilka Mindt</i>	101
Valency rules? The case of verbs with propositional complements <i>Michael Klotz</i>	117
Valency issues in FrameNet <i>Charles J. Fillmore</i>	129

Section 2: Cognitive issues and valency phenomena

Valency and cognition – a notion in transition 163
Gert Rickheit and Lorenz Sichelschmidt

Valency grammar in mind 183
Rudolf Emons

The acquisition of argument structure 193
Heike Behrens

Section 3: Contrastive aspects of valency

Valency and the errors of learners of English and German 217
Ian Roe

Temporary ambiguity of German and English term complements 229
Klaus Fischer

Sentence patterns and perspective in English and German 253
Irene Ickler

Contrasting valency in English and German 271
Brigitta Mittmann

Valency in a contrastive perspective: Structure and use 287
Stig Johansson

Section 4: Computational aspects of valency analysis

Valency and automatic syntactic and semantic analysis 309
Dieter Götz

Handling valency and coordination in Database Semantics 321
Roland Hausser

Pronominal clitics and valency in Albanian: A computational linguistics perspective and modelling within the LAG-Framework <i>Besim Kabashi</i>	339
The practical use of valencies in the Erlangen speech dialogue system CONALD <i>Günther Görz and Bernd Ludwig</i>	353
Valency data for Natural Language Processing: What can the <i>Valency Dictionary of English</i> provide? <i>Ulrich Heid</i>	365
Subject index	383
Author index	391

Section 1

Theoretical and descriptive aspects of valency

Pronominal clitics and valency in Albanian: A computational linguistics perspective and modelling within the LAG-Framework¹

Besim Kabashi

1. Theoretical aspects: Pronominal clitics and valency

When reading Albanian texts, elements like *e, i, ia* cannot be overlooked. If they immediately precede the verb or if they are part of the verb form in non-negated imperative sentences, they are instances of *trajtat e shkurtra të përëmrave vetorë* [short forms of personal pronouns; Domi 1995; 1997]. Buchholz and Fiedler (1987) call them *Objektszeichen* [object signs], Kalluli (1995) names them *clitics*. Like Newmark, Hubbard, and Prifti (1982), we will use the term *pronominal clitics (pCl)*.

1.1. Formal properties of pronominal clitics in Albanian

The following table gives an overview of the Albanian pCl forms:

Table 1. The personal pronouns in Albanian (left column) with their clitic forms (right column)

	1st Person				2nd Person			
	Singular		Plural		Singular		Plural	
Nom.	unë	-	ne	-	ti	-	ju	-
Dat.	mua	më	neve	na	ty	të	juve	ju
Acc.	mua	më	ne	na	ty	të	ju	ju
3rd Person								
	Singular				Plural			
	Masculine		Feminine		Masculine		Feminine	
Nom.	ai	-	ajo	-	ata	-	ato	-
Dat.	atij	i	asaj	i	atyre	u	atyre	u
Acc.	atë	e	atë	e	ata	i	ato	i

PCIs in Albanian indicate the person and number of the respective objects of the verb. They occur in dative and accusative case: dative pCIs can be combined with accusative pCIs. In most cases this results in *amalgamated forms* (*crasis*). For example *më* and *e* amalgamate to *ma*. The combination of *na* with *e*, *i*, or *u* does not involve amalgamation but *concatenation*, in which case the dative precedes the accusative: *na e*, *na i*, and *na u*.

There are two morphosyntactic types of pCIs: bound and free. *Bound forms* occur within positive (non-negated) imperatives after the verb stem (*enclitic position*). In the plural they appear between the verb stem and the suffix *ni*, cf. the following example:

Example 1. Pronominal clitics as bound forms²

(1)	sill	V _{Stem}	No pCI; Sg;	‘bring’
a.	sille	V _{Stem} +pCl _A	pCl _A , Sg;	‘bring it’
b.	sillma	V _{Stem} +pCl _{D+A}	pCl _{D+A} , Sg;	‘bring me it’
c.	sillni	V _{Stem} +ni	No pCI; Pl;	‘bring’
d.	silleni	V _{Stem} +pCl _A +ni	pCl _A , Pl;	‘bring it’
e.	sillmani	V _{Stem} +pCl _{D+A} +ni	pCl _{D+A} , Pl;	‘bring me it’

In cases with negation particles, pCIs cannot occur as bound forms, e.g. *Mos e sill!* [‘(You: Sg) do not **it** bring!’], i.e. ‘Do not bring it!’ and *Mos e sillni!* [‘(You: Pl) do not **it** bring!’], i.e. ‘Do not bring it!’]. As *free forms*, pCIs always precede the finite verb (*proclitic position*).³ Word order in Albanian is relatively free. Therefore, both subject and objects may appear in front of the verb complex, which allows for a large number of different sentence patterns for a given verb; the order of elements within the verb complex is fixed. If all of these elements occur, they are in the following sequence: negation, future marker/modal verb, subjunctive particle, pCIs, finite and non-finite verb. The subject can be deduced from verb inflection and thus be left out.

PCIs in Albanian appear either in addition to objects in the sentence (*object doubling*) or they replace objects in the sentence, so that the objects themselves can be left out (*object elimination*).

1.2. Object doubling

The following example (2b) shows the doubling of the accusative object:

Example 2. Accusative object doubling⁴

(2) S+V

<i>Ne</i>	<i>shohim.</i>
Subject:	Verb: Itr
N Pl1	Pl1 Ind Prs Act Nad
we	see
'We (are able to) see.'	

S+V+O_A

a. <i>Ne</i>	<i>shohim</i>	<i>studentët/ata.</i>
	Verb: Tr	O: A
		A Pl3 M
we	see	the students/them

S+pCl_A+V+O_A

b. <i>Ne</i>	<i>i</i>	<i>shohim</i>	<i>studentët/ata.</i>
	pCl: A		
	Pl3		
we	them	see	the students/them
'We see the students/them.'			

The following example shows the doubling of dative (3, 3a, and 3b) and accusative objects (3a and 3c).

Example 3. Object doubling in dative and accusative⁵

(3) S+pCl_D+V+O_D+O_A

<i>Ne</i>	<i>u</i>	<i>dhamë</i>	<i>studentëve/atyre</i>	<i>librat/ato.</i>
Subject:	pCl: u	Verb: Tr	O: D	O: A
N Pl1	D Pl3	Pl3 Ind Aor Act	D Pl3	A Pl Det
		Nad		
we	them	gave	the students/ them	the books/ them
'We gave the students/them the books/them.'				

S+pCl_{D+A}+V+O_D+O_A

a. <i>Ne</i>	<i>ua</i>	<i>dhamë</i>	<i>studentëve/atyre</i>	<i>librat/ato</i>
	pCl: u+i			
	D Pl3 +			
	A Pl3			
we	them + them	gave	the students/ them	the books/ them
'We gave the students/them the books/them.'				

S+pCl _{D+A} +V+O _D				
b.	<i>Ne</i>	<i>ua</i>	<i>dhamë</i>	<i>studentëve/atyre</i>
	we	them + them	gave	the students/ them
	‘We gave the students/them them.’			
S+pCl _{D+A} +V+O _A				
c.	<i>Ne</i>	<i>ua</i>	<i>dhamë</i>	<i>librat/ato</i>
	we	them + them	gave	the books/them
	‘We gave them the books/them.’			

Sentence (3) has two objects, one in dative and one in accusative case, and one pCl in dative case, which doubles the dative object. Whenever the verb has a dative valency, the dative pCl cannot be left out, regardless of the presence of a dative object in the sentence. In sentence (3a), both objects are doubled by means of the respective pCls. Only one object is doubled by the amalgamated pCl in sentences (3b) and (3c). Object doubling is also possible in imperatives, e.g. *Sillma librin!* [‘Bring+me+it the book!’, ‘Bring me the book!’], i.e. doubling of accusative object.

When the verb has a first or second person accusative or dative object, the pCls of the first and second person cannot be omitted, e.g. *Studentët të kuptojnë ty.* or *Studentët të kuptojnë.* [‘The students understand you’]. The form **Studentët kuptojnë ty.* is ungrammatical.

1.3. Object elimination

If a pCl occurs, the corresponding object can be left out.⁶ This phenomenon has been called *Objektseliminierung* [object elimination] by Buchholz (1977), and Buchholz and Fiedler (1987). As there is no established English term, we will use the term *object elimination* as a translation of *Objektseliminierung*.⁷ Example sentences where the objects are left out are (3b) and (3c). A sentence consisting only of pCl(s) and verb can be syntactically well-formed, cf. sentences (4a) and (4b) in the following example, where the subject is optional.

Example 4. Object elimination

(4)	*S+V+O _D +O _A			
	* <i>Ne</i>	_	<i>dhamë</i>	<i>studentëve/atyre</i>
	we		gave	the students/them
				the books/ them
	*‘We gave _ the students/them the books/them.’			

S+pCl_{D+A}+V

a.	<i>Ne</i>	<i>ua</i>	<i>dhamë.</i>
	we	them +	gave
		them	
	‘We gave them to them.’		

S+pCl_A+V

b.	<i>Ne</i>	<i>i</i>	<i>shohim</i>
		them	see
	‘We see them.’		

The following patterns⁸ are possible for doubling and elimination of accusative complements: S+V+O_A (object without pCl), S+pCl_A+V+O_A (object doubling) and S+pCl_A+V (object elimination). For dative complements, the following patterns apply: S+pCl_D+V+O_D (object doubling) and S+pCl_D+V (object elimination). As shown in the examples above, these patterns can be combined with each other.

1.4. Pronominal clitics as valency fillers

Sentences (3b-c) and (4a-b) show the ability of pCls to function as *valency fillers*, as they can replace objects. The grammatical information of the pCl is sufficient to fill the valency of the verb; only the lexical content is missing, which has to be recoverable from the linguistic or extralinguistic context. Thus, if pCls occur, objects can be left out without making the sentence ungrammatical.

If objects are left out, the omission of pCls leads to the selection of a different valency pattern. This can (but does not have to) indicate the use of the verb in a different sense, cf. the difference of valency and meaning between sentences (2) and (2a), (2) and (2b), as well as in the following example adapted from Buchholz, Fiedler, and Uhlisch (1993):

Example 5. Different patterns/meanings of the verb *flas*⁹

(5) S+V

Ai	flet.
Subject:	Verb: Itr
N Sg3 M	Sg3 Ind Prs Act Nad
He	speak
‘He speaks.’	

S+pCl _D +V+O _D						
a.	Ai	i	flet	atij.		
		pCl: D Sg3	Verb: Tr	O: D D Sg3 M		
	He	him	scold	him		
	‘He scolds him.’					
S+pCl _D +A+V+O _D +O _A						
b.	Ai	ia	flet	atij	një	libër.
		pCl: D Sg3 + A Sg3				A Sg1 M Undet
	He	him + it	promise	him	one	book
	‘He promises him a book.’					

The role of pCIs in the context of valency is different from the substitution of objects by pronouns, despite the fact that the pCIs are short forms of personal pronouns. Personal pronouns, however, offer a more precise description of case and number (first and second person plural) and gender (third person singular) than pCIs. First of all, the position of the pronoun is the same as the position of the object if a pronoun substitutes the object. A pCI has a fixed position in the verbal complex, regardless of the position of the object it replaces. While pronouns always replace objects, pCIs are capable of either eliminating or doubling the objects they refer to. The appearance of the dative pCI is obligatory while the substitution of the dative object by pronouns is optional. The substitution of objects by pronouns does not change the structure of the sentence (which means the sentence matches to the same pattern), i.e. the O_A is only realized by a pronoun instead of a noun phrase, but the pattern S+V+O_A remains the same. If a pCI appears, however, the pattern is changed, e.g. from S+V+O_A to S+pCl_A+V+O_A or to S+pCl_A+V. As shown above, pCIs and pronouns can (and in some cases must, cf. section 1.2.) occur together in a sentence.

As the dative pCI cannot be left out without making the sentence ungrammatical, it always shares the dative valency slot with the dative object if the object is present. The pCI fills the valency slot itself, if the dative object is eliminated. In the case of the accusative, there is one more option: the object alone fills the valency slot of the verb, the object and pCI share the valency slot, or the pCI alone fills the valency slot. Thus in this analysis, there is a minimum number of required complements of 1 and a maximum number of required complements of 2 for the dative and the accusative valency slot (by analogy to Herbst et al. [2004], where the minimum and maximum valency of each verb is indicated); cf. the patterns S+V+O_A, S+pCl_A+V and S+pCl_D+V for minimum, and S+pCl_A+V+O_A and

S+pCl_D+V+O_D for maximum complements.

The fact that the dative pCl is obligatory even if the dative object is also present can be accounted for using the concept of structural necessity presented by Herbst (1999), which means that it is obligatory only on the syntactic level, not as a lexical property. As pCls can be correctly predicted by a set of rules, they are no idiosyncratic properties of lexical units and thus need not be specified in the valency frame of a lexical unit. It is enough to specify valency slots such as dative or accusative. Object doubling is always optional on the syntactic level; here semantic factors (needed for lexical content) as well as pragmatic factors (emphasis) decide whether doubling takes place or not. Additional functions and properties of pCls will not be discussed here.¹⁰

The properties of the pCl in the context of valency can be summarized as follows: in case of object elimination, the pCls indicate the verb valency and function as valency fillers, whereas in case of object doubling, both pCls and objects share the valency slot of the verb.

2. Practical aspects: The computational model

The formalism used in the following computational model is *Left Associative Grammar* (LAG).¹¹ LAG was developed as a formalism for the SLIM (*Surface compositional Linear Internal Matching*) language theory intended to model and reconstruct natural language communication on a computer.

2.1. The formalism

The formalism works according to the principle of possible continuations whereby every grammar rule concatenates the sentence start read so far with the next word. The result of this concatenation becomes the new sentence start to be concatenated with the next word again; this procedure is repeated until the end of the input is reached. For example the input *a b c d* is parsed in the steps $a+b \rightarrow ab$; $ab+c \rightarrow abc$; $abc+d \rightarrow abcd$, summarized as $((a+b)+c)+d \rightarrow abcd$. Ambiguities are handled by tracing several derivation paths in parallel. The most varied sentence types can be modeled with linear effort regarding complexity and parse time, because only the possible continuations matter in each LAG rule.

2.2. An example

Below, an algorithm for the treatment of dependencies between verb, pCIs and objects is outlined using the sentence *Ai na i dha librat*. [‘He gave us the books’].¹² It starts with the first word form and the matching rule *Subject* (start rule).

Ai_N Sg^3 M

❶ RULE: Subject
 LVF = <Subject>;
 Follow RULE pCl_D ;

The subject is added to the list of valency fillers (LVF). In case of an omitted subject, the algorithm starts with the next rule, i.e. the parser searches for a rule with a matching start pattern. *Follow* means continuations, i.e. the next applicable rule, in this case pCl_D . When reading a word form, information required for recognition of the verb complex and for dealing with valency and congruency is provided by the lexicon or a morphological analysis component.

Ai + $na_{D\ P11}$

❷ RULE: pCl_D
 LVF = <Subject, $pCl_D_Object_elimination$ >;
 Follow RULE pCl_A ;

pCl_D is added to the LVF and labeled *Object_elimination* because the dative object can be left out. Since at this point it is not clear yet whether the actual pattern is *elimination* or *doubling*, the modification of the attribute to $pCl_D_Object_doubling$ can be done when a matching object is encountered. If no matching object is found, the attribute is just left at the value set here. Here, two non-amalgamated (concatenated) pCIs are used to demonstrate the *canceling* (filling) of valencies (valency slots). Amalgamated pCIs have to be analyzed in the morphology component.

Ai na + $i_{A\ P13}$

❸ RULE: pCl_A
 LVF = <Subject, $pCl_D_Object_elimination$, $pCl_A_Object_elimination$ >;
 Follow RULE *finVerb*;

pCl_A is added to the LFV just like *pCl_D* in the previous rule, also with the *Object_elimination* label.

Ai na i	+	dha _{Sg3 Aor Ind Act Nad <<D, A>, <A>>}
---------	---	--

④ RULE: *finVerb*
 LVF = <*pCl_D_Object_elimination*, *pCl_A_Object_elimination*>;
 Follow RULE *Accusative_object*;

The verb is read in and checked for agreement with the subject. The subject valency in LVF is canceled, the *pCl*s, however, remain in the LVF because they can share their valency slot with an object and thus cannot be canceled before the respective object has been read or the sentence is finished. Because *pCl*s always precede the verb, a (minimum) valency pattern that selects one or more of the possible lexical readings can be already constructed at this point.

Ai na i dha	+	librat _{A Pl3}
-------------	---	-------------------------

⑤ RULE: *Accusative_object*
 Replace *pCl_A_Object_elimination* by *pCl_A_Object_doubling*;
 LVF = <*pCl_D_Object_elimination*, *pCl_A_Object_doubling*>;
 Follow RULE *Punctuation*;

An accusative object is read in and modifies the *pCl_A_Object_elimination* entry in the LFV to *pCl_A_Object_doubling*.

Ai na i dha librat	+	.
--------------------	---	---

⑥ RULE: *Punctuation*
 LVF = <*pCl_D_Object_elimination*, *pCl_A_Object_doubling*>;
 RESULT = *pCl_D_Object_elimination*, *pCl_A_Object_doubling*;

The end of the sentence has been reached, the result is in LVF: object doubling for accusative and object elimination for dative.

Only the path that actually parses this sentence is shown above; other possible paths were left out. After various rule applications, several continuations (parallel paths) would be possible and would have to be tried out by the algorithm. For example, after the rule *finVerb* the sentence may continue with punctuation, dative object, accusative object, a preposition, an adjunct etc.

As demonstrated in the example model, valencies can be canceled immediately when a potential valency filler has been read. Another possibility is end canceling, where the properties of all word forms read are collected in attribute-value matrices and are not canceled until a punctuation mark signals the end of the sentence. This has the advantage of transparency when handling constructions such as subclauses with the verb in final position, but one disadvantage is that ungrammatical constructions may not be rejected before they have been completely parsed.

When reading amalgamated pCl word forms such as *ma*, *ta*, *t'ia* etc., it is important to read and process the individual parts of the morphosyntactic information, e.g. *ma* ($m\ddot{e}_D Sg1 + e_A Sg3$). In this way, an amalgam can replace one object (object elimination) and double another at the same time, cf. sentences (3b) and (3c).

The following figure shows an example of a morphological analysis of the imperative one-word-sentence *sillmani*. As shown under the *Clitic* attribute, the enclitic form *ma* has been recognized as an amalgam of *mē* and *e*.¹³

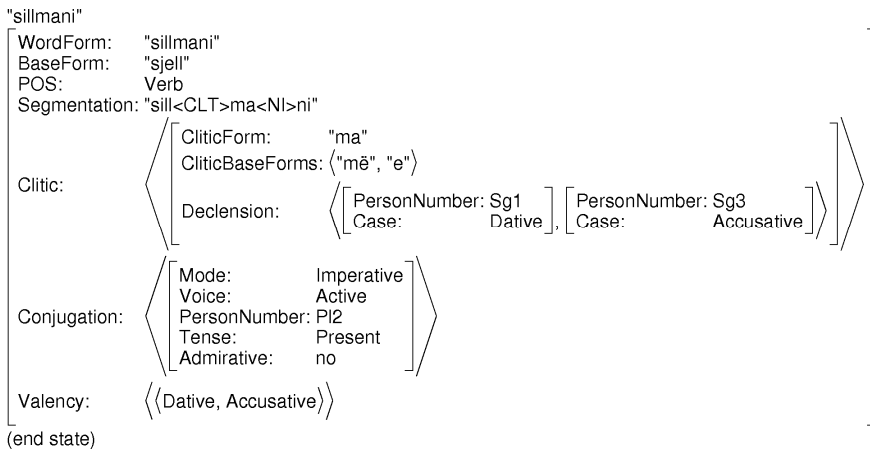


Figure 1. Morphological analysis of the word form *sillmani* from Kabashi (2003)

Here, valencies from the corresponding attribute are canceled with the matching cases from the clitic attribute *Declension*. It is necessary to check the information of lexical entries and morphological analysis to be able to select or construct the correct valency pattern, particularly with regard to the pCl's that may be used with a verb. Information on possible valency patterns of a verb comes from its lemma in the base-form lexicon.

Figure 2 shows the result of syntax analysis and valency handling of the sentence *Ai na i dha librat.* treated above. The dative object is missing there, cf. the *FilledValencyFromObjects* attribute. A dative object was expected according to the verb's valency pattern but was not present. A dative pCl was found and thus the sentence can be analyzed as well-formed, cf. the *FilledValencyFromClitics* and the *DativeSlot* attributes. On the other hand the fact that both an accusative pCl and an accusative object are found, leads to object doubling, cf. the corresponding attributes, *FilledValencyFromObjects* and *AccusativeSlot*. The *Index* attribute indicates the position of a word form in the analyzed sentence. The actual pattern is derived from the attributes *Clitic_D*, *Clitic_A*, *Verb*, and *Object_A*. The *Meaning* attribute contains the meaning of the verb in the currently selected pattern.

"Ai na i dha librat ."															
Subject:	<table border="0"> <tr><td>POS:</td><td>Pronoun</td></tr> <tr><td>Lemma:</td><td>"ai"</td></tr> <tr><td>Declension:</td><td>{ PersonNumber: Sg3 Case: Nominative }</td></tr> <tr><td>Gender:</td><td>Masculine</td></tr> <tr><td>Index:</td><td>1</td></tr> </table>	POS:	Pronoun	Lemma:	"ai"	Declension:	{ PersonNumber: Sg3 Case: Nominative }	Gender:	Masculine	Index:	1				
POS:	Pronoun														
Lemma:	"ai"														
Declension:	{ PersonNumber: Sg3 Case: Nominative }														
Gender:	Masculine														
Index:	1														
Clitic_D:	<table border="0"> <tr><td>POS:</td><td>Clitic</td></tr> <tr><td>Lemma:</td><td>"na"</td></tr> <tr><td>Declension:</td><td>{ PersonNumber: P11 Case: Dative }</td></tr> <tr><td>Index:</td><td>2</td></tr> </table>	POS:	Clitic	Lemma:	"na"	Declension:	{ PersonNumber: P11 Case: Dative }	Index:	2						
POS:	Clitic														
Lemma:	"na"														
Declension:	{ PersonNumber: P11 Case: Dative }														
Index:	2														
Clitic_A:	<table border="0"> <tr><td>POS:</td><td>Clitic</td></tr> <tr><td>Lemma:</td><td>"i"</td></tr> <tr><td>Declension:</td><td>{ PersonNumber: P13 Case: Accusative }</td></tr> <tr><td>Index:</td><td>3</td></tr> </table>	POS:	Clitic	Lemma:	"i"	Declension:	{ PersonNumber: P13 Case: Accusative }	Index:	3						
POS:	Clitic														
Lemma:	"i"														
Declension:	{ PersonNumber: P13 Case: Accusative }														
Index:	3														
Verb:	<table border="0"> <tr><td>WordForm:</td><td>"dha"</td></tr> <tr><td>BaseForm:</td><td>"jep"</td></tr> <tr><td>POS:</td><td>Verb</td></tr> <tr><td>Conjugation:</td><td>{ Mode: Indicative Voice: Active PersonNumber: Sg3 Tense: Aorist Admirative: no }</td></tr> <tr><td>Valency:</td><td>{ (Dative, Accusative) }</td></tr> <tr><td>Meaning:</td><td>{ "to give" }</td></tr> <tr><td>Index:</td><td>4</td></tr> </table>	WordForm:	"dha"	BaseForm:	"jep"	POS:	Verb	Conjugation:	{ Mode: Indicative Voice: Active PersonNumber: Sg3 Tense: Aorist Admirative: no }	Valency:	{ (Dative, Accusative) }	Meaning:	{ "to give" }	Index:	4
WordForm:	"dha"														
BaseForm:	"jep"														
POS:	Verb														
Conjugation:	{ Mode: Indicative Voice: Active PersonNumber: Sg3 Tense: Aorist Admirative: no }														
Valency:	{ (Dative, Accusative) }														
Meaning:	{ "to give" }														
Index:	4														
Object_A:	<table border="0"> <tr><td>WordForm:</td><td>"librat"</td></tr> <tr><td>BaseForm:</td><td>"libër"</td></tr> <tr><td>POS:</td><td>Substantive</td></tr> <tr><td>Declension:</td><td>{ Num: Pl Definite: yes Case: Accusative }</td></tr> <tr><td>Gender:</td><td>Masculine</td></tr> <tr><td>Index:</td><td>5</td></tr> </table>	WordForm:	"librat"	BaseForm:	"libër"	POS:	Substantive	Declension:	{ Num: Pl Definite: yes Case: Accusative }	Gender:	Masculine	Index:	5		
WordForm:	"librat"														
BaseForm:	"libër"														
POS:	Substantive														
Declension:	{ Num: Pl Definite: yes Case: Accusative }														
Gender:	Masculine														
Index:	5														
FilledValencyFromClitics:	{ Dative, Accusative }														
FilledValencyFromObjects:	{ Accusative }														
AccusativeSlot:	Filled_2_pCl_O														
DativeSlot:	Filled_1_pCl														
(end state)															

Figure 2. Syntactical analysis of the sentence *Ai na i dha librat.*

3. Conclusion

Acting as object substitutes (object elimination), pronominal clitics determine the verb's valency and assume the role of valency fillers. They play an important role in distinguishing between various possible valency frames of a verb and thus between different meanings. In the case of object doubling, they merely function as semantic and pragmatic markers for their respective object.

PCIs supply grammatical information that can be very useful both in natural language communication as well as in natural language processing, e.g. for processing discontinued sentences during turn-taking in dialogue analysis.

In spite of the treated phenomenon's complexity, an efficient implementation of a parser is possible using the LAG formalism. As shown in the algorithm, canceling (filling) of verb valencies (valency slots) is easily solved despite the multitude of possible combinations and the consequently large number of sentence patterns.

Notes

1. For comments on the draft of this paper I would like to thank Jörg Kapfer, Matthias Bethke, and Peter Uhrig (all Friedrich-Alexander-Universität Erlangen-Nürnberg). Only the properties relevant to verb valency and computational modeling will be treated here. For further information on pronominal clitics left out here, see Buchholz (1977), Buchholz and Fiedler (1987), Domi (1995; 1997), Kallulli (1995), and Newmark, Hubbard, and Prifti (1982).
2. Here, the following abbreviations are used: A=Accusative, D=Dative, pCl_{D+A} =Amalgam of pCl_D and pCl_A , Pl=Plural, Sg=Singular, and V=Verb. Forms like *silleni* have alternatives in the form *sillnie*.
3. In subjunctive clauses, pCIs are positioned after the subjunctive and future particles and precede the finite verb. In this case the pCI can be combined with these particles in one word, e.g. the amalgam *t'i* consisting of the subjunctive particle *të* and the pCI *i*.
4. Act=Active, Ind=Indicative, Itr=Intransitive, M=Masculine, Nad=Not admira-tive, N=Nominative, O=Object, Prs=Present, S=Subject, and Tr=Transitive.
5. Aor=Aorist (definite past), and Det=Determined.
6. An exception is the reflexive use of verbs, e.g. *Ai e_(pCl: A) lavdëron vetën_(O: A, Reflexive)* ['He (**him**) praised **himself**.', i.e. 'He praised himself.'], vs. *Ai e lavdëron*. ['He (**him**) praised.', i.e. 'He praised him.'].
Reflexive)
7. This term is described in Buchholz and Fiedler (1987) as *Vertretung des Objekts* (which might be translated as *object replacement* or *object substitution*).

8. As we focus on the influence of pronominal clitics on verb valency, only direct and indirect objects (without prepositions) are treated here.
9. Undet=Undetermined. Other patterns of the verb *flas*, e.g. pCl(s) + *flas* + preposition, and other meanings not presented in Buchholz, Fiedler, and Uhlisch (1993) are not treated here. For more patterns/meanings cf. Kostallari (1980), where the pattern/meaning from sentence (5b) is marked as a dialect form. Cf. also Qesku (1999). Toçi (2002) does not list this pattern.
10. Buchholz and Fiedler (1987: 445–446) have described a group of verbs which cooccur with „„pleonastisch‘ verwendete Objektszeichen“ [“pleonastically used” pronominal clitics]. These pCIs are „nicht ... systematisch aus einem zur Grundstruktur gehörenden dir[ekten] Obj[ekt] ableitbar“ [not derived in a systematic way from a direct object belonging to the base structure]. Thus sentences in which only one dative pCI would be expected can have two pCIs or an amalgamated one consisting of a dative and an accusative pCI, like for example *Ia*_(pCI: D+A) *hipi*_(V) *kalit*_(O: D). [‘He mounted the horse’]; in this example the subject is left out. These verbs cooccur only with specific lexical entries (and so they must be marked in the lexicon to be considered during automatic syntactic and semantic analysis). According to Buchholz and Fiedler (1987: 445), this group contains, among others, the following verbs: *arrin, del, fillon, hipën, hyn, kërcen, mbath, merr, nis, pëlçet, shtron, and thotë*.

PCIs can also occur as ethical datives, cf. Buchholz and Fiedler (1987: 447–448). In this case, the verb valency must be treated differently as well.

11. For the formal definition of LAG see Hausser (1992; 2001a).
12. The version illustrated here is only a simple LAG. Information about valency in *Database Semantics (DBS)* can be found in the article *Handling valency and coordination in database semantics* by R. Hausser in this volume. For information on DBS see Hausser (2001b).
13. During the syntactic analysis both pCIs would fill the corresponding attribute in the sentence structure while reading the verb.

References

- Buchholz, Oda
 1977 *Zur Verdoppelung der Objekte im Albanischen*. Linguistische Studien, Reihe A, Arbeitsberichte 34. Berlin: Akademie der Wissenschaften der DDR.
- Buchholz, Oda, and Wilfried Fiedler
 1987 *Albanische Grammatik*. Leipzig: VEB.
- Buchholz, Oda, Wilfried Fiedler, and Gerda Uhlisch
 1993 *Wörterbuch Albanisch–Deutsch*. München: Langenscheidt.
- Domi, Mahir (ed.)
 1995 *Gramatika e gjuhës shqipe*. Vëllimi I – *Morfologjia* [Grammar of the Albanian Language. Vol. 1: Morphology]. Tiranë: Akademia e Shkencave e Republikës së Shqipërisë.

- 1997 *Gramatika e gjuhës shqipe*. Vëllimi II – *Sintaksa* [Grammar of the Albanian Language. Vol. 2: Syntax]. Tiranë: Akademia e Shkencave e Republikës së Shqipërisë.
- Hausser, Roland
- 1992 Complexity in left-associative grammar. In *Theoretical Computer Science*. 106 (2): 283–308.
- 2001a *Foundations of Computational Linguistics. Human-Computer Communication in Natural Language*. 2d ed. Berlin/New York: Springer.
- 2001b Database semantics for natural language. In *Artificial Intelligence* 130: 27–74.
- 2007 Handling valency and coordination in database semantics. This volume.
- Herbst, Thomas
- 1999 English valency structures – A first sketch. *Erfurt Electronic Studies in English* (EESE).
http://webdoc.gwdg.de/edoc/ia/eese/artic99/herbst/6_99.html.
- Herbst, Thomas, David Heath, Ian F. Roe, and Dieter Götz (eds.)
- 2004 *A Valency Dictionary of English. A Corpus-Based Analysis of the Complementatation Patterns of English Verbs, Nouns and Adjectives*. Berlin/New York: Mouton de Gruyter.
- Kabashi, Besim
- 2003 Automatische Wortformererkennung für das Albanische. Master's thesis, Computational Linguistics. Universität Erlangen-Nürnberg.
- Kallulli, Dalina
- 1995 *Clitics in Albanian*. (Working Papers in Linguistics 24.) Trondheim: University of Trondheim.
- Kostallari, Androkli (ed.)
- 1980 *Fjalor i gjuhës së sotme shqipe* [Dictionary of Contemporary Albanian Language]. Tiranë: Akademia e Shkencave e RPS të Shqipërisë.
- Newmark, Leonard, Philipp Hubbard, and Peter Prifti
- 1982 *Standard Albanian. A Reference Grammar for Students*. Stanford: Stanford University Press.
- Qesku, Pavli
- 1999 *Fjalor Shqip–Anglisht. Albanian–English Dictionary*. Tiranë: EDFA.
- Toçi, Fatmir (ed.)
- 2002 *Fjalor i shqipes së sotme* [Dictionary of Contemporary Albanian]. 2d ed. Tiranë: Toena.

Pronominal clitics and valency in Albanian: A computational linguistics perspective and modelling within the LAG-Framework¹

Besim Kabashi

1. Theoretical aspects: Pronominal clitics and valency

When reading Albanian texts, elements like *e, i, ia* cannot be overlooked. If they immediately precede the verb or if they are part of the verb form in non-negated imperative sentences, they are instances of *trajtat e shkurtra të përëmrave vetorë* [short forms of personal pronouns; Domi 1995; 1997]. Buchholz and Fiedler (1987) call them *Objektszeichen* [object signs], Kalluli (1995) names them *clitics*. Like Newmark, Hubbard, and Prifti (1982), we will use the term *pronominal clitics (pCl)*.

1.1. Formal properties of pronominal clitics in Albanian

The following table gives an overview of the Albanian pCl forms:

Table 1. The personal pronouns in Albanian (left column) with their clitic forms (right column)

	1st Person				2nd Person			
	Singular		Plural		Singular		Plural	
Nom.	unë	-	ne	-	ti	-	ju	-
Dat.	mua	më	neve	na	ty	të	juve	ju
Acc.	mua	më	ne	na	ty	të	ju	ju
3rd Person								
	Singular				Plural			
	Masculine		Feminine		Masculine		Feminine	
Nom.	ai	-	ajo	-	ata	-	ato	-
Dat.	atij	i	asaj	i	atyre	u	atyre	u
Acc.	atë	e	atë	e	ata	i	ato	i

PCIs in Albanian indicate the person and number of the respective objects of the verb. They occur in dative and accusative case: dative pCIs can be combined with accusative pCIs. In most cases this results in *amalgamated forms* (*crasis*). For example *më* and *e* amalgamate to *ma*. The combination of *na* with *e*, *i*, or *u* does not involve amalgamation but *concatenation*, in which case the dative precedes the accusative: *na e*, *na i*, and *na u*.

There are two morphosyntactic types of pCIs: bound and free. *Bound forms* occur within positive (non-negated) imperatives after the verb stem (*enclitic position*). In the plural they appear between the verb stem and the suffix *ni*, cf. the following example:

Example 1. Pronominal clitics as bound forms²

(1)	sill	V _{Stem}	No pCI; Sg;	‘bring’
a.	sille	V _{Stem} +pCl _A	pCl _A , Sg;	‘bring it’
b.	sillma	V _{Stem} +pCl _{D+A}	pCl _{D+A} , Sg;	‘bring me it’
c.	sillni	V _{Stem} +ni	No pCI; Pl;	‘bring’
d.	silleni	V _{Stem} +pCl _A +ni	pCl _A , Pl;	‘bring it’
e.	sillmani	V _{Stem} +pCl _{D+A} +ni	pCl _{D+A} , Pl;	‘bring me it’

In cases with negation particles, pCIs cannot occur as bound forms, e.g. *Mos e sill!* [‘(You: Sg) do not **it** bring!’], i.e. ‘Do not bring it!’ and *Mos e sillni!* [‘(You: Pl) do not **it** bring!’], i.e. ‘Do not bring it!’]. As *free forms*, pCIs always precede the finite verb (*proclitic position*).³ Word order in Albanian is relatively free. Therefore, both subject and objects may appear in front of the verb complex, which allows for a large number of different sentence patterns for a given verb; the order of elements within the verb complex is fixed. If all of these elements occur, they are in the following sequence: negation, future marker/modal verb, subjunctive particle, pCIs, finite and non-finite verb. The subject can be deduced from verb inflection and thus be left out.

PCIs in Albanian appear either in addition to objects in the sentence (*object doubling*) or they replace objects in the sentence, so that the objects themselves can be left out (*object elimination*).

1.2. Object doubling

The following example (2b) shows the doubling of the accusative object:

Example 2. Accusative object doubling⁴

(2) S+V

<i>Ne</i>	<i>shohim.</i>
Subject:	Verb: Itr
N P11	P11 Ind Prs Act Nad
we	see
'We (are able to) see.'	

S+V+O_A

a. <i>Ne</i>	<i>shohim</i>	<i>studentët/ata.</i>
	Verb: Tr	O: A
		A P13 M
we	see	the students/them

S+pCl_A+V+O_A

b. <i>Ne</i>	<i>i</i>	<i>shohim</i>	<i>studentët/ata.</i>
	pCl: A		
	P13		
we	them	see	the students/them
'We see the students/them.'			

The following example shows the doubling of dative (3, 3a, and 3b) and accusative objects (3a and 3c).

Example 3. Object doubling in dative and accusative⁵

(3) S+pCl_D+V+O_D+O_A

<i>Ne</i>	<i>u</i>	<i>dhamë</i>	<i>studentëve/atyre</i>	<i>librat/ato.</i>
Subject:	pCl: u	Verb: Tr	O: D	O: A
N P11	D P13	P13 Ind Aor Act	D P13	A P1 Det
		Nad		
we	them	gave	the students/ them	the books/ them
'We gave the students/them the books/them.'				

S+pCl_{D+A}+V+O_D+O_A

a. <i>Ne</i>	<i>ua</i>	<i>dhamë</i>	<i>studentëve/atyre</i>	<i>librat/ato</i>
	pCl: u+i			
	D P13 +			
	A P13			
we	them + them	gave	the students/ them	the books/ them
'We gave the students/them the books/them.'				

S+pCl _{D+A} +V+O _D				
b.	<i>Ne</i>	<i>ua</i>	<i>dhamë</i>	<i>studentëve/atyre</i>
	we	them + them	gave	the students/ them
	‘We gave the students/them them.’			
S+pCl _{D+A} +V+O _A				
c.	<i>Ne</i>	<i>ua</i>	<i>dhamë</i>	<i>librat/ato</i>
	we	them + them	gave	the books/them
	‘We gave them the books/them.’			

Sentence (3) has two objects, one in dative and one in accusative case, and one pCl in dative case, which doubles the dative object. Whenever the verb has a dative valency, the dative pCl cannot be left out, regardless of the presence of a dative object in the sentence. In sentence (3a), both objects are doubled by means of the respective pCIs. Only one object is doubled by the amalgamated pCl in sentences (3b) and (3c). Object doubling is also possible in imperatives, e.g. *Sillma librin!* [‘Bring+me+it the book!’, ‘Bring me the book!’], i.e. doubling of accusative object.

When the verb has a first or second person accusative or dative object, the pCIs of the first and second person cannot be omitted, e.g. *Studentët të kuptojnë ty.* or *Studentët të kuptojnë.* [‘The students understand you’]. The form **Studentët kuptojnë ty.* is ungrammatical.

1.3. Object elimination

If a pCl occurs, the corresponding object can be left out.⁶ This phenomenon has been called *Objektseliminierung* [object elimination] by Buchholz (1977), and Buchholz and Fiedler (1987). As there is no established English term, we will use the term *object elimination* as a translation of *Objektseliminierung*.⁷ Example sentences where the objects are left out are (3b) and (3c). A sentence consisting only of pCl(s) and verb can be syntactically well-formed, cf. sentences (4a) and (4b) in the following example, where the subject is optional.

Example 4. Object elimination

(4)	*S+V+O _D +O _A				
	* <i>Ne</i>	_	<i>dhamë</i>	<i>studentëve/atyre</i>	<i>librat/ato.</i>
	we		gave	the students/them	the books/ them
	*‘We gave _ the students/them the books/them.’				

S+pCl_{D+A}+V

a.	<i>Ne</i>	<i>ua</i>	<i>dhamë.</i>
	we	them +	gave
		them	
	‘We gave them to them.’		

S+pCl_A+V

b.	<i>Ne</i>	<i>i</i>	<i>shohim</i>
		them	see
	‘We see them.’		

The following patterns⁸ are possible for doubling and elimination of accusative complements: S+V+O_A (object without pCl), S+pCl_A+V+O_A (object doubling) and S+pCl_A+V (object elimination). For dative complements, the following patterns apply: S+pCl_D+V+O_D (object doubling) and S+pCl_D+V (object elimination). As shown in the examples above, these patterns can be combined with each other.

1.4. Pronominal clitics as valency fillers

Sentences (3b-c) and (4a-b) show the ability of pCls to function as *valency fillers*, as they can replace objects. The grammatical information of the pCl is sufficient to fill the valency of the verb; only the lexical content is missing, which has to be recoverable from the linguistic or extralinguistic context. Thus, if pCls occur, objects can be left out without making the sentence ungrammatical.

If objects are left out, the omission of pCls leads to the selection of a different valency pattern. This can (but does not have to) indicate the use of the verb in a different sense, cf. the difference of valency and meaning between sentences (2) and (2a), (2) and (2b), as well as in the following example adapted from Buchholz, Fiedler, and Uhlisch (1993):

Example 5. Different patterns/meanings of the verb *flas*⁹

(5) S+V

Ai	flet.
Subject:	Verb: Itr
N Sg3 M	Sg3 Ind Prs Act Nad
He	speak
‘He speaks.’	

S+pCl _D +V+O _D						
a.	Ai	i	flet	atij.		
		pCl: D Sg3	Verb: Tr	O: D D Sg3 M		
	He	him	scold	him		
	‘He scolds him.’					
S+pCl _D +A+V+O _D +O _A						
b.	Ai	ia	flet	atij	një	libër.
		pCl: D Sg3 + A Sg3				A Sg1 M Undet
	He	him + it	promise	him	one	book
	‘He promises him a book.’					

The role of pCIs in the context of valency is different from the substitution of objects by pronouns, despite the fact that the pCIs are short forms of personal pronouns. Personal pronouns, however, offer a more precise description of case and number (first and second person plural) and gender (third person singular) than pCIs. First of all, the position of the pronoun is the same as the position of the object if a pronoun substitutes the object. A pCI has a fixed position in the verbal complex, regardless of the position of the object it replaces. While pronouns always replace objects, pCIs are capable of either eliminating or doubling the objects they refer to. The appearance of the dative pCI is obligatory while the substitution of the dative object by pronouns is optional. The substitution of objects by pronouns does not change the structure of the sentence (which means the sentence matches to the same pattern), i.e. the O_A is only realized by a pronoun instead of a noun phrase, but the pattern S+V+O_A remains the same. If a pCI appears, however, the pattern is changed, e.g. from S+V+O_A to S+pCl_A+V+O_A or to S+pCl_A+V. As shown above, pCIs and pronouns can (and in some cases must, cf. section 1.2.) occur together in a sentence.

As the dative pCI cannot be left out without making the sentence ungrammatical, it always shares the dative valency slot with the dative object if the object is present. The pCI fills the valency slot itself, if the dative object is eliminated. In the case of the accusative, there is one more option: the object alone fills the valency slot of the verb, the object and pCI share the valency slot, or the pCI alone fills the valency slot. Thus in this analysis, there is a minimum number of required complements of 1 and a maximum number of required complements of 2 for the dative and the accusative valency slot (by analogy to Herbst et al. [2004], where the minimum and maximum valency of each verb is indicated); cf. the patterns S+V+O_A, S+pCl_A+V and S+pCl_D+V for minimum, and S+pCl_A+V+O_A and

S+pCl_D+V+O_D for maximum complements.

The fact that the dative pCl is obligatory even if the dative object is also present can be accounted for using the concept of structural necessity presented by Herbst (1999), which means that it is obligatory only on the syntactic level, not as a lexical property. As pCls can be correctly predicted by a set of rules, they are no idiosyncratic properties of lexical units and thus need not be specified in the valency frame of a lexical unit. It is enough to specify valency slots such as dative or accusative. Object doubling is always optional on the syntactic level; here semantic factors (needed for lexical content) as well as pragmatic factors (emphasis) decide whether doubling takes place or not. Additional functions and properties of pCls will not be discussed here.¹⁰

The properties of the pCl in the context of valency can be summarized as follows: in case of object elimination, the pCls indicate the verb valency and function as valency fillers, whereas in case of object doubling, both pCls and objects share the valency slot of the verb.

2. Practical aspects: The computational model

The formalism used in the following computational model is *Left Associative Grammar* (LAG).¹¹ LAG was developed as a formalism for the SLIM (*Surface compositional Linear Internal Matching*) language theory intended to model and reconstruct natural language communication on a computer.

2.1. The formalism

The formalism works according to the principle of possible continuations whereby every grammar rule concatenates the sentence start read so far with the next word. The result of this concatenation becomes the new sentence start to be concatenated with the next word again; this procedure is repeated until the end of the input is reached. For example the input *a b c d* is parsed in the steps $a+b \rightarrow ab$; $ab+c \rightarrow abc$; $abc+d \rightarrow abcd$, summarized as $((a+b)+c)+d \rightarrow abcd$. Ambiguities are handled by tracing several derivation paths in parallel. The most varied sentence types can be modeled with linear effort regarding complexity and parse time, because only the possible continuations matter in each LAG rule.

2.2. An example

Below, an algorithm for the treatment of dependencies between verb, pCIs and objects is outlined using the sentence *Ai na i dha librat*. [‘He gave us the books’].¹² It starts with the first word form and the matching rule *Subject* (start rule).

Ai_N Sg^3 M

❶ RULE: Subject
 LVF = <Subject>;
 Follow RULE pCl_D ;

The subject is added to the list of valency fillers (LVF). In case of an omitted subject, the algorithm starts with the next rule, i.e. the parser searches for a rule with a matching start pattern. *Follow* means continuations, i.e. the next applicable rule, in this case pCl_D . When reading a word form, information required for recognition of the verb complex and for dealing with valency and congruency is provided by the lexicon or a morphological analysis component.

Ai + $na_{D\ P11}$

❷ RULE: pCl_D
 LVF = <Subject, $pCl_D_Object_elimination$ >;
 Follow RULE pCl_A ;

pCl_D is added to the LVF and labeled *Object_elimination* because the dative object can be left out. Since at this point it is not clear yet whether the actual pattern is *elimination* or *doubling*, the modification of the attribute to $pCl_D_Object_doubling$ can be done when a matching object is encountered. If no matching object is found, the attribute is just left at the value set here. Here, two non-amalgamated (concatenated) pCIs are used to demonstrate the *canceling* (filling) of valencies (valency slots). Amalgamated pCIs have to be analyzed in the morphology component.

$Ai\ na$ + $i_{A\ P13}$

❸ RULE: pCl_A
 LVF = <Subject, $pCl_D_Object_elimination$, $pCl_A_Object_elimination$ >;
 Follow RULE *finVerb*;

pCl_A is added to the LFV just like *pCl_D* in the previous rule, also with the *Object_elimination* label.

Ai na i	+	dha _{Sg3 Aor Ind Act Nad <<D, A>, <A>>}
---------	---	--

④ RULE: *finVerb*
 LVF = <*pCl_D_Object_elimination*, *pCl_A_Object_elimination*>;
 Follow RULE *Accusative_object*;

The verb is read in and checked for agreement with the subject. The subject valency in LVF is canceled, the *pCl*s, however, remain in the LVF because they can share their valency slot with an object and thus cannot be canceled before the respective object has been read or the sentence is finished. Because *pCl*s always precede the verb, a (minimum) valency pattern that selects one or more of the possible lexical readings can be already constructed at this point.

Ai na i dha	+	librat _{A Pl3}
-------------	---	-------------------------

⑤ RULE: *Accusative_object*
 Replace *pCl_A_Object_elimination* by *pCl_A_Object_doubling*;
 LVF = <*pCl_D_Object_elimination*, *pCl_A_Object_doubling*>;
 Follow RULE *Punctuation*;

An accusative object is read in and modifies the *pCl_A_Object_elimination* entry in the LFV to *pCl_A_Object_doubling*.

Ai na i dha librat	+	.
--------------------	---	---

⑥ RULE: *Punctuation*
 LVF = <*pCl_D_Object_elimination*, *pCl_A_Object_doubling*>;
RESULT = *pCl_D_Object_elimination*, *pCl_A_Object_doubling*;

The end of the sentence has been reached, the result is in LVF: object doubling for accusative and object elimination for dative.

Only the path that actually parses this sentence is shown above; other possible paths were left out. After various rule applications, several continuations (parallel paths) would be possible and would have to be tried out by the algorithm. For example, after the rule *finVerb* the sentence may continue with punctuation, dative object, accusative object, a preposition, an adjunct etc.

As demonstrated in the example model, valencies can be canceled immediately when a potential valency filler has been read. Another possibility is end canceling, where the properties of all word forms read are collected in attribute-value matrices and are not canceled until a punctuation mark signals the end of the sentence. This has the advantage of transparency when handling constructions such as subclauses with the verb in final position, but one disadvantage is that ungrammatical constructions may not be rejected before they have been completely parsed.

When reading amalgamated pCl word forms such as *ma*, *ta*, *t'ia* etc., it is important to read and process the individual parts of the morphosyntactic information, e.g. *ma* ($m\ddot{e}_D Sg1 + e_A Sg3$). In this way, an amalgam can replace one object (object elimination) and double another at the same time, cf. sentences (3b) and (3c).

The following figure shows an example of a morphological analysis of the imperative one-word-sentence *sillmani*. As shown under the *Clitic* attribute, the enclitic form *ma* has been recognized as an amalgam of *mē* and *e*.¹³

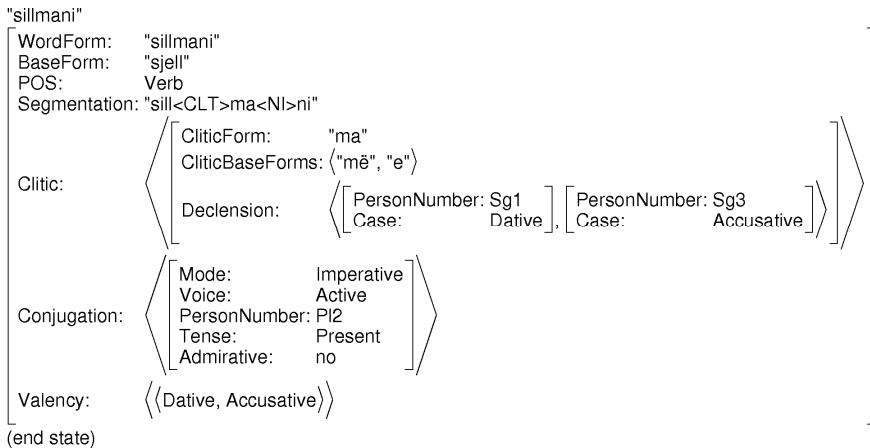


Figure 1. Morphological analysis of the word form *sillmani* from Kabashi (2003)

Here, valencies from the corresponding attribute are canceled with the matching cases from the clitic attribute *Declension*. It is necessary to check the information of lexical entries and morphological analysis to be able to select or construct the correct valency pattern, particularly with regard to the pCl's that may be used with a verb. Information on possible valency patterns of a verb comes from its lemma in the base-form lexicon.

Figure 2 shows the result of syntax analysis and valency handling of the sentence *Ai na i dha librat.* treated above. The dative object is missing there, cf. the *FilledValencyFromObjects* attribute. A dative object was expected according to the verb's valency pattern but was not present. A dative pCl was found and thus the sentence can be analyzed as well-formed, cf. the *FilledValencyFromClitics* and the *DativeSlot* attributes. On the other hand the fact that both an accusative pCl and an accusative object are found, leads to object doubling, cf. the corresponding attributes, *FilledValencyFromObjects* and *AccusativeSlot*. The *Index* attribute indicates the position of a word form in the analyzed sentence. The actual pattern is derived from the attributes *Clitic_D*, *Clitic_A*, *Verb*, and *Object_A*. The *Meaning* attribute contains the meaning of the verb in the currently selected pattern.

"Ai na i dha librat ."															
Subject:	<table border="0"> <tr><td>POS:</td><td>Pronoun</td></tr> <tr><td>Lemma:</td><td>"ai"</td></tr> <tr><td>Declension:</td><td>{ PersonNumber: Sg3 Case: Nominative }</td></tr> <tr><td>Gender:</td><td>Masculine</td></tr> <tr><td>Index:</td><td>1</td></tr> </table>	POS:	Pronoun	Lemma:	"ai"	Declension:	{ PersonNumber: Sg3 Case: Nominative }	Gender:	Masculine	Index:	1				
POS:	Pronoun														
Lemma:	"ai"														
Declension:	{ PersonNumber: Sg3 Case: Nominative }														
Gender:	Masculine														
Index:	1														
Clitic_D:	<table border="0"> <tr><td>POS:</td><td>Clitic</td></tr> <tr><td>Lemma:</td><td>"na"</td></tr> <tr><td>Declension:</td><td>{ PersonNumber: P11 Case: Dative }</td></tr> <tr><td>Index:</td><td>2</td></tr> </table>	POS:	Clitic	Lemma:	"na"	Declension:	{ PersonNumber: P11 Case: Dative }	Index:	2						
POS:	Clitic														
Lemma:	"na"														
Declension:	{ PersonNumber: P11 Case: Dative }														
Index:	2														
Clitic_A:	<table border="0"> <tr><td>POS:</td><td>Clitic</td></tr> <tr><td>Lemma:</td><td>"i"</td></tr> <tr><td>Declension:</td><td>{ PersonNumber: P13 Case: Accusative }</td></tr> <tr><td>Index:</td><td>3</td></tr> </table>	POS:	Clitic	Lemma:	"i"	Declension:	{ PersonNumber: P13 Case: Accusative }	Index:	3						
POS:	Clitic														
Lemma:	"i"														
Declension:	{ PersonNumber: P13 Case: Accusative }														
Index:	3														
Verb:	<table border="0"> <tr><td>WordForm:</td><td>"dha"</td></tr> <tr><td>BaseForm:</td><td>"jep"</td></tr> <tr><td>POS:</td><td>Verb</td></tr> <tr><td>Conjugation:</td><td>{ Mode: Indicative Voice: Active PersonNumber: Sg3 Tense: Aorist Admirative: no }</td></tr> <tr><td>Valency:</td><td>{ (Dative, Accusative) }</td></tr> <tr><td>Meaning:</td><td>{ "to give" }</td></tr> <tr><td>Index:</td><td>4</td></tr> </table>	WordForm:	"dha"	BaseForm:	"jep"	POS:	Verb	Conjugation:	{ Mode: Indicative Voice: Active PersonNumber: Sg3 Tense: Aorist Admirative: no }	Valency:	{ (Dative, Accusative) }	Meaning:	{ "to give" }	Index:	4
WordForm:	"dha"														
BaseForm:	"jep"														
POS:	Verb														
Conjugation:	{ Mode: Indicative Voice: Active PersonNumber: Sg3 Tense: Aorist Admirative: no }														
Valency:	{ (Dative, Accusative) }														
Meaning:	{ "to give" }														
Index:	4														
Object_A:	<table border="0"> <tr><td>WordForm:</td><td>"librat"</td></tr> <tr><td>BaseForm:</td><td>"libër"</td></tr> <tr><td>POS:</td><td>Substantive</td></tr> <tr><td>Declension:</td><td>{ Num: Pl Definite: yes Case: Accusative }</td></tr> <tr><td>Gender:</td><td>Masculine</td></tr> <tr><td>Index:</td><td>5</td></tr> </table>	WordForm:	"librat"	BaseForm:	"libër"	POS:	Substantive	Declension:	{ Num: Pl Definite: yes Case: Accusative }	Gender:	Masculine	Index:	5		
WordForm:	"librat"														
BaseForm:	"libër"														
POS:	Substantive														
Declension:	{ Num: Pl Definite: yes Case: Accusative }														
Gender:	Masculine														
Index:	5														
FilledValencyFromClitics:	{ Dative, Accusative }														
FilledValencyFromObjects:	{ Accusative }														
AccusativeSlot:	Filled_2_pCl_O														
DativeSlot:	Filled_1_pCl														
(end state)															

Figure 2. Syntactical analysis of the sentence *Ai na i dha librat.*

3. Conclusion

Acting as object substitutes (object elimination), pronominal clitics determine the verb's valency and assume the role of valency fillers. They play an important role in distinguishing between various possible valency frames of a verb and thus between different meanings. In the case of object doubling, they merely function as semantic and pragmatic markers for their respective object.

PCIs supply grammatical information that can be very useful both in natural language communication as well as in natural language processing, e.g. for processing discontinued sentences during turn-taking in dialogue analysis.

In spite of the treated phenomenon's complexity, an efficient implementation of a parser is possible using the LAG formalism. As shown in the algorithm, canceling (filling) of verb valencies (valency slots) is easily solved despite the multitude of possible combinations and the consequently large number of sentence patterns.

Notes

1. For comments on the draft of this paper I would like to thank Jörg Kapfer, Matthias Bethke, and Peter Uhrig (all Friedrich-Alexander-Universität Erlangen-Nürnberg). Only the properties relevant to verb valency and computational modeling will be treated here. For further information on pronominal clitics left out here, see Buchholz (1977), Buchholz and Fiedler (1987), Domi (1995; 1997), Kallulli (1995), and Newmark, Hubbard, and Prifti (1982).
2. Here, the following abbreviations are used: A=Accusative, D=Dative, pCl_{D+A}=Amalgam of pCl_D and pCl_A, Pl=Plural, Sg=Singular, and V=Verb. Forms like *silleni* have alternatives in the form *sillnie*.
3. In subjunctive clauses, pCIs are positioned after the subjunctive and future particles and precede the finite verb. In this case the pCl can be combined with these particles in one word, e.g. the amalgam *t'i* consisting of the subjunctive particle *të* and the pCl *i*.
4. Act=Active, Ind=Indicative, Itr=Intransitive, M=Masculine, Nad=Not admira-tive, N=Nominative, O=Object, Prs=Present, S=Subject, and Tr=Transitive.
5. Aor=Aorist (definite past), and Det=Determined.
6. An exception is the reflexive use of verbs, e.g. *Ai e_(pCl: A) lavdëron vetën_(O: A, Reflexive)* ['He (**him**) praised **himself**.', i.e. 'He praised himself.'], vs. *Ai e lavdëron*. ['He (**him**) praised.', i.e. 'He praised him.'].
 Reflexive) ['He (**him**) praised himself.', i.e. 'He praised himself.'], vs. *Ai e lavdëron*. ['He (**him**) praised.', i.e. 'He praised him.'].
7. This term is described in Buchholz and Fiedler (1987) as *Vertretung des Objekts* (which might be translated as *object replacement* or *object substitution*).

8. As we focus on the influence of pronominal clitics on verb valency, only direct and indirect objects (without prepositions) are treated here.
9. Undet=Undetermined. Other patterns of the verb *flas*, e.g. pCl(s) + *flas* + preposition, and other meanings not presented in Buchholz, Fiedler, and Uhlisch (1993) are not treated here. For more patterns/meanings cf. Kostallari (1980), where the pattern/meaning from sentence (5b) is marked as a dialect form. Cf. also Qesku (1999). Toçi (2002) does not list this pattern.
10. Buchholz and Fiedler (1987: 445–446) have described a group of verbs which cooccur with „„pleonastisch‘ verwendete Objektszeichen“ [“pleonastically used” pronominal clitics]. These pCIs are „nicht ... systematisch aus einem zur Grundstruktur gehörenden dir[ekten] Obj[ekt] ableitbar“ [not derived in a systematic way from a direct object belonging to the base structure]. Thus sentences in which only one dative pCI would be expected can have two pCIs or an amalgamated one consisting of a dative and an accusative pCI, like for example *Ia*_(pCI: D+A) *hipi*_(V) *kalit*_(O: D). [‘He mounted the horse’]; in this example the subject is left out. These verbs cooccur only with specific lexical entries (and so they must be marked in the lexicon to be considered during automatic syntactic and semantic analysis). According to Buchholz and Fiedler (1987: 445), this group contains, among others, the following verbs: *arrin, del, fillon, hipën, hyn, kërcen, mbath, merr, nis, pëlçet, shtron, and thotë*.

PCIs can also occur as ethical datives, cf. Buchholz and Fiedler (1987: 447–448). In this case, the verb valency must be treated differently as well.

11. For the formal definition of LAG see Hausser (1992; 2001a).
12. The version illustrated here is only a simple LAG. Information about valency in *Database Semantics (DBS)* can be found in the article *Handling valency and coordination in database semantics* by R. Hausser in this volume. For information on DBS see Hausser (2001b).
13. During the syntactic analysis both pCIs would fill the corresponding attribute in the sentence structure while reading the verb.

References

- Buchholz, Oda
 1977 *Zur Verdoppelung der Objekte im Albanischen*. Linguistische Studien, Reihe A, Arbeitsberichte 34. Berlin: Akademie der Wissenschaften der DDR.
- Buchholz, Oda, and Wilfried Fiedler
 1987 *Albanische Grammatik*. Leipzig: VEB.
- Buchholz, Oda, Wilfried Fiedler, and Gerda Uhlisch
 1993 *Wörterbuch Albanisch–Deutsch*. München: Langenscheidt.
- Domi, Mahir (ed.)
 1995 *Gramatika e gjuhës shqipe*. Vëllimi I – *Morfologjia* [Grammar of the Albanian Language. Vol. 1: Morphology]. Tiranë: Akademia e Shkencave e Republikës së Shqipërisë.

- 1997 *Gramatika e gjuhës shqipe*. Vëllimi II – *Sintaksa* [Grammar of the Albanian Language. Vol. 2: Syntax]. Tiranë: Akademia e Shkencave e Republikës së Shqipërisë.
- Hausser, Roland
- 1992 Complexity in left-associative grammar. In *Theoretical Computer Science*. 106 (2): 283–308.
- 2001a *Foundations of Computational Linguistics. Human-Computer Communication in Natural Language*. 2d ed. Berlin/New York: Springer.
- 2001b Database semantics for natural language. In *Artificial Intelligence* 130: 27–74.
- 2007 Handling valency and coordination in database semantics. This volume.
- Herbst, Thomas
- 1999 English valency structures – A first sketch. *Erfurt Electronic Studies in English* (EESE).
http://webdoc.gwdg.de/edoc/ia/eese/artic99/herbst/6_99.html.
- Herbst, Thomas, David Heath, Ian F. Roe, and Dieter Götz (eds.)
- 2004 *A Valency Dictionary of English. A Corpus-Based Analysis of the Complementatation Patterns of English Verbs, Nouns and Adjectives*. Berlin/New York: Mouton de Gruyter.
- Kabashi, Besim
- 2003 Automatische Wortformererkennung für das Albanische. Master's thesis, Computational Linguistics. Universität Erlangen-Nürnberg.
- Kallulli, Dalina
- 1995 *Clitics in Albanian*. (Working Papers in Linguistics 24.) Trondheim: University of Trondheim.
- Kostallari, Androkli (ed.)
- 1980 *Fjalor i gjuhës së sotme shqipe* [Dictionary of Contemporary Albanian Language]. Tiranë: Akademia e Shkencave e RPS të Shqipërisë.
- Newmark, Leonard, Philipp Hubbard, and Peter Prifti
- 1982 *Standard Albanian. A Reference Grammar for Students*. Stanford: Stanford University Press.
- Qesku, Pavli
- 1999 *Fjalor Shqip–Anglisht. Albanian–English Dictionary*. Tiranë: EDFA.
- Toçi, Fatmir (ed.)
- 2002 *Fjalor i shqipes së sotme* [Dictionary of Contemporary Albanian]. 2d ed. Tiranë: Toena.